ESSAYS ON LIQUIDITY RISK AND ASSET PRICING

A Dissertation

Presented to

The Faculty of the C.T. Bauer College of Business

University of Houston

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

Anandi Banerjee

December, 2016

ESSAYS ON LIQUIDITY RISK AND ASSET PRICING

APPROVED:

Thomas J. George, Bauer Professor of Finance
Chairperson of Committee

Kris Jacobs, Bauer Professor of Finance

Beau Page, Assistant Professor of Finance

Seshadri Tirunillai, Assistant Professor of Marketing

Latha Ramchand, Dean
C.T. Bauer College of Business

DEDICATION

To my Dad

ACKNOWLEDGEMENTS

I am deeply indebted to my advisor, Dr. Thomas George for his fundamental role in my doctoral work. He has mentored me since the first year of the Ph.D. program and this dissertation would have not been possible without his guidance. He has inspired and encouraged me during difficult times and truly has been a blessing in my life. I doubt that I will ever be able to convey my appreciation fully, but I owe him my eternal gratitude.

I would also like to thank Dr. Beau Page who has helped me complete this dissertation. I am grateful to him for devoting so much time, sharing his profound knowledge and deepening my understanding about concepts that were integral to my research. He has supported me every step of the way, and has been a friend, philosopher and guide to me.

I am grateful to Dr. Kris Jacobs for his invaluable counsel. His insightful comments and feedback has helped me widen my research from various perspectives. His advice on my research as well as on my career has been priceless. I would also like to thank Dr. Hitesh Doshi for his caring guidance and encouraging words. I am grateful to Dr. Rauli Susmel for his valuable suggestions during my presentations. I would like to thank Dr. Darla Chisholm for being a friend and a confidant.

I thank my family members for their unconditional love and support. I am grateful to my father for believing in me. Special thanks to my mother for dedicating so many years of her life to make things easier for me, to fulfill my duties, and of course, for taking care of the light of my life, Aahnik. I am grateful to Maharajji for being a source of strength to me during tough times. Last, but not the least, I thank my son, Aahnik Vrishank for filling my life with unadulterated joy and optimism.

ESSAYS ON LIQUIDITY RISK AND ASSET PRICING

Abstract

Anandi Banerjee

December, 2016

This dissertation consists of two essays on liquidity risk and asset pricing. In the first essay, I diagnose the impact of error-in-variables (EIV) on inferences in asset pricing models. I test the CAPM and the liquidity-adjusted CAPM in a manner that explicitly accounts for EIV, without pooling stocks into portfolios. I find that the single-factor CAPM beta is not priced. I document that the aggregate liquidity risk in the liquidity-adjusted CAPM of Acharya and Pedersen (2005) is priced, and the portfolio-based approach is unable to capture this relationship. The cumulant-based approach used in my paper to handle EIV enables me to test the effects of the individual components of aggregate liquidity risk, and I find that the risk associated with the commonality in illiquidity has a positive premium and the risk associated with the sensitivity of a stock's illiquidity to the value-weighted market return has a negative premium. I also show that for microcap stocks, the risk attributable to the covariance between stock return and market-wide liquidity has a negative relationship with average returns. I find that the LCAPM cannot be rejected when the betas are estimated at the stock-level, and the intercept of the model is insignificant.

In the second essay, I explore the relation between idiosyncratic volatility and the cross-section of expected returns. I use an EGARCH model to estimate the forecasted idiosyncratic volatility (FIVOL) and find that this estimate is not affected by the microstructure biases embodied by bid-ask spreads and the percentage of zero returns. I document a positive relation between FIVOL and expected returns. However, contrary to the models in the existing literature (such as Merton (1987)), I find that the cross-sectional differences in levels of idiosyncratic volatility are not priced. The positive relation is mainly driven by stocks that rise in their FIVOL quintile ranking. These transitions in FIVOL ranking are a consequence of return shocks that result in the sudden changes in FIVOL. I explore earnings surprises as a potential explanation for these return shocks and find that standardized unexpected earnings cannot completely explain the pricing ability of these transitions in FIVOL. Even after controlling for earnings surprises, I find that the stocks that move from a low FIVOL quintile to a higher quintile earn high returns.

Table of Contents

List of Tables

Chapter 1

Error-in-Variables and Tests of Asset Pricing Models with Liquidity Risk

1.1 Introduction

True security betas are unobservable. However, accurate measurement of betas is of paramount importance for evaluating the performance of asset pricing models in explaining the cross-section of average returns. In the traditional two-pass estimation methodology, which is usually used to test beta-pricing models, the betas estimated from the first stage are used as the explanatory variables in the second stage. Estimation error in the first stage leads to error-in-variables (EIV) in the second pass.

The EIV problem is widespread in economic statistics and distorts inferences if it is not taken into account explicitly. In a linear regression model, with one mismeasured regressor and an intercept, measurement error biases the slope coefficient towards zero. This is known as the attenuation effect due to EIV. This error may also have an effect in the opposite direction on the intercept, and bias the intercept away from zero. This is a large issue in asset pricing, where finding a significant intercept means the model is a failure.

In a model with one mismeasured regressor, EIV does not change the sign on the coefficient estimate. This does not always hold true when more than one regressor is affected by EIV. In a model with many mismeasured regressors, the slope coefficient on a mismeasured regressor is affected not only by its own measurement error, but also by the measurement error in the other mismeasured regressors. The second effect, which may bias the coefficients away from zero, is known as the contamination effect and is severe if the measurement errors are correlated. Because of the two opposing effects, all the coefficient estimates in a multiple regression are inconsistent, and no definite conclusion can be drawn about the direction of the bias or whether the coefficients are significantly different from zero.

EIV correction is of utmost importance in multivariate models plagued with measurement error. In asset pricing tests assessing the role of beta in explaining expected returns, researchers recognize this problem and are willing to trade-off power for precision. Therefore, in most of these tests, portfolios are used instead of individual stocks to address measurement error. It is argued that grouping stocks into portfolios diversifies away the estimation error in the betas. However, this approach has its own drawbacks. Ang et al.(2010) show that aggregating stocks into portfolios decreases the cross-sectional dispersion of betas. Thus using portfolios as test assets lowers the efficiency of the tests due to the inherent loss of information. Liang (2000) contends that if the sorting variable used to form portfolios is also measured with error, then this approach biases the results. Lewellen et. al (2010) argue that the portfolios commonly used by researchers in the tests of asset pricing models bias these tests towards accepting the model. These papers highlight the need for a viable alternate technique to address EIV in the betas in asset pricing tests.

The primary implication of the single-factor capital asset pricing model (CAPM) of Sharpe (1964), Lintner (1965), and Black (1972) suggests that market beta, which captures the return covariance of an individual security with the return on the market portfolio, should be priced. Nevertheless, prior research has suggested that market betas have limited ability to explain asset returns. Despite its repeated failure to explain the cross-section of expected stock returns, the popularity of the CAPM persists because of its intuitive appeal and simplicity. Since the beta estimates in CAPM are measured with error, it is crucial to determine whether the failure of the market beta to explain average stock returns is a result of model misspecification or the attenuation effect of EIV.

Liquidity is another attribute that could affect an investment's required return. It describes the ease of buying or selling an asset without affecting its price. It is a multi-dimensional concept that is difficult to measure. Empirical studies have used various proxies for liquidity. Some of the different measures of liquidity used in literature, such as bid-ask spreads (Amihud and Mendelson, 1986), price-impact (Brennan and Subrahanyam, 1996; Amihud, 2002), and share turnover (Brennan, Chordia and Subrahmanyam, 1998) have been shown to affect asset returns. While each of these proxies likely captures some aspect of the theoretical concept of "liquidity", none of them embeds their proxy explicitly into an asset pricing model.

In contrast, Acharya and Pedersen (2005), henceforth referred to as AP2005, propose a liquidity-adjusted capital asset pricing model (LCAPM), which is a generalization of the CAPM in an economy with trading frictions. In the LCAPM, the "net market risk" is given by the sensitivity of stock returns net of its illiquidity costs to the market return net of the market illiquidity cost. This "net market risk" can be decomposed into four components, three of which are related to stock and market-wide liquidity. AP2005 start with a simple theoretical model based on the CAPM and show that the overall risk in a market with trading costs can be attributed to the beta due to the sensitivity of stock returns to fluctuations in market return, and three liquidity betas stemming from the commonality in liquidity, the sensitivity of stock returns to fluctuations in market-wide liquidity, and the sensitivity of stock illiquidity to fluctuations in market return. Similar to the market beta, these liquidity betas are unobservable. They test the LCAPM using a two-pass procedure, where liquidity betas are estimated in the first stage from the moments of past returns and illiquidity. These beta estimates are used in the second stage regressions. Hence, their test of LCAPM suffers from EIV.

This paper tests the CAPM and the LCAPM in a manner that explicitly accounts for EIV, without pooling stocks into portfolios. I employ a unique method based on cumulants developed in Geary (1942) to deal with EIV in asset pricing tests. One of the main advantages of this methodology is that it enables me to use individual stocks as test assets, and eliminates the need to form portfolios. I use the information contained in the third and higher order cumulants of the joint distribution of the observable variables. This technique has previously been used by Erickson, Jiang and Whited (2014), henceforth referred to as EJW2014, to address EIV in areas of corporate finance. Their paper developed a convenient two-step minimum distance estimator with a simple closed-form solution. The applications covered in their paper dealt with a maximum of two mismeasured regressors, whereas this paper will deal with up to five mismeasured regressors. I use simulated data to highlight the importance of accounting for measurement error in a model with multiple mismeasured regressors. A comparison of small sample performance of cumulant-based estimators with OLS is also presented, which highlights the advantages of using the cumulant approach.

I estimate the CAPM with portfolio as well as stock-level betas. I find that the relation between expected returns and portfolio beta is flat. Then I estimate the market beta for individual stocks and use the cumulant-based approach to correct for EIV. Even after correcting for measurement error in the market beta, I find that the CAPM is still misspecified.

Next, I present the empirical results of tests of the LCAPM. Previous papers have estimated the betas in the LCAPM against an equal-weighted market portfolio. But, the CAPM states that the value-weighted portfolio of all stocks in the economy should be the tangency portfolio. This motivates the estimation of betas in my paper based on a value-weighted instead of an equal-weighted market portfolio. Using the more theoretically appropriate regressor leads to results that are stronger, more intuitive and more consistent than results found using the equal-weighted market portfolio.

First, I group stocks and estimate betas at the portfolio level, as done in previous papers, and find that liquidity risk is not priced. But grouping stocks not only diversifies away the estimation error, but also diversifies away information contained in individual stock-level betas. Next, I estimate the betas for individual stocks and address the inherent EIV in these stock-level betas using the cumulant-based approach. Using these stock-level betas, I find that liquidity risk is priced and the LCAPM cannot be rejected. This is the first paper to document the positive relation between aggregate liquidity risk and expected returns after accounting for the estimation error in the stock-level betas.

This aggregate liquidity beta can be decomposed into three separate betas arising from different components of a stock's exposure to systematic risk. AP2005 find that the three liquidity betas are highly correlated with each other and a model estimated with the individual betas suffers from severe multicollinearity. Thus the existing papers have not been able to disentangle the effects of these betas on stock returns. However, when I estimate the betas at the stock level, based on a value-weighted market portfolio, I find that the correlations between the betas are quite low. Therefore, another advantage of the technique used in my paper is that it allows me to disaggregate the liquidity beta into its individual components. Thus the statistical identification of the separate effects of different liquidity risks is possible using the cumulant-based approach. I find that the return premium due to the risk associated with the commonality in illiquidity is positive and significant. I also find that the risk due to the covariance between an asset's illiquidity and the market return is negatively priced. To the best of my knowledge, this is the first paper to document the relationship between individual liquidity risks and expected returns after controlling for measurement error explicitly.

Next, I investigate whether the pricing ability of liquidity risk is driven by microcap stocks. Fama and French (2008) define the microcaps as stocks whose market capitalization is below the 20^{th} NYSE percentile. They show that though these stocks account for only 3% of the market capitalization of the NYSE-Amex-NASDAQ universe, they account for about 60% of the stocks. Hou et al.(2015) note that the microcap stocks are highly illiquid and have higher transaction costs. I find that the coefficient on liquidity beta is much smaller when we exclude the microcaps from the sample. I also find evidence of a negative premium due to the risk associated with the return sensitivity to market illiquidity.

The remainder of the paper is structured as follows. Section 1.2 reviews the existing literature. Section 1.3 explains the classical EIV model and how it affects returns. It also describes the EJW2014 method to handle the bias due to EIV. Section 2.2 discusses the data and methodology employed to estimate the LCAPM. Section 1.5 uses simulations to compare the small-sample performance of OLS and higher order cumulants. Section 2.3 describes the empirical applications and results of the tests of the CAPM and the LCAPM.

Section 1.7 conducts robustness checks. Section 2.6 concludes.

1.2 Relevant Literature

This paper forms a nexus between two different lines of research, namely, the literature on asset pricing models that seek to explain the relationship between risk and expected returns of securities and the literature on measurement error and techniques adopted to get consistent estimates in error-laden models.

The first strand of literature explores whether systematic risk is priced. It takes us back to the seminal question in asset pricing that asks whether differences in exposure to market-wide risk factors can explain the differences in the expected returns of assets. A variety of asset pricing models have been proposed in the literature to understand why different assets earn different rates of return and if it can be attributable to the difference in their sensitivities to systematic risk. The factors used in these models to explain returns are often different.

According to the single-factor CAPM, the return on the market portfolio is the only source of non-diversifiable risk and an asset's exposure to this factor determines its expected return. However, Reinganum (1981), Lakonishok and Shapiro (1986) and Fama and French (1992) find that a relation between expected returns and market beta does not exist in the 1963-1990 period.

Richer models have been proposed that use other economic variables as systematic risk factors. Unanticipated changes in the term premium, default premium, the growth rate of industrial production and inflation constitute the factors in Chen, Roll and Ross (1986). Jagannathan and Wang (1996) extend the proxy for market return to include the return on human capital as a factor. Firm characteristics have also been used to create factors that affect expected returns. Fama and French (1993) show that a three-factor model, consisting of the market factor, a factor based on the market value of equity (small-minus-big, SMB), and a factor based on the book-to-market equity (high-minus-low, HML), jointly do a reasonable job at explaining the cross section of stock returns. However, post 1990s, the model's performance deteriorates. The three-factor model fails to account for the significant alphas generated by momentum strategies. Carhart (1997) introduces a momentum factor, which is based on the prior returns (high prior returns-minus-low prior returns, MOM), as an extension to the Fama-French three-factor model and he finds that short-term persistence in equity mutual fund returns can be explained by the MOM factor.

Most of these models assume frictionless markets. But, in reality there are no truly frictionless markets since trading is always associated with certain costs or restraints. Hence researchers have incorporated the effect of trading frictions into factor models. Pastor and Stambaugh (2003) measure illiquidity as the return reversal in response to volume shocks and examine whether market-wide liquidity is a priced factor. They find that the innovations in market illiquidity is a priced factor that is related to the average returns of an asset. The LCAPM proposed by AP2005 provides a unified framework that encompasses the different channels via which liquidity affects stock returns. In the LCAPM, the expected return of a security depends on its expected illiquidity level, market risk and three liquidity risks. But, as mentioned earlier, it is difficult to investigate the pricing ability of the liquidity risks because they are highly correlated with each other. The results in AP2005 are not robust to various portfolio formation techniques and not consistent across different model specifications. I find that the incongruity is a consequence of the measurement error embedded in these liquidity-based models. The measurement error in these models stem from the error in the liquidity estimate as well as the estimation error in beta estimates. Thus it is crucial to address EIV in tests of the LCAPM.

The second strand of literature studies the effect of measurement error in economic models and explores methods to overcome the EIV problem. Papers in corporate finance often use proxy variables when the independent variables in a predicted relationship are not observable. Traditionally, instrumental variables have been used to address the bias due to measurement error. However, it is often difficult to find a good instrument that is correlated with the mismeasured variable, but uncorrelated with any other determinant of the dependent variable. Erickson and Whited (2000, 2002), henceforth referred to as

EW2000 and EW2002 respectively, circumvent the need to find suitable instruments in such models. They develop consistent estimators using information in third and higher order moments of the observable variables. Nevertheless, the estimating equations used in these papers are complicated non-linear functions of the parameters to be estimated, and estimation is sensitive to the starting values used in the numerical optimization process, as local optima may exist.

Geary (1942) proposes a method to derive the true relationship between variables, when the independent variables are measured with error. He develops a system of estimating equations using high order cumulants to determine the true underlying relation between the variables. Building on Geary, EJW2014 propose simple linear estimating equations with a closed-form solution using high-order cumulants, instead of moments. Furthermore, their paper extends Geary's results by employing minimum distance estimation to solve models which have overidentifying cumulant equations.

In asset pricing papers that seek to estimate the price of risk in linear factor models, the most common approach used to address EIV is to form diversified portfolios based on common characteristics. Fama and MacBeth (1973) show that this method reduces the estimation error in the individual beta estimates. However, the portfolio grouping method may conceal information that exists in individual stocks. Roll (1977) proposes that mispricing in individual assets can be diversified away in portfolios. Shanken (1992) suggests a correction factor under the assumption of conditional homoscedasticity to account for the estimation error in betas. Kim and Skoulakis (2015) use a regression-calibration method to correct the betas estimated in the first pass in the two-pass cross-sectional regression (CSR) method. A correction factor is used to calibrate the betas. These calibrated betas, which are used in the second-pass, satisfy the orthogonality conditions necessary for N-consistency. The authors estimate risk premia using individual stock-level data and over short time horizons. They develop an entirely new set of asymptotic results specialized for their regression-calibration approach.

The approach in this paper is much simpler. I use the higher order cumulants technique introduced by Geary (1942) to tackle the bias and possible inconsistency in asset pricing tests due to measurement error. This method eliminates the need to form portfolios, and allows me to use individual stock level data. It also relies on standard asymptotic distribution theory. EJW2014 use a maximum of two mismeasured regressors in their applications. I apply the EJW2014 methodology to asset pricing models which have greater numbers of mismeasured regressors.

1.3 Errors in Variables and Estimated Betas

A common procedure for investigating the relation between betas and expected asset returns is based on Fama and MacBeth (1973), henceforth referred to as FM1973. It involves two steps. For the single-factor CAPM, risk factor betas for individual stocks are estimated in the first step from time-series regressions given by

$$R_{i,t} = \alpha_i + \beta_i F_{m,t} + \nu_{i,t} \tag{1.1}$$

where $F_{m,t}$ is the market realization for month t, and β_i is the beta for stock (or portfolio) i. The second step estimates risk premia from monthly cross-sectional regressions of returns on the beta estimates obtained from the first step

$$R_{i,t} = \rho_t^0 + \lambda_t^1 \widehat{\beta}_i + u_{i,t} \tag{1.2}$$

where λ_t^1 is the market factor risk premium and $\widehat{\beta}_i$ is the estimated beta of each stock (or portfolio) from the first pass. A time series of $\widehat{\lambda_t^1}$ is obtained from monthly estimates of Equation 2.2 and the sample mean of this distribution is used as the final estimate λ^1. The standard error of the estimate is based on the assumption that $\widehat{\lambda_t^1}$ estimates are independent and identically distributed. Inferences are drawn based on these values.

This method is easy to implement and has become a standard methodology in the finance literature. The true beta, β, is unobservable and the estimated beta, $\widehat{\beta}$, from the first pass is used as the independent variable in the second pass cross-sectional regression. The difference between β and $\widehat{\beta}$ constitutes the measurement error in the model. Hence this method suffers from EIV. The next subsection describes a model with EIV and how to get consistent estimates in such a model.

1.3.1 Classical Error-in-variables problem

Let $(y_{i,t}, x_{i,t}, z_{i,t}), i = 1, ..., n, t = 1, ...T$, be sequences of observable variables. $(u_{i,t}, \epsilon_{it}, \chi_{it})$ are sequences of unobservable variables. The classical EIV model is described by

$$y_{i,t} = z_{i,t}\rho + \chi_{i,t}\lambda + u_{i,t} \tag{1.3}$$

$$x_{i,t} = \chi_{i,t} + \epsilon_{i,t} \tag{1.4}$$

where λ and ρ are the unknown vectors. Regressing $y_{i,t}$ on $x_{i,t}$ gives inconsistent estimates of λ. This issue is categorically addressed in EW2000, EW2002 and EJW2014, which develop consistent estimators for models plagued with EIV. They use estimating equations involving higher order moments (EW2000, EW2002) and higher order cumulants (EJW2014), which are functions of moments, in order to estimate consistently the slope parameters in Equation 2.3. The next subsection gives a brief summary of cumulants and the advantages of using cumulant based estimators.

1.3.2 Cumulant estimators

Cumulants were first studied in 1889 by T.N.Thiele, who had termed them as semi-invariants. They are polynomial functions of moments. Cumulants κ_r of a random variable X are defined by a cumulant-generating function $K(\xi)$, which is the natural logarithm of the moment-generating function $M(\xi)$. κ_r are the coeficients in the Taylor expansion of the cumulant generating function about the origin, and are given by

$$K_x(\xi) \equiv lnM(\xi) = lnE[e^{X\xi}] = \sum_{r=0}^{\infty} \kappa_r \xi^r / r! \tag{1.5}$$

where $\kappa_0 = 0$, $\kappa_1 = E[X]$, $\kappa_2 = E[X^2] - E[X]^2$, $\kappa_3 = E[X^3] - 3E[X^2]E[X] + 2E[X]^3$.

A property of cumulants that makes them useful is additivity - the rth cumulant of the sum of two independent random variables equals the sum of the rth cumulant of the individual variables. This property also makes cumulants an attractive choice in estimating models with measurement error.

EJW2014 exploits properties of cumulants to derive two-step minimum distance estimators. To apply this method, the variables in Equations 2.3 and 2.4 should satisfy the following assumptions: (i) the elements of $(u_{i,t}, \chi_{i,t}, z_{i,t})$ should have finite moments of every order, (ii) $(u_{i,t}, \epsilon_{i,t})$ should be independent of $(z_{i,t}, \chi_{i,t})$, (iii) the elements of $(u_{i,t}, \epsilon_{i,t})$ should be independent of each other, (iv) $E(u_{i,t}) = 0$ and $E(\epsilon_{i,t}) = 0$, (v) $E[(\chi_{it}, z_{i,t})'(\chi_{it}, z_{i,t})]$ should be positive definite.

EJW2014 use the relations between cumulants to form an over-identified system of estimating equations. Using a minimum distance estimator to efficiently combine information from the high order cumulants, they solve the equations for λ. In a model with mismeasured and perfectly measured regressors, as described in Equations 2.3 and 2.4, the perfectly measured regressors are first partialled out and the system is expressed in terms of regression residuals. Population linear regression of $x_{i,t}$ on $z_{i,t}$ yields the residual $x_{i,t} - z_{i,t}\mu_x$, where $\mu_x \equiv [E({z_{it}}'z_{it})]^{-1}[E({z_{it}}'x_{it})]$. Linear regression of $\chi_{i,t}$ on $z_{i,t}$ yields the residual $\chi_{i,t} - z_{i,t}\mu_\chi$. But since $z_{i,t}$ and $\epsilon_{i,t}$ are independent of each other,

$$\mu_x \equiv [E({z_{it}}'z_{it})]^{-1}[E({z_{it}}'(\chi_{it} + \epsilon_{i,t})] = [E({z_{it}}'z_{it})]^{-1}[E({z_{it}}'\chi_{it})] \equiv \mu_\chi \tag{1.6}$$

The residual $\chi_{i,t} - z_{i,t}\mu_x$ is denoted by $\eta_{i,t}$.

Subtracting $z_{i,t}\mu_x$ from both sides of Equation 2.4 yields $\dot{x}$ which is defined as

$$\dot{x} \equiv x_{i,t} - z_{i,t}\mu_x = \chi_{i,t} - z_{i,t}\mu_x + \epsilon_{i,t} = \eta_{i,t} + \epsilon_{i,t} \tag{1.7}$$

The residual of population linear regression of y_i on z_i is $y_i - z_i\mu_y$, where $\mu_y \equiv [E({z_{it}}'z_{it})]^{-1}[E({z_{it}}'y_{it})]$. Subtracting $z_{i,t}\mu_y$ from both sides of Equation 2.3 yields $\dot{y}$ which is defined as

$$\dot{y} \equiv y_{i,t} - z_{i,t}\mu_y = \eta_{i,t}\lambda + u_{i,t} \tag{1.8}$$

In the two-step minimum distance estimation, the first step comprises of substituting the least square estimates of $\widehat{\mu_x}$ and $\widehat{\mu_y}$ in Equations 2.7 and 2.8 respectively, as they are consistent estimates of μ_x and μ_y. The second step estimates λ using the sample cumulants of $y_i - z_i\widehat{\mu_y}$ and $x_i - z_i\widehat{\mu_x}$. The λ estimation approach is based on the results derived in Geary(1942). EJW2014 show that the system of cumulant based estimating equations is identified if $\eta_{i,t}$ is skewed. Thus the third moment, and thereby the third order cumulant of $\eta_{i,t}$ should be non-zero.

Consider a model with J mismeasured regressors. If $\kappa(s_0, s_1, ..., s_J)$ is the multivariate cumulant of order s_0 in $\dot{y}$, order s_1 in $\dot{x_1}$ and order s_J in $\dot{x_J}$, then Geary proves that

$$\kappa(s_0+1, s_1, s_2, ..., s_J) = \lambda_1\kappa(s_0, s_1+1, s_2, ..., s_J) + ... + \lambda_J\kappa(s_0, s_1, s_2, ..., s_J+1) \quad (1.9)$$

as long as at least two elements in $(s_0, s_1, s_2, ..., s_J)$ are different from zero. The system of equations based on Equation 2.9 can be represented by

$$K_y = K_x\lambda \quad (1.10)$$

J independent equations can identify λ. If the number of equations in the system, M, is less than J, then λ is indeterminate. If $M > J$, then the system is overidentified and EJW2014 discusses a method of estimating λ from minimum distance estimation of

$$\hat{\lambda} \equiv argmin_l(\hat{K_y} - \hat{K_x}l)'\hat{W}(\hat{K_y} - \hat{K_x}l) \quad (1.11)$$

where $\hat{W}$ is a symmetric positive definite weighting matrix.

The cumulants in Equation 1.10 can be obtained from the moments of the distribution of the observable variables. Cumulants can be expressed as the sums of products of moments. Chapter 2 of McCullagh (1987) gives an expression for any cumulant of a distribution as a function of the moments of the distribution. An example of the relationship between third order cumulants and moments involving four random variables is given by:

$$
\begin{aligned}
\kappa(1,1,1,0) &\equiv E(\dot{y}\dot{x_1}\dot{x_2}) - E(\dot{y}\dot{x_1})E(\dot{x_2}) - E(\dot{y}\dot{x_2})E(\dot{x_1}) - E(\dot{x_2}\dot{x_3})E(\dot{y}) + 2E(\dot{y})E(\dot{x_1})E(\dot{x_2}) \\
&\equiv E(\dot{y}\dot{x_1}\dot{x_2})
\end{aligned}
\tag{1.12}
$$

where $\kappa(1,1,1,0)$ is a third order cumulant of degree 1 in $\dot{y}$, degree 1 in $\dot{x_1}$, degree 1 in $\dot{x_2}$ and degree 0 in $\dot{x_3}$. Thus the matrices $(\hat{K_y}, \hat{K_x})$ can be estimated from the sample moments of $\widehat{\dot{y}} \equiv y_{i,t} - z_{i,t}\widehat{\mu_y}$ and $\widehat{\dot{x}} \equiv x_{i,t} - z_i\widehat{\mu_x}$.

In this paper, I consider models with one, three and five mismeasured regressors. In a model with three mismeasured regressors, χ contains three elements. Examples of third order cumulant estimating equations represented by Equation 1.10 are:

$$K(2,1,0,0) = \lambda_1 K(1,2,0,0) + \lambda_2 K(1,1,1,0) + \lambda_3 K(1,1,0,1) \tag{1.13}$$

$$K(2,0,0,1) = \lambda_1 K(1,1,0,1) + \lambda_2 K(1,0,1,1) + \lambda_3 K(1,0,0,2) \tag{1.14}$$

With three mismeasured regressors, we have 18, 66 and 159 estimating equations using cumulants up to degree three, four and five respectively. Hence, the model is always overidentified, and the minimum distance estimator of λ in Equation 1.11 can be obtained from

$$\hat{\lambda} = (\hat{K_x}'\hat{W}\hat{K_x})'\hat{K_x}'\hat{W}\hat{K_y} \tag{1.15}$$

I use this method to obtain consistent estimates in tests of asset pricing models with mismeasured regressors. The cumulant based estimating equations have a closed-form solution, which eliminates the need to find suitable starting values and iterating to a numerical minimization of the objective function given in Equation 1.11. As demonstrated in EW2012, performance of moment estimators is highly sensitive to the selection of starting values when numerical optimization is used. This is one of the main advantages of using cumulant over moment based estimators.

An increase in the order of cumulants substantially increases the number of estimating equations. However, it does not necessarily improve the accuracy of the estimation. The performance of different orders of cumulant estimators can be compared by comparing the percentage of cumulants that are statistically different from zero. The next section describes the data and beta estimation methodology used in this paper.

1.4 Data and Methodology

1.4.1 Data

I use daily return and volume data from January 1st, 1963 until December 31st, 2014 for all common stocks listed on NYSE, AMEX and NASDAQ, available from CRSP. Accounting information is obtained from Compustat Annual and Quarterly Fundamental Files. A firm's book-to-market ratio is computed by dividing its last fiscal year-end book value by its fiscal year-end market equity. Stocks with share prices in the top 1% or bottom 1% at the end of the previous month are excluded. Volume is measured in millions of dollars. The rate on 30-day US Treasury bill is used as the risk-free rate. I define an all-but-microcap sample as the universe of stocks excluding the microcaps. The next subsection describes the LCAPM proposed in AP2005, which is estimated using these data.

1.4.2 Liquidity-adjusted Capital Asset Pricing Model

The LCAPM is a CAPM in returns net of illiquidity costs. In this model, the expected net return of a stock (stock return net of its illiquidity cost) depends on its net market beta, which is defined as the sensitivity of the net return of the stock to the market net return (return on market portfolio net of market illiquidity cost). In other words, it is the relation in gross returns that must hold if the CAPM holds in net returns. A conditional version of the LCAPM is given by

$$E_t(R_{i,t+1} - C_{i,t+1}) = R_{f,t} + \lambda_t \frac{Cov_t(R_{i,t+1} - C_{i,t+1}, R_{m,t+1} - C_{m,t+1})}{var_t(R_{m,t+1} - C_{m,t+1})} \tag{1.16}$$

where $R_{i,t+1}$ is an asset's gross return, $C_{i,t+1}$ is the illiquidity cost, R_f is the risk free rate, λ_t is the risk premium associated with net beta, $R_{m,t+1}$ is the market return and $C_{m,t+1}$ is the market illiquidity cost.

Following Lesmond, Ogden and Trzcinka (1999), henceforth referred to as LOT1999, I use the percentage of zero daily returns as a proxy of illiquidity cost. This measure of illiquidity has also been used in Bekaert, Harvey, and Lundblad (2007) and Lee (2011). Fong, Holden, and Trzcinka (2011) find that zero returns efficiently capture the time-series patterns of stock market liquidity compared to effective spread-based benchmarks. The percentage of zero returns is defined by

$$ZR_{i,t} = \frac{N_{i,t}}{T_t} \tag{1.17}$$

where $N_{i,t}$ is the number of trading days of stock i in month t that experience no price movement from the prior end-of-day price. T_t is the number of trading days in month t, which is defined by the number of days with non-missing returns. $ZR_{i,t}$ is estimated using CRSP daily stock returns. If a stock has less than ten trading days in a month, then it is dropped from the data. I also exclude observations that have a zero-return proportion greater than 80%. The average of the monthly percentage of zero returns over the past 12 months is denoted by ZR_12 and is used as a proxy for the illiquidity cost, C, in Equation 1.16.

The net beta in Equation 1.16 can be decomposed into four separate betas and the gross return on a stock can be expressed as

$$\begin{aligned} E_t(R_{i,t+1}) - R_{f,t} = E_t(C_{i,t+1}) &+ \lambda_t \frac{Cov_t(R_{i,t+1}, R_{m,t+1})}{var_t(R_{m,t+1} - C_{m,t+1})} + \lambda_t \frac{Cov_t(C_{i,t+1}, C_{m,t+1})}{var_t(R_{m,t+1} - C_{m,t+1})} \\ &- \lambda_t \frac{Cov_t(R_{i,t+1}, C_{m,t+1})}{var_t(R_{m,t+1} - C_{m,t+1})} - \lambda_t \frac{Cov_t(C_{i,t+1}, R_{m,t+1})}{var_t(R_{m,t+1} - C_{m,t+1})} \end{aligned} \tag{1.18}$$

Equation 1.18 gives a straight forward relation between the expected gross return, $E_t(R_{i,t+1})$, the expected illiquidity cost, $E_t(C_{i,t+1})$, and four covariances which represent the components of a stock's sensitivity to systematic risk. An unconditional LCAPM described by

$$E(R_{i,t} - R_{f,t}) = E(C_{i,t}) + \lambda\beta_i^1 + \lambda\beta_i^2 - \lambda\beta_i^3 - \lambda\beta_i^4 \tag{1.19}$$

can be obtained by assuming constant λ and constant conditional covariances of innovations in illiquidity and returns, where

$$\beta_i^1 = \frac{cov(R_{i,t+1}, R_{m,t+1})}{var(R_{m,t+1} - [C_{m,t+1} - E_t(C_{m,t+1})])} \tag{1.20}$$

$$\beta_i^2 = \frac{cov(C_{i,t+1} - E_t(C_{i,t+1}), C_{m,t+1} - E_t(C_{m,t+1}))}{var(R_{m,t+1} - [C_{m,t+1} - E_t(C_{m,t+1})])} \tag{1.21}$$

$$\beta_i^3 = \frac{cov(R_{i,t+1}, C_{m,t+1} - E_t(C_{m,t+1}))}{var(R_{m,t+1} - [C_{m,t+1} - E_t(C_{m,t+1})])} \tag{1.22}$$

$$\beta_i^4 = \frac{cov(C_{i,t+1} - E_t(C_{i,t+1}), R_{m,t+1})}{var(R_{m,t+1} - [C_{m,t+1} - E_t(C_{m,t+1})])} \tag{1.23}$$

β^1 is proportional to the covariance of a stock's return with the return on the market portfolio. Hence it is similar in flavor to the market beta. β^2 represents the risk due to commonality in illiquidity. Chordia et al.(2000) and Hasbrouck and Seppi (2001) find that variation in stock liquidity is positively correlated with variation in market liquidity. Thus, Equation 1.21 suggests that if shocks to liquidity cannot be diversified away, then the sensitivity of a stock to such shocks could be regarded as a component of a stock's exposure to systematic risk. Hence stocks with higher sensitivities to broad illiquidity shocks may demand a higher return. β^3 depends on the sensitivity of the stock return to fluctuations in market illiquidity. A stock with high β^3 has a high return when the market is illiquid, which works as a hedge against market illiquidity. Pastor and Stambaugh (2003) document that after controlling for exposure to other priced factors, the average return for stocks with high covariation with market liquidity exceeds that for stocks with low covariation with market liquidity by 7.5% annually. However, their paper does not control for the other components of liquidity risk. β^4 captures the sensitivity of a security's illiquidity to the market return. A stock with high β^4 is illiquid when market returns are high and is liquid in a down market. Thus an asset with high β^4 works as a hedge against market downturns since it has low illiquidity costs during states of low market return.

In AP2005, the total effect of systematic risk is captured by the combination of the three liquidity betas and the market beta. The authors argue that the liquidity betas are highly correlated with each other, and including them separately in the cross-sectional regression leads to a collinearity problem. Thus they aggregate the betas, and their framework does not allow them to identify the effects of the individual liquidity risks on asset returns. They do this by defining the net liquidity beta as

$$\beta_i^{net} \equiv \beta_i^2 - \beta_i^3 - \beta_i^4 \tag{1.24}$$

and the condensed LCAPM is given by

$$E(R_{i,t} - R_{f,t}) = E(C_{i,t}) + \lambda^1 \beta_i^1 + \lambda^{net} \beta_i^{net} \tag{1.25}$$

The next subsection describes how betas in Equations 1.24 and 1.25 are estimated.

1.4.3 Beta Estimation

This paper studies how liquidity affects stock returns using test portfolios (as done in previous papers) as well as individual stocks. Grouping stocks into portfolios is the standard method adopted to address EIV. Black et al. (1972) show that grouping can substantially reduce the bias due to measurement error and for large sample size, sampling error in the estimated betas can be eliminated. Their paper suggests estimating the group risk parameter (portfolio beta) on sample data that is not used in the ranking procedure in order to prevent an association of the measurement error in the β estimates with the errors in the coefficients used in ranking the portfolios. To implement this technique, $\widehat{\beta_{pre}}$ is computed from five years of previous monthly data for each stock. Individual securities are then assigned to groups based on their $\widehat{\beta_{pre}}$ ranking. Portfolio data from a subsequent time period is then used to estimate portfolio $\widehat{\beta_k}$.

However, Liang (2000) contends that sorting based on variables computed using a preceding sample does not completely eliminate the possibility of biased inferences due to measurement error. In the portfolio formation process, estimation errors embedded in the sorting variable can cause systematic bias in the results. He shows that grouping stocks may aggregate these measurement errors, which results in positive or negative errors for extreme portfolios that further biases the results. Moreover, grouping stocks into portfolios also causes loss of information present in individual stock data.

Though the existing literature has identified problems associated with portfolio formation procedures in asset pricing tests, this method is still used to deal with measurement error. In contrast, I use higher order cumulant based estimators to tackle the bias and inconsistency caused by EIV. This method does away with the need to use portfolios as test assets and I use individual stock-level data. To compare the performance of the two approaches mentioned above, I also use portfolio betas in the tests in this paper.

To contrast the ability of portfolio betas with that of individual betas to explain returns, betas are estimated at both the portfolio and the stock level as described below. I compute the market return for each month t based on a value-weighted average of returns on all stocks in the market portfolio in that month. Similarly, market illiquidity for month t is defined as the value-weighted average of $ZR_{i,t}$ in that month. The first order autocorrelation in market illiquidity is 0.99. Thus innovations in market illiquidity, $C_{m,t} - E_{t-1}(C_{m,t})$ are obtained from the first-differences in illiquidity levels. Similarly, for stock i, the change in illiquidity is used as its innovation.

Tests using individual stock level data use individual stock-level betas, which are obtained as follows. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23. The 60-month window rolls forward every month. Individual stock windows with less than 36 prior monthly returns or innovations in illiquidity are dropped.

Next, I describe the portfolio formation approach used in this paper to obtain the portfolio betas. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}$, $(k = 1, 2, 3, 4)$ of month t

is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months. If a stock has less than 36 valid observations in the $t-60$ to $t-1$ monthly window, then $\beta_{i,t}^{k,pre}$ for that stock is set to missing. Thus the $\beta_{i,t}^{k,pre}$ of a stock is the same as $\beta_{i,t}^{k}$ described earlier. Stocks are then sorted monthly into ten equal-weighted portfolios based on $\beta_{i,t}^{k,pre}$ for month t. The post-ranking portfolio beta, β_{p}^{k}, is then estimated for each of the ten portfolios over the entire sample period using Equations 1.20-1.23.

The post-ranking portfolio beta estimation procedure to obtain β_{p}^{1} is illustrated as follows. First, I calculate the pre-ranking $\beta_{i,t}^{1,pre}$ for stock i in month t using Equation 1.20, based on the time-series of previous 60 months of returns and illiquidity innovations. Then, at the beginning of month t, stocks are sorted into ten equally weighted portfolios based on $\beta_{i,t}^{1,pre}$. Subsequently, the post-ranking beta for portfolio p, denoted by β_{p}^{1}, is estimated over the entire sample period, using Equation 1.20. β_{p}^{1} is then assigned to all the individual stocks i, which belong to portfolio p in a given month t. The same technique is repeated to sort stocks into deciles based on $\beta_{i,t}^{k,pre}$ and calculate β_{p}^{k} for k=2,3,4. This β_{p}^{k} is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta_{i,t}^{k,pre}$. This approach is similar to that used in Lee (2011). But he uses an equal-weighted market portfolio instead of a value-weighted market portfolio to estimate the betas. Additionally, the 5-year window for $\beta_{i,t}^{k,pre}$ estimation in Lee (2011) rolls forward on a yearly basis, and the decile portfolios are formed at the beginning of every year, instead of every month as I do here.

1.5 Simulations

In this section I use simulations in order to highlight the importance of an EIV correction. Small sample performance of the cumulant estimators is compared with that of OLS estimators. I consider two models, similar to the models estimated later in the paper using real data. The first model has three mismeasured regressors and two perfectly measured regressors. This corresponds to the condensed LCAPM. The second model has five mismeasured regressors and two correctly measured regressors, thus corresponding to the LCAPM with individual liquidity risks.

1.5.1 Simulation Setup

The data generating process (DGP) should match the characteristics of the real data set as closely as possible. Hence, I generate panel data of length 564, which is equal to the number of months in the second stage CSR, and width 3000, which is equal to the average number of stocks per month. For the first model, I select values for λ_1, λ_{net}, $\lambda_{zr_{12}}$ and ρ which are the unknown parameters in the model. The entire panel is then generated from a system of equation as follows:

$$\chi^j_{i,t} = \delta_{\chi^j} + \upsilon^{\chi^j_{i,t}} \tag{1.26}$$

$$z^p_{i,t} = \delta_{z^p} + \upsilon^{z^p_{i,t}} \tag{1.27}$$

$$u_{i,t} = \upsilon^u_{i,t} \tag{1.28}$$

$$\epsilon^j_{i,t} = \upsilon^{\epsilon^j_{i,t}} \tag{1.29}$$

where $(\upsilon^{\chi^j_{i,t}}, \upsilon^{z^p_{i,t}}, \upsilon^u_{i,t}, \upsilon^{\epsilon^j_{i,t}})$ are drawn from zero-mean, unit-variance gamma distributions, j is the number of mismeasured regressors and p is the number of perfectly measured regressors in the model. Gamma distributions are used to ensure that all the assumptions stated in Section 1.3 are satisfied. $(\delta^j_\chi, \delta^p_z)$ are chosen such that the means of the simulated $(\chi^j_{i,t}, z^p_{i,t})$ equal the means of $(x^j_{i,t}, z^p_{i,t})$ in the real data. The simulated $(x^j_{i,t}, y_{i,t})$ are generated from the simulated $(\chi^j_{i,t}, z^p_{i,t})$ using Equations 2.3 and 2.4.

The measurement errors $(\epsilon^j_{i,t})$ and the regression error $(u_{i,t})$ in the simulation can be controlled by adjusting the shape parameters of the gamma distributions for $(\upsilon^{\epsilon^j_{i,t}}, \upsilon^u_{i,t})$. The measurement quality of the proxy variable is given by the R^2 of Equation 2.4. In order to test the power of the methodology used in this paper, I simulate data with high as well as low levels of measurement error in the mismeasured regressors. The results help us determine if higher order estimators are effective in eliminating bias in parameter estimates due to EIV in models with varying levels of measurement error. The following two sections give the simulation results using three and five mismeasured regressors.

Three mismeasured regressors

Table 1.1 presents the results from a model that is similar to the condensed LCAPM. In this set-up, j =3 and p =2. The table reports the slope estimates (λ_1, λ_2, λ_3), which are the coefficients on the mismeasured regressors, the slope estimates (ρ_1, ρ_2), which are the coefficients on the two correctly measured regressors and the intercept. The measurement quality of the mismeasured regressor is given by its coefficient of determination, which indicates the proportion of the variance in the mismeasured regressor that is explained by the true regressor. This is set to range from 45% to 97%, with χ^1 having the highest degree of measurement error and χ^3 the lowest. The values of the parameters used in the DGP are $\lambda_1 = 0.70$, $\lambda_2 = 0.80$, $\lambda_3 = 0.90$, $\rho_1 = 0.40$ and $\rho_2 = 0.50$. The intercept in the DGP is designed to be zero.

The first row labeled Fama-MacBeth exhibits the attenuation bias due to measurement error. The parameter estimates of the mismeasured regressors are biased downward. The intercept ρ_0 is biased upward and appears to be statistically significant when estimated using FM. The true DGP has an intercept of zero. Thus, the FM technique rejects a correctly specified asset pricing model due to measurement error. The results obtained from third, fourth and fifth order cumulant estimators are denoted by CUMD3, CUMD4 and CUMD5 respectively, and are reported in the second, third and fourth rows in Table 1.1. In contrast to FM, the intercept is not significantly different from zero for CUMD3, CUMD4 and CUMD5.

Next, I compare the estimates of λ_1, λ_2 and λ_3 obtained from Fama-MacBeth technique with those obtained from cumulant-based estimators. I find that that λ_1 and λ_2 are statistically different from zero, but biased downward (attenuation bias), when estimated using FM. However, the λ_1, λ_2 and λ_3 estimates obtained from CUMD3, CUMD4 and CUMD5 are close to the true values used in the DGP and are statistically different from zero. The results in Table 1.1 indicate that cumulant estimators perform better than FM in addressing bias caused by EIV.

Five mismeasured regressors

This section compares the performance of the Fama-MacBeth technique with that of estimators using higher order information in a model with five mismeasured regressors and two perfectly measured regressors. The slope estimates used in the DGP are λ_1= 0.1, λ_2= 0.2, λ_3= 0.3, λ_4= 0.4, λ_5= 0.5, ρ_1= 1.0, $\rho_2 = 0.5$ and the intercept is zero. I generate data with low quality proxy for χ^1, χ^3, χ^5 and high quality proxy for χ^2 and χ^4.

The first row in Table 1.2 highlights the contamination bias and the attenuation bias induced by EIV. The FM estimates are biased. The estimates of λ_1 and λ_2 have the wrong sign and λ_2 is statistically insignificant. λ_3, λ_4 and λ_5 estimates are biased downward. Similar to the model with three mismeasured regressors, the intercept is biased upward and is statistically different from zero, though in the DGP, the intercept is zero. In contrast, the estimates from CUMD3, CUMD4 and CUMD5 are unbiased and similar to each other. The values of λ_1, λ_2, λ_3, λ_4 and λ_5 are close to the true values used in the DGP, and the intercept is statistically insignificant.

The results from Table 1.2 show that in a model with many mismeasured regressors, the slope parameters obtained from FM may even have the wrong sign. These results highlight the need for EIV correction in multivariate models. The simulations give evidence of the superior performance of higher order cumulant estimators compared to FM estimators. The next section applies these estimators to asset pricing models.

1.6 Results

1.6.1 Summary Statistics

Descriptive statistics of the sample are reported in Table 1.3. The number of firm-month observations is 1,303,337 for the full sample and 655,590 for the all-but-microcap sample. Panel A1 and Panel A2 give the summary statistics of β_i^1, β_i^2, β_i^3 and β_i^4 estimated at the stock-level based on a value-weighted market portfolio, using Equations 1.20-1.23. Panel A1 reports the characteristics for all stocks and Panel A2 excludes the microcap stocks.

The striking difference between the two panels is that the average β_i^4 is negative for the whole sample, but positive when we exclude the microcap stocks. This shows that the β_i^4 for the microcap stocks is strongly negative. The table also indicates that the average β_i^3 is smaller for the microcap stocks than for the larger stocks, which results in the larger value of average β_i^3 for the all-but-microcap sample. As expected, the illiquidity cost, ZR_12 is higher in case of the full sample than for the all-but-microcap sample. The average market capitalization increases and the book-to-market ratio decreases when we exclude the microcap stocks. Univariate and multivariate tests of normality are performed to ensure that the stock-level betas are non-normal. Doornik-Hansen (2008) test of β_i^k and ZR_12 rejects the null hypothesis of normality for all the variables.

Panels B1 and B2 give the descriptive statistics of β_p^1, β_p^2, β_p^3 and β_p^4 for the ten equal-weighted portfolios for the full sample and the all-but-microcap sample respectively. Doornik-Hansen (2008) test is used to check for univariate and multivariate normality of the portfolio betas, and the p-values (not reported) from the test are greater than 0.10. Thus the portfolio betas fail to reject univariate and multivariate normality. The differences between these two panels are similar to the differences when using stock-level betas. I find that the average β_p^4 is more negative for the whole sample than for the all-but-microcap sample. This shows that the β_p^4 for microcap stocks is very negative.

Table 1.4 reports the correlations between different measures of liquidity risk. Panel A presents the correlations between the portfolio betas, β_p^1, β_p^2, β_p^3 and β_p^4 for the entire sample. Panel B reports the correlations between the stock-level betas. Similar to AP2005, I calculate the cross-sectional pair-wise correlations between the betas for individual stocks for each month. The averages of the time-series of these correlations give the correlations between the beta measures for the entire sample period.

The correlations between the portfolio betas are very small, ranging from 2.5% to 7.4%. The correlations between the stock-level betas are also very small, ranging from 2.4% to 7.9%. I find that the correlations between the portfolio betas reported in this paper are significantly lower than those reported in AP2005, which ranged from 44.1% to 97.1%. In AP2005, stocks are sorted into 25 illiquidity portfolios and the betas for each portfolio are computed as per Equations 1.20-1.23 based on an equal-weighted market portfolio, using the entire monthly time-series from 1964 to 1999. Thus, the higher correlations in AP2005 may be driven by the greater emphasis on small stocks.

1.6.2 Pricing of market risk

This section tests the validity of the single-factor CAPM and investigates whether market risk is priced. Market beta, $\beta_{i,t}^{mkt}$ for each firm i in month t is estimated by using the previous 60 monthly returns. To obtain the portfolio betas, stocks are then sorted monthly into decile portfolios based on $\beta_{i,t}^{mkt}$, and these ten equally weighted portfolios are used as test assets. The following model is then estimated using portfolio or stock-level data to estimate the market risk premium

$$E(R_{i,t} - R_{f,t}) = \alpha_0 + \lambda_{mkt}\beta_{i,t}^{mkt} \tag{1.30}$$

Panel A of Table 1.5 reports the estimated risk premium using the portfolio approach. The Fama-MacBeth results indicate that market risk is not priced. The coefficient estimate of market beta is insignificant and α_0 is positive and highly significant.

Panel B of Table 1.5 reports the results when the CAPM is estimated using stock level data. FM cross-sectional regression results are presented in the first row of Panel B. The second, third and fourth rows present the results obtained by using cumulants of order three, four and five respectively. Standard errors clustered by time are reported in parenthesis.

The Fama-MacBeth results show that the relation between expected return and market beta is flat. The intercept is positive and highly significant. Thus the CAPM is rejected. For the cumulant-based methodology, an important check is whether the assumptions underlying high order cumulant estimators are satisfied. Geary (1942) states that in order to use the cumulant based estimators, at least two of the cumulants should be different from zero. The results using CUMD3 are similar to the FM results mainly because CUMD3 has only one cumulant estimator equation, and this cumulant is not statistically different from zero. CUMD4 has three cumulants and only one of them is different from zero. Three out of the six cumulants in CUMD5 are different from zero, and hence it is the only specification that satisfies the necessary condition in Geary(1942). The results of CUMD5 show that the coefficient estimate of market beta is negative. However, alpha is positive and highly significant. According to this test, the CAPM is misspecified.

1.6.3 Is liquidity risk priced?

This section explores the relationship between expected asset returns and liquidity risk in the condensed LCAPM. Subsection 1.6.3.1 examines portfolio betas, and Subsection 1.6.3.2 examines stock level liquidity betas. The results are reported in Table 1.6.

Condensed LCAPM estimated with portfolio betas

This section reports empirical results of the test of the condensed LCAPM, which seeks to answer whether liquidity risk is priced. First, I follow the portfolio-based approach to compute the four liquidity betas based on a value-weighted market portfolio. These portfolio betas, β_p^k, are then assigned to each stock in portfolio p. This method is traditionally used in asset pricing models to address EIV in the estimated betas.

The 12-month average illiquidity cost ZR_12, the log of market capitalization and the log of the book-to-market ratio of each stock at the end of the previous month is also included in the model. The 12-month average zero-return proportion for the lagged month is used as the proxy for expected illiquidity at time t, $E(ZR\_12_{i,t})$. Fama and French (1992) find that size and the book-to-market ratio can explain the cross-section of stock returns. To control for these effects, I include the log of market capitalization and the log of the book-to-market ratio in the regression. These two variables may be considered as refinements to the intercept. If we find that the intercept is non-zero, then this specification allows us to check whether the significant intercept is due to stylized deviations in the model caused by the well-researched effects of size and book-to-market. The risk premia are estimated from the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1 \beta^1_{i,p,t} + \lambda^{net} \beta^{net}_{i,p,t} + \lambda^{zr} E(ZR\_12_{i,t}) + \rho^1 ln(MV)_{i,t} + \rho^2 ln(B/M)_{i,t} \quad (1.31)$$

where $\beta^{net}_{i,p,t} = \beta^2_{i,p,t} - \beta^3_{i,p,t} - \beta^4_{i,p,t}$ and $\beta^2_{i,p,t}$ is defined as the β^2 of portfolio p to which stock i belongs in month t. $\beta^3_{i,p,t}$ and $\beta^4_{i,p,t}$ are defined similarly. Panel A of Table 1.6 reports the means of the estimated premia. This specification separates the premium due to liquidity risk from that due to market beta and level of illiquidity cost of an asset. The FM results indicate that β^{net}_p is insignificant and has a coefficient of 0.0004 with a t-statistic of 0.667. β^1_p and ZR_12 are insignificant, and so is the intercept. Thus neither the liquidity risk nor the market risk is priced and the illiquidity cost is insignificant. The log of the book-to-market ratio is used in Equation 1.31 as a stylized deviation to the intercept. Hence, the significance on ln(B/M) is equivalent to the model's intercept being significant. Thus according to this test using portfolio betas, the LCAPM is misspecified.

Panel B1 of Table 1.3 shows that the skewness and kurtosis of the portfolio liquidity betas are low. Hence, we cannot apply the portfolio-beta method using higher order cumulants to estimate risk premia in the condensed specification. However, Panel A1 of Table 1.3 shows that the stock-level liquidity betas are non-normal. Thus we can apply the cumulant approach to the model comprising of individual betas. The next section implements this approach.

Condensed LCAPM estimated with stock-level betas

Panel B of Table 1.6 reports the empirical results illustrating the EIV correction method developed in Section 1.3. Betas are estimated from monthly data over the prior five years. Individual stock-level data are used to estimate the risk premia

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1 \beta^1_{i,t} + \lambda^{net} \beta^{net}_{i,t} + \lambda^{zr} E(ZR\_12_{i,t}) + \rho^1 ln(MV)_{i,t} + \rho^2 ln(B/M)_{i,t} \quad (1.32)$$

where $\beta_{i,t}^{net} = \beta_{i,t}^{2} - \beta_{i,t}^{3} - \beta_{i,t}^{4}$. Results from the FM cross-sectional regressions are reported in the first row of Panel B. These results are similar to those in Panel A. The results using third through fifth order cumulant estimators are reported in second, third and fourth rows of Panel B. Standard errors clustered by time are reported in parenthesis.

FM produces a small coefficient on β^{net} and it is insignificant. In contrast, the results from the cumulant estimators are sharply different from the FM results but nearly identical to each other. The coefficient on β^{net} in CUMD3 through CUMD5 is larger in magnitude than the coefficient in FM. This difference stems from the attenuation bias in the FM estimate and it highlights the need to control for the measurement error in the betas computed at the stock-level.

I find that β^{net} is positive and significant for all the specifications using the cumulant-based approach. This draws our attention to the advantages of using stock-level betas. In CUMD3, β^{net} is positive and significant, with a t-statistic of 2.21. Thus liquidity risk is priced. The coefficient on β^{1} is negative and significant after controlling for EIV. This is similar to the results in Kan, Robotti and Shanken (2013) and Shanken and Zhou (2007), which employ alternative methods to address model misspecification in beta-pricing models. The coefficients on ZR_12, ln(MV) and ln(BM) are insignificant. Thus the model is not rejected in explaining the cross-section of expected returns. Employing fourth order cumulant estimators, I get similar results. β^{net} is positive with a t-statistic of 3.65. The results get stronger as we move from CUMD3 to CUMD4. In CUMD5, β^{1}, β^{net}, ZR_12 and ln(MV) are significant.

Using higher order cumulants, identification comes from the non-normality of the true regressors. When all of the high order cumulants are different from zero, then all of the cumulant equations fully contribute to identifying the parameters. In both CUMD3 and CUMD4, 54% of the cumulants are statistically different from zero. However, in CUMD5, only 37% of cumulants are different from zero. This may indicate that CUMD3 and CUMD4 are better specifications than CUMD5.

A striking result in Table 1.6 is that the LCAPM fails to explain expected stock returns when we use the portfolio-based approach, as shown in Panel A. However, as evident from Panel B, when we use stock-level betas and use the cumulant-based approach to handle EIV, the LCAPM performs very well. These results show that the portfolio-based approach was unable to capture the true positive relation between liquidity risk and expected stock returns. Overall, I find that the LCAPM is not rejected as an explanation of the cross-section of average returns.

1.6.4 Pricing of the individual liquidity risks

As discussed earlier, Acharya and Pedersen (2005) reported a severe multi-collinearity problem when they included the three liquidity betas separately in the cross-sectional regressions. This is mainly because of the high correlation coefficients between the portfolio betas when they are estimated using the entire monthly series and based on an equal-weighted market portfolio. Their paper states that due to this reason, statistical identification of the separate effects of the three liquidity betas is difficult and they cannot conclude which of these risks are empirically relevant. To answer this question, Lee (2011) estimates the model with one liquidity beta at a time along with the market beta and cost of illiquidity. However, his method omits two of the three theoretically relevant liquidity risks in each model specification. Hence, the results may suffer from omitted variable bias.

Panel A of Table 1.7 reports the coefficients from FM regression of the full LCAPM with all betas estimated at the portfolio level. It shows that β^4 is negatively priced and significant (t-statistic=3.75). However, the direction on β^3 is counter-intuitive. Despite using portfolio betas, these results seem to suffer from contamination bias due to the measurement error in this multivariate model.

The correlations between the individual stock-level liquidity betas, as shown in Panel B of Table 1.4, are much lower than the correlations reported in previous literature. Previous papers did not deal with individual stock betas due to the noise introduced in the betas when they are estimated at stock level. But I address this error in estimated betas using the cumulant approach described in Section 1.3. Thus it is possible to empirically test the

LCAPM with the individual liquidity risks, without the need of aggregating the betas. This section investigates which of the liquidity risks is significant. I test the LCAPM given in Equation 1.33 using individual stock level data

$$E(R_{i,t}-R_{f,t}) = \rho^0+\lambda^1\beta^1_{i,t}+\lambda^2\beta^2_{i,t}+\lambda^3\beta^3_{i,t}+\lambda^4\beta^4_{i,t}+\lambda^{zr}E(ZR\_12_{i,t})+\rho^1 ln(MV)_{i,t}+\rho^2 ln(B/M)_{i,t} \tag{1.33}$$

Results are reported in Panel B of Table 1.7. The Fama-MacBeth results show that β^4 is negative and significant. However, the other liquidity betas are not significant. Moreover, the signs on β^2 and β^3 are opposite to theoretical implications, which may be attributable to the contamination bias in the model.

The second and third rows report the results obtained from CUMD3 and CUMD4 respectively. The signs on all of the liquidity betas align with theory. I find that the return premium due to β^2, which represents $cov(c_i, c_m)$, is positive and significant at the 10% level in CUMD3 and at the 1% level in CUMD4. This documents that investors demand a premium for holding a stock that is illiquid when the overall market is illiquid. Thus commonality in illiquidity is priced.

I find that β^4, which represents $cov(c_i, r_m)$, is negatively priced. It is significant at the 5% level in CUMD3 and 1% level in CUMD4. This demonstrates that investors are willing to pay a premium to hold a security that becomes more liquid when the market return is low. Stocks with high values of β^4 have lower illiquidity costs in states of poor market return and hence work as a hedge against market downturns. Thus investors have a preference for these stocks. Panel B of Table 1.7 also shows that β^3 is insignificant. Thus, in general, investors are not willing to accept a lower expected return on stocks that have a higher return when the market as a whole is more illiquid. The coefficient on the illiquidity cost is insignificant. A notable result in this table is that the intercept, the coefficient on ln(MV) and the coefficient on ln(B/M) are not significant. This implies that the model is not rejected.

Overall, I find that the risks due to $cov(c_i, c_m)$ and $cov(c_i, r_m)$ are the two most important sources of systematic liquidity risks that are related to expected stock returns and the LCAPM cannot be rejected. Moreover, the results get stronger as we move from third order to fourth order estimators. I also find that 44 % of the cumulants in CUMD3 and 47% of the cumulants in CUMD4 are statistically different from zero. This may explain the reason behind the stronger results in CUMD4.

1.7 Robustness Tests

To check the robustness of the results in Tables 1.6 and 1.7, I consider different specifications and portfolios. First, I consider whether the results are robust after I exclude the microcap stocks from the sample. The results are presented in Section 1.7.1. As a further robustness check, I re-estimate the model with an equal-weighted market portfolio in Section 1.7.2.

1.7.1 Controlling for microcap stocks

Pricing of liquidity risk after excluding the microcap stocks

This section aims to explore if the results in Tables 1.6 and 1.7 are driven by the microcap stocks. In my sample, I find that the microcap stocks account for around 50% of the observations but less than 5% of the total market capitalization. These stocks have high illiquidity and transaction costs. It is interesting to investigate whether the pricing ability of liquidity risk holds for all stocks or is limited to these highly illiquid stocks that represent only a small portion of market wealth.

This section presents empirical results of the test of the condensed LCAPM after excluding the microcap stocks from the sample. The results from the portfolio-based approach are reported in Panel A of Table 1.8. I find that β^{net} is still insignificant. However, the intercept on the model is positive and significant. This is in contrast to the insignificant intercept obtained in Panel A of Table 1.6 using the full sample. The insignificant intercept in Table 1.6 may have been caused by the contamination bias due to the mismeasured betas in the model. This indicates that the grouping method does not completely eliminate the measurement error in the betas for the microcap stocks.

Panel B reports the results from the cumulant-based approach with betas estimated at the stock level. The coefficient on β^{net} for the all-but-microcap sample is less than half its magnitude for the full sample. Furthermore, this coefficient is not significant in CUMD3 and CUMD4. Thus the positive relation between β^{net} and expected returns is mostly driven by the microcap stocks and the price of liquidity risk decreases substantially after excluding these stocks. The proportion of cumulants that are different from zero is 54% in CUMD3 and 37% in CUMD4. The intercept and the coefficient on β^{net} are significant in CUMD5. However, only 18% of the fifth order cumulants are different from zero, which may be the reason behind the difference in results between CUMD5 and the lower order cumulants. Overall, the results in Panel B are in line with the effect we would expect to see when we exclude the tiny stocks, which are also usually the most illiquid, from the sample.

Pricing of the individual liquidity betas after excluding the microcap stocks

The effect of excluding the microcap stocks in the estimation of the LCAPM with disaggregated betas is shown in Table 1.9. The results from the portfolio-based approach are reported in Panel A. I find that β^4 is still negatively priced and significant at the 5% level. However, similar to the result obtained in the previous section, I find that the intercept on the model is positive and significant. This suggests that for the microcap stocks, the grouping method may not be very efficient in eliminating the measurement errors in the individual betas.

Panel B presents the results when the model is estimated with stock-level betas. The coefficient on β^2 for the all-but-microcap sample is insignificant and much smaller in magnitude than the coefficient for the full sample. Thus the microcap stocks drive the risk premium due to the commonality in liquidity. The coefficient on β^3 is positive for the all-but-micro sample but insignificant for the full sample. This implies that the coefficient on β^3 for the microcap stocks must be very large and negative in order to cause the negative coefficient on β^3 for the full sample. β^4 is still negative, but the coefficient has a smaller magnitude than that in the full sample.

This table indicates that the effects of microcap stocks are very important when analyzing the relation between liquidity risk and expected returns. The main inference that stands out from these results is that the risk due to the covariation between a stock's return and the market liquidity is negatively priced for the microcap stocks. Thus investors are willing to accept a lower return to hold these microcap stocks that have a high return when market is illiquid.

1.7.2 LCAPM estimation with betas based on an equal-weighted market portfolio

AP2005 and Lee (2011) use an equal-weighted market portfolio to estimate the liquidity betas. They estimate the betas at the beginning of every year, based on the monthly data for the last five years. However, prior literature has highlighted the shortcomings of using an equal-weighted market portfolio. Equal-weighted market return and market illiquidity are dominated by the microcap stocks. When I use an equal-weighted instead of a value-weighted market portfolio to estimate the portfolio liquidity betas, the correlations between these betas are quite substantial, ranging from 10.35% to 42.49%. Thus, it is problematic to correctly identify the separate effects of these three liquidity betas. To circumvent this multicollinearity problem, AP2005 estimates the LCAPM with β^{net}, which is a linear combination of the three liquidity betas. Using monthly data from January 1988 to December, Lee (2011) finds that β^{net} is positive and significant at the 5% level.

Table 1.10 reports the results when the beta estimation is based on an equal-weighted market portfolio. The five-year rolling window for beta estimation rolls forward every month. I find that for the entire sample from January 1963 to December 2014, β^{net} is positive but insignificant. Panel B reports the results of the test of the LCAPM with stock-level betas. I find that the coefficient on β^{net} in CUMD3 and CUMD4 is positive and has a higher magnitude than in the value-weighted results reported in Panel B of Table 1.6. This indicates that the results using an equal-weighted market portfolio are driven to a greater degree by the microcap stocks. The coefficients on net liquidity risk and illiquidity cost are positive and significant with t-statistics of 3.83 and 3.16 respectively in CUMD4. These coefficients are insignificant in CUMD3 and CUMD5. Amongst the three specifications, the proportion of cumulants different from zero is the highest in CUMD4, which may be the reason behind the differences in coefficients estimated from using cumulants of different orders.

Table 1.11 reports the results obtained when the net liquidity beta is decomposed into its individual components. Panel A presents the coefficients from FM regression of the LCAPM using portfolio betas. I find that β^4 is negatively priced and significant. The intercept is positive and significant. Panel B reports the results when betas are calculated at the stock-level. I find that the risk due to the covariation between a stock's return and the market illiquidity, β^3, is significant and negatively priced for both CUMD3 and CUMD4. This is because these results are dominated by the microcap stocks, which have been shown to have a negative price of risk for β^3. The risk due to commonality in liquidity, β^2, is positive and insignificant. Furthermore, unlike the results obtained using CUMD3 and CUMD4 in Panel B of Table 1.7, I find that the intercept is positive and significant when I use an equal-weighted market portfolio . Thus the LCAPM is misspecified when an equal-weighted portfolio is cast as the market.

1.8 Conclusion

Measurement error is endemic in asset pricing models employing the standard two-pass cross-sectional regression methodology. The sensitivities to factors are estimated in the first stage, and used as independent variables in the second stage. Error in the estimates may lead to inconsistent estimates of the slope coefficients and the intercept in the second stage.

This paper implements a method that is new to the estimation of risk premia in asset pricing models to address EIV. I propose an alternative to the cross-sectional regression step in a Fama-MacBeth framework, and use higher order information in the data to estimate the price of risk. This is achieved by using a system of equations that express higher order cumulants of observable variables as a linear function of the coefficients to be estimated and other higher order cumulants. It does not necessitate sorting stocks into groups to test hypotheses and hence circumvents the information loss caused due to portfolio formation.

I apply this methodology to the CAPM and LCAPM. I find that the single factor CAPM still fails to explain the cross-section of average returns. In testing the LCAPM, I use stock level betas and find that liquidity risk is priced. Then I decouple the net liquidity risk into three components and empirically test the significance of these factors. Betas estimated at the stock level are not highly correlated with each other, and hence we can bypass the collinearity problem that was a challenge in the previous studies.

I find that the risk premium due to the commonality in liquidity is positive. I also find that the premium due to the sensitivity of a stock's illiquidity to the market return is negatively related to the cross-section of expected return. This paper sheds light on the true relationship between liquidity, liquidity risk and asset returns after controlling for contamination and attenuation biases. I also show that β^3, which measures the exposure of a stock's return to the market illiquidity, is negatively related to expected returns for microcap stocks.

The scope of applications of higher-order cumulants in handling EIV is expansive. Future research could investigate whether measurement error distorts the inferences in the q-factor model or the Fama-French five-factor model.

Acharya, V.V. and Pedersen, L.H., 2005. Asset pricing with liquidity risk. Journal of financial Economics, 77(2), pp.375-410.

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of financial markets, 5(1), pp.31-56.

Amihud, Y. and Mendelson, H., 1986. Asset pricing and the bid-ask spread. Journal of financial Economics, 17(2), pp.223-249.

Ang, A., Liu, J. and Schwarz, K., 2010. Using Stocks Or Portfolios in Tests of Factor Models. Working paper, Columbia University, 2010.

Bekaert, G., Harvey, C.R. and Lundblad, C., 2007. Liquidity and expected returns: Lessons from emerging markets. Review of Financial studies, 20(6), pp.1783-1831.

Berk, J.B., 2000. Sorting out sorts. The Journal of Finance, 55(1), pp.407-427.

Black, F., Jensen, M.C. and Scholes, M.S., 1972. The capital asset pricing model: Some empirical tests.

Brennan, M.J., Chordia, T. and Subrahmanyam, A., 1998. Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. Journal of Financial Economics, 49(3), pp.345-373.

Brennan, M.J. and Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of financial economics, 41(3), pp.441-464.

Brockman, P., Chung, D.Y. and Pérignon, C., 2009. Commonality in liquidity: A global perspective. Journal of Financial and Quantitative Analysis, 44(04), pp.851-882.

Brunnermeier, M.K. and Pedersen, L.H., 2009. Market liquidity and funding liquidity. Review of Financial studies, 22(6), pp.2201-2238.

Carhart, M.M., 1997. On persistence in mutual fund performance. The Journal of finance, 52(1), pp.57-82.

Chen, N.F., Roll, R. and Ross, S.A., 1986. Economic forces and the stock market. Journal of business, pp.383-403.

Chordia, T., Roll, R. and Subrahmanyam, A., 2000. Commonality in liquidity. Journal of financial economics, 56(1), pp.3-28.

Chordia, T., Roll, R. and Subrahmanyam, A., 2001. Market liquidity and trading activity. The Journal of Finance, 56(2), pp.501-530.

Coughenour, J.F. and Saad, M.M., 2004. Common market makers and commonality in liquidity. Journal of Financial Economics, 73(1), pp.37-69.

Cragg, J.G., 1994. Making good inferences from bad data. Canadian Journal of Economics, pp.776-800.

Cragg, J.G., 1997. Using higher moments to estimate the simple errors-in-variables model. Rand Journal of Economics, pp.S71-S91.

Dagenais, M.G. and Dagenais, D.L., 1997. Higher moment estimators for linear regression models with errors in the variables. Journal of Econometrics, 76(1), pp.193-221.

Doornik, J.A. and Hansen, H., 2008. An omnibus test for univariate and multivariate normality. Oxford Bulletin of Economics and Statistics, 70(s1), pp.927-939.

Erickson, T. and Whited, T.M., 2000. Measurement error and the relationship between investment and q. Journal of political economy, 108(5), pp.1027-1057.

Erickson, T. and Whited, T.M., 2002. Two-step GMM estimation of the errors-in-variables model using high-order moments. Econometric Theory, 18(03), pp.776-799.

Erickson, T. and Whited, T.M., 2012. Treating measurement error in Tobin's q. Review of Financial Studies, 25(4), pp.1286-1329.

Erickson, T., Jiang, C.H. and Whited, T.M., 2014. Minimum distance estimation of the errors-in-variables model using linear cumulant equations. Journal of Econometrics, 183(2), pp.211-221.

Fama, E.F. and French, K.R., 1992. The crosssection of expected stock returns. the Journal of Finance, 47(2), pp.427-465.

Fama, E.F. and French, K.R., 2008. Dissecting anomalies. The Journal of Finance, 63(4), pp.1653-1678.

Fama, E.F. and MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. The journal of political economy, pp.607-636.

Fong, K., Holden, C.W. and Trzcinka, C.A., 2011. Can global stock liquidity be measured. Unpublished working paper. University of New South Wales and Indiana University.

Geary, R.C., 1942. The estimation of many parameters. Journal of the Royal Statistical Society, 105(3), pp.213-217.

Goyenko, R.Y., Holden, C.W. and Trzcinka, C.A., 2009. Do liquidity measures measure liquidity?. Journal of financial Economics, 92(2), pp.153-181.

Han, Y. and Lesmond, D., 2011. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. Review of Financial Studies, 24(5), pp.1590-1629.

Harvey, C.R., 1991. The world price of covariance risk. The Journal of Finance, 46(1), pp.111-157.

Hasbrouck, J. and Seppi, D.J., 2001. Common factors in prices, order flows, and liquidity. Journal of financial Economics, 59(3), pp.383-411.

Hou, K., Xue, C. and Zhang, L., 2015. Digesting Anomalies: An Investment Approach. Review of Financial Studies, 28(3), pp.650-705.

Huberman, G. and Halka, D., 2001. Systematic liquidity. Journal of Financial Research, 24(2), pp.161-178.

Jagannathan, R., Kim, S. and Skoulakis, G., 2010. Revisiting the Errors in Variables Problem in Studying the Cross Section of Stock Returns. Unpublished Working Paper. Northwestern University.

Jagannathan, R. and Wang, Z., 1996. The conditional CAPM and the crosssection of expected returns. The Journal of finance, 51(1), pp.3-53.

Jagannathan, R. and Wang, Z., 1998. An asymptotic theory for estimating betapricing models using crosssectional regression. The Journal of Finance, 53(4), pp.1285-1309.

Kamara, A., Lou, X. and Sadka, R., 2008. The divergence of liquidity commonality in the cross-section of stocks. Journal of Financial Economics, 89(3), pp.444-466.

Kan, R., Robotti, C. and Shanken, J., 2013. Pricing Model Performance and the TwoPass CrossSectional Regression Methodology. The Journal of Finance, 68(6), pp.2617-2649.

Karolyi, G.A., Lee, K.H. and Van Dijk, M.A., 2012. Understanding commonality in liquidity around the world. Journal of Financial Economics, 105(1), pp.82-112.

Kim, D., 1995. The errors in the variables problem in the crosssection of expected stock returns. The Journal of Finance, 50(5), pp.1605-1634.

Kim, S. and Skoulakis, G., 2015. Ex-post Risk Premia: Estimation and Inference using Large Cross Sections. Unpublished Working Paper.

Korajczyk, R.A. and Sadka, R., 2008. Pricing the commonality across alternative measures of liquidity. Journal of Financial Economics, 87(1), pp.45-72.

Lintner, J., 1965. Security prices, risk, and maximal gains from diversification. The Journal of Finance, 20(4), pp.587-615.

Lee, K.H., 2011. The world price of liquidity risk. Journal of Financial Economics, 99(1), pp.136-161.

Lesmond, D.A., 2005. Liquidity of emerging markets. Journal of Financial Economics, 77(2), pp.411-452.

Lesmond, D.A., Ogden, J.P. and Trzcinka, C.A., 1999. A new estimate of transaction costs. Review of Financial Studies, 12(5), pp.1113-1141.

Lewbel, A., 1997. Constructing instruments for regressions with measurement error when no additional data are available, with an application to patents and R&D. Econometrica, 65(5), pp.1201-1213.

Lewbel, A., 2012. Using heteroscedasticity to identify and estimate mismeasured and endogenous regressor models. Journal of Business & Economic Statistics 30, 67-80.

Lewellen, J., Nagel, S. and Shanken, J., 2010. A skeptical appraisal of asset pricing tests. Journal of Financial economics, 96(2), pp.175-194.

Liang, B., 2000. Portfolio formation, measurement errors, and beta shifts: A random sampling approach. Journal of Financial Research, 23(3), pp.261-284.

Liu, W., 2006. A liquidity-augmented capital asset pricing model. Journal of financial Economics, 82(3), pp.631-671.

Lo, A.W. and MacKinlay, A.C., 1990. Data-snooping biases in tests of financial asset pricing models. Review of financial studies, 3(3), pp.431-467.

Madansky, A., 1959. The fitting of straight lines when both variables are subject to error. Journal of the American Statistical Association 54, 173-205.

Martin, I.W., 2013. Consumption-based asset pricing with higher cumulants. The Review of Economic Studies, 80(2), pp.745-773.

McCullagh, P., 1987. Tensor Methods in Statistics. Chapman and Hall, New York.

Newey, W.K. and West, K.D., 1994. Automatic lag selection in covariance matrix estimation. The Review of Economic Studies, 61(4), pp.631-653.

Pakes, A., 1982. On the asymptotic bias of Wald-type estimators of a straight line when both variables are subject to error. International Economic Review, 23(2), pp.491-497.

Pal, M., 1980. Consistent moment estimators of regression coefficients in the presence of errors in variables. Journal of Econometrics, 14(3), pp.349-364.

Pastor, L. and Stambaugh, R.F., 2001. Liquidity risk and expected stock returns (No. w8462). National Bureau of Economic Research.

Pukthuanthong, K., Roll, R. and Wang, J.L., 2014. Resolving the errors-in-variables bias in risk premium estimation. Working Paper,UCLA.

Richardson, M. and Smith, T., 1993. A test for multivariate normality in stock returns. Journal of Business, pp.295-321.

Roll, R., 1977. A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory. Journal of financial economics, 4(2), pp.129-176.

Sadka, R., 2006. Momentum and post-earnings-announcement drift anomalies: The role of liquidity risk. Journal of Financial Economics, 80(2), pp.309-349.

Sargan, J.D., 1958. The estimation of economic relationships using instrumental variables. Econometrica: Journal of the Econometric Society, pp.393-415.

Shanken, J., 1985. Multivariate tests of the zero-beta CAPM. Journal of financial economics, 14(3), pp.327-348.

Shanken, J., 1992. On the estimation of beta-pricing models. Review of Financial studies, 5(1), pp.1-33.

Shanken, J. and Zhou, G., 2007. Estimating and testing beta pricing models: Alternative methods and their performance in simulations. Journal of Financial Economics, 84(1), pp.40-86.

Sharpe, W.F., 1964. Capital asset prices: A theory of market equilibrium under conditions of risk. The journal of finance, 19(3), pp.425-442.

Spiegelman, C., 1979. On estimating the slope of a straight line when both variables are subject to error. The Annals of Statistics, pp.201-206.

Van Montfort, K., Mooijaart, A. and Leeuw, J.D., 1987. Regression with errors in variables: estimators based on third order moments. Statistica Neerlandica, 41(4), pp.223-238.

Van Montfort, K., Mooijaart, A. and De Leeuw, J., 1989. Estimation of regression coefficients with the help of characteristic functions. Journal of Econometrics, 41(2), pp.267-278.

Wansbeek, T.J. and Meijer,E., 2000. Measurement error and latent variables in econometrics. Elsevier, Amsterdam.

Watanabe, A. and Watanabe, M., 2008. Time-varying liquidity risk and the cross section of stock returns. Review of Financial Studies, 21(6), pp.2449-2486.

Table 1.1 Simulation Results: Three mismeasured regressors

The data are simulated by the following model:

$$y_{i,t} = 0.7\chi_{i,t}^1 + 0.8\chi_{i,t}^2 + 0.9\chi_{i,t}^3 + 0.4z_{i,t}^1 + 0.5z_{i,t}^2$$

The estimated model is

$$y_{i,t} = \lambda_1\chi_{i,t}^1 + \lambda_2\chi_{i,t}^2 + \lambda_3\chi_{i,t}^3 + \rho_0 + \rho_1 z_{i,t}^1 + \rho_2 z_{i,t}^2$$

The total number of time periods is 564, and each time period has 3000 observations. This table reports the estimates obtained by applying the Fama-MacBeth methodology and higher order cumulant estimators. The standard errors are reported in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

Three mismeasured regressors						
	λ_1	λ_2	λ_3	ρ_0	ρ_1	ρ_2
Fama-MacBeth	0.3198***	0.6159***	0.8737***	0.3653***	0.4002***	0.4999***
	(0.0008)	(0.0007)	(0.0003)	(0.0017)	(0.0006)	(0.0003)
CUMD3	0.6989***	0.8006***	0.8993***	0.0020	0.3999***	0.4997***
	(0.0032)	(0.0017)	(0.0009)	(0.0035)	(0.0007)	(0.0003)
CUMD4	0.6950***	0.8008***	0.8992***	0.0056	0.3999***	0.4998***
	(0.0027)	(0.0014)	(0.0007)	(0.0031)	(0.0007)	(0.0003)
CUMD5	0.6969***	0.8010***	0.8993***	0.0038	0.3999***	0.4998***
	(0.0023)	(0.0012)	(0.0007)	(0.0028)	(0.0007)	(0.0003)

Table 1.2 Simulation Results: Five Mismeasured regressors

The data are simulated by the following model:

$$y_{i,t} = 0.1\chi^1_{i,t} + 0.2\chi^2_{i,t} + 0.3\chi^3_{i,t} + 0.4\chi^4_{i,t} + 0.5\chi^5_{i,t} + 0.4z^1_{i,t} + 0.5z^2_{i,t}$$

The estimated model is

$$y_{i,t} = \lambda_1\chi^1_{i,t} + \lambda_2\chi^2_{i,t} + \lambda_3\chi^3_{i,t} + \lambda_4\chi^4_{i,t} + \lambda_5\chi^5_{i,t} + \rho_0 + \rho_1 z^1_{i,t} + \rho_2 z^2_{i,t}$$

The total number of time periods is 564, and each time period has 3000 observations. I report the estimates obtained by applying the Fama-MacBeth methodology and higher order cumulant estimators.The standard errors are reported in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

	Five mismeasured regressors							
	λ_1	λ_2	λ_3	λ_4	λ_5	ρ_0	ρ_1	ρ_2
Fama-MacBeth	-0.0476***	-0.0219	0.1483***	0.3619***	0.3979***	1.1090***	1.0001***	0.5002***
	(0.0148)	(0.0178)	(0.0119)	(0.0111)	(0.0118)	(0.1244)	(0.0164)	(0.0253)
CUMD3	0.1005***	0.1989***	0.2945***	0.4019***	0.4971***	-0.0035	1.0000***	0.5005***
	(0.0032)	(0.0033)	(0.0041)	(0.0034)	(0.0031)	(0.0193)	(0.0007)	(0.0011)
CUMD4	0.1011***	0.2011***	0.2938***	0.4035***	0.4968***	-0.0109	1.0000***	0.5005***
	(0.0026)	(0.0026)	(0.0035)	(0.0028)	(0.0023)	(0.0153)	(0.0007)	(0.0011)

Table 1.3 Summary Statistics: Stock-level betas

This table gives the descriptive statistics of the data. Panels A1 and A2 report the mean, standard deviation, skewness and kurtosis of the stock-level betas in the LCAPM estimated from January 1963 to December 2014 for all and the all-but-microcap stocks respectively. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is estimated from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23. The innovations in illiquidity are obtained from the first differences in illiquidity over the 60-month window. This window rolls forward every month. Individual stock windows with less than 36 prior monthly returns or innovations in illiquidity are dropped. ZR_12 is the average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio.

Panel A1: All stocks				
Variable	Mean	Std Dev	Skewness	Kurtosis
β^1_i	1.1181	0.7006	1.1115	4.6298
β^2_i	0.0581	0.1509	0.6012	4.7382
β^3_i	0.0169	0.1422	0.3248	18.5410
β^4_i	-0.0099	0.3667	-0.1060	2.4426
ZR_12	0.1729	0.1479	0.9632	0.6384
ln(MV)	4.9541	2.1104	0.3374	-0.2008
ln(B/M)	-0.4251	0.8929	-0.5936	2.1015

Panel A2: Excluding microcap stocks				
Variable	Mean	Std Dev	Skewness	Kurtosis
β^1_i	1.1087	0.6274	1.3814	5.0206
β^2_i	0.0610	0.1302	1.0014	4.5511
β^3_i	0.0197	0.1023	0.3787	9.9942
β^4_i	0.0035	0.3065	-0.0178	2.8852
ZR_12	0.1118	0.1079	1.4589	3.5614
ln(MV)	6.6272	1.6393	0.4499	0.2057
ln(B/M)	-0.6232	0.8048	-0.6386	1.8132

Table 1.3 Summary Statistics: Portfolio betas

This table gives the descriptive statistics of the data. Panels B1 and B2 report the mean, standard deviation, skewness and kurtosis of portfolio betas in the LCAPM estimated from January 1963 to December 2014 for all and the all-but-microcap stocks respectively. For each stock i, the pre-ranking beta, $\beta_{i,t}^{k,pre}$, $(k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the value-weighted market return or the innovations in value-weighted market illiquidity. The innovations in illiquidity are obtained from the first differences in illiquidity over the 60-month window. This window rolls forward every month. If a stock has less than 36 valid observations in the $t-60$ to $t-1$ monthly window, then $\beta_{i,t}^{k,pre}$ for that stock is set to missing. Stocks are then sorted into ten equal-weighted portfolios based on $\beta_{i,t}^{k,pre}$ for month t. The post-ranking portfolio beta, β_p^k, is then estimated for each of the ten portfolios over the entire sample period using Equations 1.20-1.23.

Panel B1: All stocks

	Mean	Std Dev	Skewness	Kurtosis
β_p^1	1.0969	0.5958	0.4025	-0.2503
β_p^2	0.0549	0.0824	0.0199	-0.3740
β_p^3	0.0061	0.0859	-0.2119	0.3176
β_p^4	-0.0124	0.3200	-0.0446	-0.2304

Panel B2: Excluding microcap stocks

Variable	Mean	Std Dev	Skewness	Kurtosis
β_p^1	1.0876	0.5476	0.5056	0.0868
β_p^2	0.0560	0.0708	0.0644	0.0789
β_p^3	0.0098	0.0692	-0.1554	1.3495
β_p^4	-0.0014	0.2666	-0.0708	0.3918

Table 1.4 Correlation between the betas

This table reports the correlations between the different beta measures. The correlations are computed monthly for all eligible stocks and then averaged over the sample period. Panel A presents the correlations of β_p^1, β_p^2, β_p^3 and β_p^4 for the portfolio betas formed each month using data from January 1963 to December 2014. For each stock i, the pre-ranking beta, $\beta_{i,t}^{k,pre}$, $(k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the value-weighted market return or the innovations in value-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta_{i,t}^{k,pre}$ for month t. The post-ranking portfolio beta, β_p^k, is then estimated for each of the ten equal-weighted portfolios over the entire sample period using Equations 1.20-1.23. Panel B reports the correlations of β_i^1, β_i^2, β_i^3 and β_i^4 estimated at the stock-level. For each stock i in month t, $\beta_{i,t}^k$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23.

Panel A :
Beta correlations for portfolios

	β_p^1	β_p^2	β_p^3	β_p^4
β_p^1	1.000	0.025	0.058	-0.067
β_p^2		1.000	-0.074	0.059
β_p^3			1.000	0.026
β_p^4				1.000

Panel B:
Beta correlations for individual stocks

	β_i^1	β_i^2	β_i^3	β_i^4
β_i^1	1.000	0.026	0.058	-0.073
β_i^2		1.000	-0.079	0.061
β_i^3			1.000	0.024
β_i^4				1.000

Table 1.5 Pricing of market risk

This table presents the estimated coefficients of the single-factor CAPM. I consider the following model

$$E(R_t^p - R_{f,t}) = \alpha_0 + \lambda_{mkt}\beta^{mkt}$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third through fifth order cumulants. The standard errors are reported in parenthesis. Panel A reports the estimated coefficients from Fama-MacBeth cross-sectional regressions of the CAPM based on portfolio beta. Market beta, $\beta_{i,t}^{mkt}$ for each firm i in month t is estimated by using the previous 60 monthly returns. To obtain the portfolio betas, stocks are then sorted into deciles portfolios based on $\beta_{i,t}^{mkt}$, and these ten equally weighted portfolios are used as test assets. Panel B reports the estimates from estimating market beta at the stock level and then controlling for EIV explicitly using third through fifth order cumulants. ** and *** denote significance at the 5% and 1% level respectively.

Panel A : CAPM estimated with portfolio beta		
	β_p^{mkt}	α_0
Fama-MacBeth	0.0023	0.0095***
	(0.0019)	(0.0009)

Panel B: CAPM estimated with stock level beta		
	β_i^{mkt}	α_0
Fama-MacBeth	-0.0001	0.0091***
	(0.0011)	(0.0013)
CUMD3	-4.609	4.915***
	(4.474)	(0.0020)
CUMD4	-0.5225***	0.5689***
	(0.1727)	(0.0002)
CUMD5	-0.4444***	0.4858***
	(0.0658)	(0.0002)

Table 1.6 Estimation of the Condensed LCAPM

This table presents the estimated coefficients of the LCAPM using data on all stocks from January 1963 to December 2014. I consider the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1\beta^1_{i,t} + \lambda^{net}\beta^{net}_{i,t} + \lambda^{zr}E(ZR\_12_{i,t}) + \rho^1 ln(MV_{i,t}) + \rho^2 ln(B/M_{i,t})$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third and fourth order cumulant estimators. The standard errors are reported in parenthesis. Panel A reports the coefficients obtained from estimating the LCAPM using portfolio betas, β^k_p. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}$, $(k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the value-weighted market return or the innovations in value-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta^{k,pre}_{i,t}$ for month t. The post-ranking portfolio beta, β^k_p, is then estimated for each of the ten equally weighted portfolios over the entire sample period using Equations 1.20-1.23. This β^k_p is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta^{k,pre}_{i,t}$. $\beta^{net}_p = \beta^2_p$ - β^3_p - β^4_p. Panel B reports the coefficients obtained from estimating the LCAPM using stock-level betas and then controlling for EIV explicitly using third and fourth order cumulant estimators. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t - 60$ to $t - 1$, using Equations 1.20-1.23. $\beta^{net}_i = \beta^2_i$ - β^3_i - β^4_i. ZR_12 is the previous month's average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio at the end of the previous year. ** and *** denote significance at the 5% and 1% level respectively.

Panel A : Condensed LCAPM estimated with portfolio betas

	β^1_p	β^{net}_p	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0019	0.0004	0.0075	0.0030	0.0001	0.0010**
	(0.0020)	(0.0006)	(0.0079)	(0.0026)	(0.0004)	(0.0005)

Panel B: Condensed LCAPM estimated with stock-level betas

	β^1	β^{net}	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0016	0.0002	0.0140	0.0012	0.0001	0.0009
	(0.0018)	(0.0008)	(0.0076)	(0.0027)	(0.0004)	(0.0005)
CUMD3	-0.0237***	0.0062**	0.0928	-0.0039	0.0044	0.0001
	(0.0053)	(0.0028)	(0.1283)	(0.0510)	(0.0056)	(0.0025)
CUMD4	-0.0173***	0.0073***	0.0163	0.0190	0.0012	0.0019
	(0.0025)	(0.0020)	(0.0377)	(0.0167)	(0.0019)	(0.0013)
CUMD5	-0.0038**	0.0125***	0.0648***	-0.0161	0.0035***	0.0025
	(0.0017)	(0.0016)	(0.0223)	(0.0102)	(0.0012)	(0.0014)

Table 1.7 Pricing of the individual liquidity betas

This table presents the estimated coefficients of the LCAPM using data on all stocks from January 1963 to December 2014. I consider the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1 \beta^1_{i,t} + \lambda^2 \beta^2_{i,t} + \lambda^3 \beta^3_{i,t} + \lambda^4 \beta^4_{i,t} + \lambda^{zr} E(ZR\_12_{i,t}) + \rho^1 ln(MV_{i,t}) + \rho^2 ln(B/M_{i,t})$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third and fourth order cumulant estimators. The standard errors are reported in parenthesis. Panel A reports the coefficients obtained from estimating the LCAPM using portfolio betas, β^k_p. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}$, $(k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the value-weighted market return or the innovations in value-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta^{k,pre}_{i,t}$ for month t. The post-ranking portfolio beta, β^k_p, is then estimated for each of the ten equally weighted portfolios over the entire sample period using Equations 1.20-1.23. This β^k_p is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta^{k,pre}_{i,t}$. Panel B reports the coefficients obtained from estimating the LCAPM using stock-level betas and then controlling for EIV explicitly using third and fourth order cumulant estimators. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23. ZR_12 is the previous month's average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio at the end of the previous year. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: LCAPM estimated with portfolio betas

	β^1_p	β^2_p	β^3_p	β^4_p	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0018	0.0020	0.0143	-0.0015***	0.0081	0.0027	0.0001	0.0010**
	(0.0020)	(0.0017)	(0.0081)	(0.0004)	(0.0079)	(0.0026)	(0.0004)	(0.0005)

Panel B: LCAPM estimated with stock-level betas

	β^1	β^2	β^3	β^4	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0020	-0.0033	0.0165	-0.0015**	0.0152**	0.0005	0.0002	0.0009
	(0.0020)	(0.0056)	(0.0147)	(0.0006)	(0.0076)	(0.0027)	(0.0004)	(0.0005)
CUMD3	-0.0102***	0.0076	-0.0119	-0.0043**	0.0127	0.0126	0.0011	0.0027
	(0.0035)	(0.0045)	(0.0146)	(0.0022)	(0.0414)	(0.0180)	(0.0019)	(0.0015)
CUMD4	-0.0151***	0.0092***	-0.0059	-0.0052***	0.0188	0.0155	0.0013	0.0021
	(0.0016)	(0.0021)	(0.0047)	(0.0012)	(0.0186)	(0.0087)	(0.0010)	(0.0014)

Table 1.8 Robustness Check: Estimation of the LCAPM excluding microcap stocks

This table presents the estimated coefficients of the LCAPM using data from January 1963 to December 2014, excluding microcap stocks. I consider the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1\beta^1_{i,t} + \lambda^{net}\beta^{net}_{i,t} + \lambda^{zr}E(ZR\_12_{i,t}) + \rho^1 ln(MV_{i,t}) + \rho^2 ln(B/M_{i,t})$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third through fifth order cumulant estimators. The standard errors are reported in parenthesis. Panel A reports the coefficients obtained from estimating the LCAPM using portfolio betas, β^k_p. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}, (k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the value-weighted market return or the innovations in value-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta^{k,pre}_{i,t}$ for month t. The post-ranking portfolio beta, β^k_p, is then estimated for each of the ten equal-weighted portfolios over the entire sample period using Equations 1.20-1.23. This β^k_p is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta^{k,pre}_{i,t}$. $\beta^{net}_p = \beta^2_p$ - β^3_p - β^4_p. Panel B reports the coefficients obtained from estimating the LCAPM using stock-level betas and then controlling for EIV explicitly using third through fifth order cumulant estimators. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23. $\beta^{net}_i = \beta^2_i$ - β^3_i - β^4_i. ZR_12 is the previous month's average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio at the end of the previous year. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: Condensed LCAPM estimated with portfolio betas

	β^1_p	β^{net}_p	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0009	0.0006	0.0449**	0.0237***	-0.0025***	-0.0005
	(0.0020)	(0.0007)	(0.0221)	(.0031)	(0.0031)	(0.0006)

Panel B: Condensed LCAPM estimated with stock-level betas

	β^1	β^{net}	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0008	0.0009	0.0417	0.0236***	-0.0025***	-0.0006
	(0.0021)	(0.0013)	(0.0209)	(0.0032)	(0.0004)	(0.0006)
CUMD3	-0.0066	0.0032	0.0018	0.0269	-0.0014	0.0004
	(0.0052)	(0.0034)	(0.052)	(0.0204)	(0.0021)	(0.002)
CUMD4	-0.0151***	0.0029	0.0573***	0.0182	0.0002	-0.0019
	(0.0027)	(0.0018)	(0.0179)	(0.0104)	(0.0012)	(0.0019)
CUMD5	-0.0122***	0.0062***	0.0182	0.0278***	-0.001	-0.0007
	(0.0019)	(0.0014)	(0.0118)	(0.0083)	(0.001)	(0.0018)

Table 1.9 Robustness Check: Pricing of the individual liquidity betas excluding the microcap stocks

This table presents the estimated coefficients of the LCAPM using data from January 1963 to December 2014, excluding microcap stocks. I consider the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1\beta^1_{i,t} + \lambda^2\beta^2_{i,t} + \lambda^3\beta^3_{i,t} + \lambda^4\beta^4_{i,t} + \lambda^{zr}E(ZR\_12_{i,t}) + \rho^1 ln(MV_{i,t}) + \rho^2 ln(B/M_{i,t})$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third and fourth order cumulant estimators. The standard errors are reported in parenthesis. Panel A reports the coefficients obtained from estimating the LCAPM using portfolio betas, β^k_p. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}, (k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the value-weighted market return or the innovations in value-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta^{k,pre}_{i,t}$ for month t. The post-ranking portfolio beta, β^k_p, is then estimated for each of the ten equal-weighted portfolios over the entire sample period using Equations 1.20-1.23. This β^k_p is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta^{k,pre}_{i,t}$. Panel B reports the coefficients obtained from estimating the LCAPM using stock-level betas and then controlling for EIV explicitly using third and fourth order cumulant estimators. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the value-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23. ZR_12 is the previous month's average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio at the end of the previous year. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: LCAPM estimated with portfolio betas								
	β^1_p	β^2_p	β^3_p	β^4_p	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0008	0.0024	0.0081	-0.0013**	0.0438	0.0233***	-0.0025***	-0.0005
	(0.0020)	(0.0020)	(0.0077)	(0.0006)	(0.0222)	(0.0031)	(0.0004)	(0.0006)

Panel B: LCAPM estimated with stock-level betas								
	β^1	β^2	β^3	β^4	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0008	0.0008	0.004	-0.0015	0.0398	0.0229***	-0.0025***	-0.0006
	(0.0021)	(0.0074)	(0.0153)	(0.0008)	(0.0206)	(0.0032)	(0.0004)	(0.0006)
CUMD3	-0.0054	-0.0118	0.0072	-0.0025	0.0256	0.0189	-0.0007	0.0001
	(0.0039)	(0.006)	(0.019)	(0.0017)	(0.0238)	(0.0131)	(0.0014)	(0.0017)
CUMD4	-0.0116***	0.0015	0.0553***	-0.0027***	0.0220**	0.0258***	-0.0009	0.0000
	(0.0019)	(0.0022)	(0.0074)	(0.0010)	(0.0098)	(0.0079)	(0.001)	(0.0018)

Table 1.10 Condensed LCAPM estimated with an equal-weighted market portfolio

This table presents the estimated coefficients of the LCAPM using data from January 1963 to December 2014 with the beta estimation based on an equal-weighted market portfolio. I consider the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1 \beta^1_{i,t} + \lambda^{net} \beta^{net}_{i,t} + \lambda^{zr} E(ZR\_12_{i,t}) + \rho^1 ln(MV_{i,t}) + \rho^2 ln(B/M_{i,t})$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third through fifth order cumulant estimators. The standard errors are reported in parenthesis. Panel A reports the coefficients obtained from estimating the LCAPM using portfolio betas, β^k_p. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}, (k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the equal-weighted market return or the innovations in equal-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta^{k,pre}_{i,t}$ for month t. The post-ranking portfolio beta, β^k_p, is then estimated for each of the ten equal-weighted portfolios over the entire sample period using Equations 1.20-1.23. This β^k_p is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta^{k,pre}_{i,t}$. $\beta^{net}_p = \beta^2_p$ - β^3_p - β^4_p. Panel B reports the coefficients obtained from estimating the LCAPM using stock-level betas and then controlling for EIV explicitly using third through fifth order cumulant estimators. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the equal-weighted market portfolio, over months $t - 60$ to $t - 1$, using Equations 1.20-1.23. $\beta^{net}_i = \beta^2_i$ - β^3_i - β^4_i. ZR_12 is the previous month's average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio at the end of the previous year. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: Condensed LCAPM estimated with portfolio betas

	β^1_p	β^{net}_p	ZR_12	intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0009	0.0011	0.0041	0.0048**	0.00004	0.0006
	(0.0028)	(0.0009)	(0.0079)	(0.0023)	(0.0003)	(0.0004)

Panel B: Condensed LCAPM estimated with stock-level betas

	β^1	β^{net}	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	-0.0005	0.0007	0.0045	0.0042	0.0003	0.0008**
	(0.0024)	(.0012)	(0.0087)	(0.0023)	(0.0003)	(0.0004)
CUMD3	-0.0595***	0.0204	0.7897	-0.2419***	0.0321***	-0.0159***
	(0.0219)	(0.0294)	(0.4954)	(0.0081)	(0.0072)	(0.0073)
CUMD4	-0.0441***	0.0181***	0.1171***	0.0072	0.0036	-0.0034
	(0.0034)	(0.0047)	(0.0371)	(0.0049)	(0.0031)	(0.0033)
CUMD5	-0.0079***	-0.0051	0.0371	0.0020	0.0016**	0.0024
	(0.0022)	(0.0036)	(0.0267)	(0.0039)	(0.0008)	(0.0015)

Table 1.11 Pricing of the individual liquidity betas with an equal-weighted market portfolio

This table presents the estimated coefficients of the LCAPM using data from January 1963 to December 2014 with the beta estimation based on an equal-weighted market portfolio. I consider the following model

$$E(R_{i,t} - R_{f,t}) = \rho^0 + \lambda^1\beta^1_{i,t} + \lambda^2\beta^2_{i,t} + \lambda^3\beta^3_{i,t} + \lambda^4\beta^4_{i,t} + \lambda^{zr}E(ZR\_12_{i,t}) + \rho^1 ln(MV_{i,t}) + \rho^2 ln(B/M_{i,t})$$

I report the estimates obtained using Fama-MacBeth methodology and correcting for EIV using third and fourth order cumulant estimators. The standard errors are reported in parenthesis. Panel A reports the coefficients obtained from estimating the LCAPM using portfolio betas, β^k_p. For each stock i, the pre-ranking beta, $\beta^{k,pre}_{i,t}, (k = 1, 2, 3, 4)$ of month t is estimated using the time-series of monthly returns and innovations in illiquidity for the previous 60 months with respect to either the equal-weighted market return or the innovations in equal-weighted market illiquidity. Stocks are then sorted into ten portfolios based on $\beta^{k,pre}_{i,t}$ for month t. The post-ranking portfolio beta, β^k_p, is then estimated for each of the ten equal-weighted portfolios over the entire sample period using Equations 1.20-1.23. This β^k_p is assigned to all stocks belonging to portfolio p ranked on the basis of $\beta^{k,pre}_{i,t}$. Panel B reports the coefficients obtained from estimating the LCAPM using stock-level betas and then controlling for EIV explicitly using third and fourth order cumulant estimators. For each stock i in month t, $\beta^k_{i,t}$ where k = (1,2,3,4), is computed from monthly returns and innovations in illiquidity for stock i and for the equal-weighted market portfolio, over months $t-60$ to $t-1$, using Equations 1.20-1.23. ZR_12 is the previous month's average zero-return proportion. Ln(MV) is the log of the market capitalization and ln(B/M) is the log of the book-to-market ratio at the end of the previous year. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: LCAPM estimated with portfolio betas

	β^1_p	β^2_p	β^3_p	β^4_p	ZR_12	intercept	ln(MV)	ln(BM)
Fama-MacBeth	0.0016	0.0022	0.0241	-0.0026***	0.0061	0.0049**	-0.0000	0.0006
	(0.0029)	(0.0029)	(0.0134)	(0.0006)	(0.0079)	(0.0023)	(0.0003)	(0.0004)

Panel B: LCAPM estimated with stock-level betas

	β^1	β^2	β^3	β^4	ZR_12	Intercept	ln(MV)	ln(B/M)
Fama-MacBeth	0.0006	-0.0041	0.0239	-0.0026***	0.0062	0.0039	0.0002	0.0007
	(0.0029)	(0.0100)	(0.0217)	(0.0008)	(0.0084)	(0.0022)	(0.0002)	(0.0004)
CUMD3	-0.0126***	0.0188	-0.0402***	-0.0077***	-0.0437	0.0336***	-0.0015	0.0033***
	(0.0051)	(0.0109)	(0.0163)	(0.0031)	(0.0398)	(0.0039)	(0.0021)	(0.0014)
CUMD4	-0.0209***	0.0015	-0.0272***	0.0001	-0.0059	0.0310***	-0.0006	0.0015
	(0.0021)	(0.0043)	(0.0079)	(0.0017)	(0.0147)	(0.0039)	(0.0007)	(0.0014)

Chapter 2

Pricing of Idiosyncratic Volatility: Levels and Differences

2.1 Introduction

The trade-off between risk and return has always been fundamental to asset pricing. The capital asset pricing model of Sharpe (1964) and Lintner (1965) captures only the systematic risk as the priced element. It assumes frictionless markets and is based on the supposition that all investors hold the market portfolio in equilibrium. However, complete diversification may not be the case in reality. Extensive literature has shown that since investors hold undiversified portfolios, firm-specific risk is an important factor that affects investor returns. Nevertheless, the relationship between idiosyncratic risk and average returns is controversial.

Merton (1987) and Malkiel and Xu (2004) relaxed the complete diversification assumption, and developed models that predict that idiosyncratic risk is positively related to the cross-section of expected returns. They assert that investors require a premium for bearing idiosyncratic risk in the less diversified portfolios, which results in the pricing ability of idiosyncratic volatility.

Contrary to these theoretical models, there is a second line of literature, which documents a negative relation between idiosyncratic volatility (IVOL) and the cross-section of average returns. Ang et al. (2006), henceforth referred to as AHXZ (2006), show that stocks with high idiosyncratic volatility earn low expected returns (henceforth referred to as the AHXZ result). They report that the difference between the average returns earned by the highest and the lowest idiosyncratic volatility quintile portfolios is -1.06% per month. Ang et al. (2009) provide evidence that this phenomenon holds true across 23 developed international markets.

Han and Lesmond (2011) approach the "puzzle" posed by AHXZ (2006) from a different perspective. They propose that microstructure influences on the estimation of idiosyncratic volatility lead to the AHXZ result. They show that the bid-ask spread and the percentage of zero returns biases the IVOL estimate in AHXZ (2006). After controlling for these biases, they find that the pricing ability of idiosyncratic volatility is reduced to insignificance.

George and Hwang (2012), henceforth referred to as GH2012, investigate the weaknesses of the AHXZ result reported by other studies (Bali and Cakici (2008), Huang et al. (2010), and Bali, Cakici and Whitelaw (2011)). They find that the weaknesses are mainly due to the effect of January returns and the influence of penny stocks. After controlling for these effects, the negative relation between high IVOL stocks and average returns holds true over return horizons up to two years after portfolio formation. They propose that the high IVOL stocks have low returns because disagreement among traders following news shocks results in optimistic mispricing, which is later corrected.

On the other hand, Fu (2009) questions the construction of lagged idiosyncratic volatility used in AHXZ (2006). He asserts that idiosyncratic volatility is time-varying and the one-month lagged (i.e., realized) value is not a good proxy for the expected value. So this measure should not be used to study the relation between expected returns and idiosyncratic risk. Fu (2009) employs forecasts of idiosyncratic volatility based on an exponential generalized autoregressive conditional heteroskedasticity (EGARCH) model and finds that the forecasted idiosyncratic volatility is positively related to the cross-section of expected returns. He documents that a zero investment portfolio that is long in the 10% of the highest and short in the 10% of the lowest forecasted idiosyncratic volatilities earns a significantly high positive return of 1.75% per month. However he does not verify whether these results hold true after controlling for the microstructure effects highlighted in Han and Lesmond (2011).

Liquidity may affect stock returns in several ways. It may directly affect returns via the return equation, or it may indirectly affect returns by influencing the estimate of idiosyncratic volatility. This paper elucidates whether idiosyncratic volatility can explain the cross-section of expected returns even after accounting for the liquidity biases embedded in its estimate. In my work, I use the Fu (2009) measure of forecasted idiosyncratic volatility (FIVOL) to explain expected returns. Univariate tests indicate that stock returns are positively related to idiosyncratic volatility. The relationship between idiosyncratic volatility and bid-ask spread is shown to be positive and so is the relationship between idiosyncratic volatility and the percentage of zero returns. The high correlation between spread and idiosyncratic volatility leads us to question whether the relation between returns and idiosyncratic volatility is driven by the bid-ask spreads. A Fama-MacBeth regression of realized idiosyncratic volatility on the microstructure variables indicates that 38.99% of the cross-sectional variation in realized idiosyncratic volatility can be explained by the spreads and the zero returns. As opposed to this, the influence of the microstructure variables on FIVOL is less pronounced and I find that these variables explain only 9.28% of the cross-sectional variation in the FIVOL estimate.

I also study whether the bid-ask spreads and the percentage of zero returns directly affect returns via the return equation. Fama-MacBeth cross-sectional regressions that examine the relationship between average returns and firm size, book-to-market ratio, idiosyncratic volatility and the liquidity variables are employed. The results indicate that FIVOL plays an important role in explaining returns, even after I account for the microstructure variables. However these variables do not have significant explanatory power when used in conjunction with FIVOL. The results are robust to sub-period analysis. The explanatory power of FIVOL decreases when I exclude the penny stocks and the January returns from the sample, but it is still positive and significant. This positive relation is robust to accounting for the Pastor and Stambaugh (2003) systematic liquidity factor and the Carhart momentum factor.

Interestingly, I find that the positive relation between FIVOL and expected returns does not hold true for a longer return horizon. I decompose FIVOL at t, $FIVOL_t$, into the forecasted IVOL at $t-1$ and the "innovation", and I study how these two measures affect average returns. I find that the forecasted IVOL estimate at $t-1$, $FIVOL_{t-1}$, does not have a significant relationship with expected returns. The "innovation", $Innov$, which is a manifestation of the return shock between $t-2$ and $t-1$, drives the positive relation.

To highlight the importance of these return shocks, I study the impact of changes in FIVOL on the cross-section of expected returns. I find that a portfolio composed of stocks that move from a low FIVOL quintile at $t-1$ to the highest FIVOL quintile at t, earns a very high return in month t. In contrast to this, a portfolio composed of stocks that fall from a higher FIVOL quintile to the lowest FIVOL quintile earns a negative return. Hence, the transitions in FIVOL quintile ranking are an extremely important factor in explaining the relation between FIVOL and expected returns. Models in the existing literature, such as Levy (1978) and Merton (1987), propose that the cross-sectional differences in levels of idiosyncratic volatility are priced. However, contrary to these models, I find that the transitions in idiosyncratic volatility drive the differences in the cross-section of expected returns.

I find that FIVOL and lagged realized IVOL cannot be regarded as substitutes and when used in conjunction, FIVOL increases the significance of the negative relation between lagged IVOL and expected returns. This implies that these two measures of idiosyncratic volatility convey different information. Experiments using a dummy variable to indicate a negative or no transition in FIVOL quintile ranking show that the results in Fu (2009) are mainly driven by the stocks that move from a lower FIVOL quintile to a higher one. A dummy variable that indicates a negative transition in IVOL quintile ranking shows that the result in AHXZ (2006) is driven by the stocks that move from a higher to a lower IVOL quintile.

I then investigate the reason behind these return shocks that cause the sudden changes in FIVOL and explore whether earnings surprises drive the pricing ability of idiosyncratic volatility. I use dummy variables that indicate extreme standardized unexpected earnings and find that the relation between FIVOL and expected returns is stronger for stocks with large positive earnings surprises and weaker (or less positive) for stocks with negative earnings surprises. Furthermore, I find that FIVOL is positively related to expected returns even after excluding the stocks that have extreme positive, negative or no earnings surprises. These results show that though earnings surprises contribute to the positive relation between FIVOL and the cross-section of expected returns, standardized unexpected earnings (SUE) cannot completely explain this relation.

This paper also explores the relation between transitions in FIVOL ranking and expected returns for the stocks in the most positive and the most negative SUE quintile. After controlling for the level of SUE, I find that these transitions still matter. The results show that the pricing of the transitions in FIVOL ranking is not driven by earnings surprises.

The rest of the paper is organized as follows: Section 2.2 discusses the data and the empirical methodology, and then estimates the one-month-ahead FIVOL using an EGARCH model. In section 2.3, I explore the relationship between idiosyncratic volatility and the microstructure variables. This section also investigates the relationship between idiosyncratic volatility, liquidity and future returns. The effect of controlling for January returns and penny stocks is also examined. Section 2.4 studies the transitions in FIVOL quintile ranking and its effect on expected returns in the cross-section. Section 2.5 investigates the role of standardized unexpected earnings in the relation between idiosyncratic volatility and the cross-section of expected returns. Section 2.6 concludes.

2.2 Data and methodology

2.2.1 Data

The data consist of daily and monthly prices, returns and other firm characteristics of the NYSE, Amex and NASDAQ companies covered by CRSP from January 1983 to December 2006. Price, return and volume data are obtained from CRSP. Financial information is obtained from Compustat. The daily factor data for the Fama-French three factor model are obtained from Kenneth French's website.

The Trades and Quotes (TAQ), the Institute for the Study of Security Markets (ISSM), and the CRSP databases are used to estimate the proportional spreads. The ISSM database is used for the trades and quotes data for all NYSE and Amex firms from January 1983 to December 1992. I utilize the CRSP and the TAQ databases for NYSE, Amex and NASDAQ firms from January 1993 to December 2006 to complete the sample.

Following Han and Lesmond (2011), henceforth referred to as HL2011, I define the proportional spread as the ask quote minus the bid quote divided by the quote midpoint. For each firm i, I average the daily proportional spread over each month to calculate the monthly spread. The percentage of zero returns for each month t is also obtained from CRSP. It is given by the number of zero return days in a month divided by the total number of trading days in the same month.

2.2.2 Estimating Fama-French based idiosyncratic volatility

Idiosyncratic risk is defined as the risk that is endemic to a particular asset. It is independent of the common movement of the market. Following AHXZ (2006), the idiosyncratic volatility of each stock is estimated relative to the Fama-French three-factor model:

$$R_{it}^d - R_{ft}^d = \alpha_{it} + \beta_{mkt,it}(R_{mkt,t}^d - R_{ft}^d) + \beta_{smb,it}R_{smb,t}^d + \beta_{hml,it}R_{hml,t}^d + \epsilon_{it}^d \qquad (2.1)$$
$$\epsilon_{it}^d \sim N(0, \varsigma_{it}^2)$$

where R_{it}^d is the daily return for firm i on day d of month t, $R_{mkt,t}^d$ is the excess daily return on a broad market portfolio, $R_{smb,t}^d$ is the daily average return on the three small portfolios minus the average return on the three big portfolios, $R_{hml,t}^d$ is the daily average return on the two value portfolios minus the average return on the two growth portfolios, and R_{ft}^d is the daily risk-free rate. For each stock i, I perform the time-series regression given by Equation 2.1 within each month. The realized idiosyncratic volatility is then defined as the standard deviation of the regression residuals or as $\sqrt{((Var(\epsilon_{it})))}$, and is denoted by IVOL.

AHXZ (2006) compute $IVOL_t$ from Equation (2.1) over a one month period from $t-1$ to t. At t, they construct value-weighted portfolios based on these idiosyncratic volatilities and hold these portfolios for the next one month. They find that the difference in average raw returns between the highest and the lowest IVOL quintile portfolios in month $t+1$ is -1.06%. In their paper, they use the estimate of the one-month lagged realized idiosyncratic volatility to find the relationship between idiosyncratic risk and expected returns. Fu (2009) suggests that investors should be compensated for bearing risk in the same period. In other words, if a stock has high idiosyncratic risk during month $t+1$ (i.e., $IVOL_{(t,t+1)}$ is high), then as a premium for bearing this risk, the returns earned by investors during month $t+1$ (i.e., $R_{(t,t+1)}$) should be high. He indicates that the research design in AHXZ (2006) is flawed since IVOL computed over $t-1$ to t (i.e., $IVOL_{(t-1,t)}$), may not be a good estimate of $IVOL_{(t,t+1)}$ and should not be used to draw an inference on the relation between idiosyncratic risk and expected returns. He emphasizes the need to find a better estimate for future idiosyncratic risk. The next section describes such a model that has been used in the literature to forecast idiosyncratic volatility.

2.2.3 Forecasting idiosyncratic volatility

Researchers have often used various autoregressive conditional heteroskedasticity (ARCH) models to estimate volatility. These autoregressive models capture the time-varying property of idiosyncratic volatility and may be used to forecast out-of-sample idiosyncratic volatility. Following Fu (2009) and Spiegel and Wang (2005), I use an EGARCH model which is estimated as follows.

Idiosyncratic risk is estimated using the three Fama-French factors as proxies for systematic risk in monthly returns:

$$R_{it} - R_{ft} = \alpha_i + \beta_{mkt,i}(R_{mkt,t} - R_{ft}) + \beta_{smb,i}R_{smb,t} + \beta_{hml,i}R_{hml,t} + \epsilon_{it} \tag{2.2}$$

where R_{it} is the return on stock i in month t and R_{ft} is the monthly risk free rate. $R_{mkt,t}$, $R_{smb,t}$ and $R_{hml,t}$ are the excess monthly market return, the size premium and the value premium, respectively. The idiosyncratic return, ϵ_{it}, is assumed to be drawn from a normal distribution defined by $\epsilon_{it} \sim (0, \sigma_{it}^2)$, where the conditional variance is described by the following EGARCH(p,q) process

$$ln\ \sigma_{it}^2 = a_i + \sum_{l=1}^{p} b_{i,l}\ ln\ \sigma_{i,t-l}^2 + \sum_{k=1}^{q} c_{i,k} \left\{ \Theta \left(\frac{\epsilon_{i,t-k}}{\sigma_{i,t-k}} \right) + \gamma \left[\left| \frac{\epsilon_{i,t-k}}{\sigma_{i,t-k}} \right| - \sqrt{\left(\frac{2}{\pi} \right)} \right] \right\} \tag{2.3}$$

Investors incorporate the newly revealed surprises in returns into their estimates of the mean and the variance of returns in the next period. This behavior of investors can be modeled in the EGARCH and these models are well suited to accommodate any asymmetric effect in the evolution of the volatility process. Thus, for $\gamma, b > 0$, large price changes are still followed by large price changes, but with $\Theta < 0$, this effect is accentuated for negative price changes, a stylized feature of equity returns often referred to as the "leverage effect". Moreover unlike the GARCH models, no restrictions are imposed on the parameter values in an EGARCH model to ensure positive values of variance. I use different EGARCH models with values of $p = 1, 2, 3$ and $q = 1, 2, 3$ and the estimate generated by the model with the maximum log-likelihood is chosen. Firms with less than 30 monthly returns are excluded from the sample.

The conditional idiosyncratic volatility estimated from the EGARCH model is called the forecasted idiosyncratic volatility and is denoted by $FIVOL$. It has a mean of 13.95% and a standard deviation of 7.74%. The correlation between $IVOL$ and $FIVOL$ is 0.51 and statistically significant at the 1% level.

2.2.4 Systematic risk factors

Fama and French (1992) document that firm size and the ratio of the book value of equity to the market value of equity are important characteristics that are related to expected returns. I calculate the beta, the size and the book-to-market ratio of each firm based on Fama and French (1992). To ensure that accounting variables are known before they are used to explain the cross-section of stock returns, these characteristics are calculated in June of each year, and used from July of that year until June of the following year.

In June of each year, I form 10 size portfolios based on the market capitalization of all stocks traded on the NYSE. Then for each firm-month observation, the preceding 60 months of returns is used to estimate the pre-ranking β based on the market model. Stocks in each size decile are then assigned to 10 portfolios based on their pre-ranking β. The equal-weighted monthly returns for the next 12 months of these 100 portfolios, which are formed on the basis of size and pre-ranking β, are calculated. $BETA$ is estimated as the sum of the slopes in the time-series regression of the portfolio return on the contemporaneous and the prior month's value-weighted market returns. Each stock in a specific size-β portfolio is assigned the $BETA$ of that portfolio.

The market value of equity (ME), which is given by the product of the monthly closing price of a stock and the number of shares outstanding, is used to define firm size. Book-to-market equity (BE/ME) is defined as the ratio of the book value of equity in the month of June to the market value of equity in the month of December. Since ME and BE/ME have substantial skewness, they are transformed to their natural logarithm in the cross-sectional tests.

2.2.5 Summary Statistics

Table 2.1 reports the summary statistics for the variables used in this paper. The numbers reported are time-series averages of the cross sectional mean, standard deviation, median, lower quartile, upper quartile and skewness of each variable. It is evident from the table that firm size and book-to-market ratio display considerable skewness.

Table 2.2 reports time-series averages of the cross-sectional correlations between the variables used in this paper. Correlation coefficients that are significant at the 1% level are marked with ***. The correlation between monthly returns and the forecasted idiosyncratic volatility, $FIVOL$, is 13% and is statistically significant at the 1% level. The correlation between monthly returns and the contemporaneous idiosyncratic volatility, $IVOL$, is 14% and is also significant at the 1% level. The high correlation of 30% between spread and $IVOL$ implies that the contemporaneous idiosyncratic volatility may have an embedded component of spread in its estimate. However, the correlation between the forecasted idiosyncratic volatility and spread is much lower (8%, but still significant).

The results in Table 2.2 also indicate that spread is inversely proportional to firm size. In accordance with Fama and French (1992), I find that the book-to-market ratio is positively correlated with returns, whereas the size of the firm is negatively correlated with returns.

The results from these univariate tests suggest that the relation between $FIVOL_t$ and R_t is positive. Furthermore, the relation between realized idiosyncratic volatility at time t, $IVOL_t$ and R_t is also positive. The next section studies the relation between idiosyncratic risk and expected returns after accounting for the various control variables.

2.3 Results

2.3.1 Relation between returns and idiosyncratic volatility

This section employs the Fama and MacBeth (1973) methodology to explore the relation between returns and idiosyncratic volatility. I estimate a model that is nested in the following cross-sectional regression.

$$
\begin{aligned}
R_{i,t} = \alpha_{0t} + \beta_{1t}\ BETA_{i,t-1} + \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ FIVOL_{i,t} \\
+ \beta_{5t}\ IVOL_{i,t} + \beta_{6t}\ IVOL_{i,t-1} + e_{i,t}
\end{aligned}
\tag{2.4}
$$

where $i = 1, 2, ..., N_t$ and $t = 1, 2, 3, ..., T$. The time-series means and the Newey-West (1987) t-statistics of the parameter estimates are reported in Table 2.3. Model 1 replicates the main findings in Fama and French (1992) for this sample and indicates that firm size and book-to-market ratio are important determinants of cross-sectional returns. Consistent with the prior literature, the results find a flat relation between expected returns and beta, a negative relation between returns and size and a positive relation between returns and the book-to-market ratio.

Models 2 and 3 indicate that the forecasted idiosyncratic volatility as well as the contemporaneous idiosyncratic volatility is positively related to average returns in the cross-section. Model 5, which includes the Fama-French explanatory variables and $FIVOL$, shows that $FIVOL$ is still positively priced and this model has a higher R^2 (4.17%) than model 1. In model 6, I replace $FIVOL$ with the lagged idiosyncratic volatility, $IVOL_{t-1}$, and find that the relationship between lagged idiosyncratic volatility and returns is negative and has a low t-statistic. Model 7 indicates that the contemporaneous idiosyncratic volatility, $IVOL_t$, is positively related to returns and the coefficient is highly significant.

These results indicate that the relationship between contemporaneous idiosyncratic risk and returns is positive. Han and Lesmond (2011) propose that the microstructure variables, embodied by the incidence of zero returns and the proportional spread, bias the AHXZ (2006) estimate of idiosyncratic volatility. They show that the negative relation between $IVOL_{t-1}$ and expected returns is driven by these liquidity biases. In the next section, I explore the mechanism through which these variables affect the different idiosyncratic volatility estimates used in this paper.

2.3.2 Relation between different measures of idiosyncratic volatility, bid-ask spread and zero returns

Preliminary evidence from the correlation analysis indicates that there is a strong positive correlation between contemporaneous idiosyncratic volatility and the liquidity components. This section takes a more rigorous look at the issue. Panel A of Table 2.4 reports the coefficients from a Fama-MacBeth based regression test of the contemporaneous idiosyncratic volatility on the proportional spread and the percentage of zero returns. I estimate a model that is nested in the following cross-sectional regression.

$$IVOL_{i,t} = \alpha_{0t} + \alpha_{1t}\,\%Zeros_{i,t} + \alpha_{2t}\,Spread_{i,t} + \alpha_{3t}\,Spread^2_{i,t} + \alpha_{4t}\,\%Zeros_{i,t} * Spread_{i,t} + \upsilon_{i,t} \tag{2.5}$$

The coefficients reported in Table 2.4 are the mean coefficients estimated from the above regression. From the results in this table it is evident that spread alone explains a very significant portion (24%) of the cross-sectional variation in idiosyncratic volatility. The squared spread and the percentage of zero returns are also highly significant. Model 1, which accounts for the spread, the squared spread, the percentage of zero returns and the term representing the interaction of spread and zero returns explains 38.99% of the cross-sectional variation in idiosyncratic volatility.

The initial results from Table 2.2 suggest that the correlations between the forecasted idiosyncratic volatility and these microstructure variables are weaker than the correlations between realized idiosyncratic volatility and these variables. To explore the relationship between $FIVOL$ and liquidity further, I carry out a Fama-MacBeth based regression test of $FIVOL$ on the liquidity components.

Here I consider two possibilities. $FIVOL_{it}$ is the forecasted idiosyncratic volatility for stock i in month t, conditional on the information set available till month $t-1$. So we might expect $FIVOL_{it}$ to be related to the bid-ask spread and the percentage of zero returns in $t-1$. I estimate the following cross-sectional regression,

$$\begin{aligned} FIVOL_{i,t} = \alpha_{0t} + \alpha_{1t}\%Zeros_{i,t-1} + \alpha_{2t}Spread_{i,t-1} + \alpha_{3t}Spread^2_{i,t-1} \\ + \alpha_{4t}\%Zeros_{i,t-1} * Spread_{i,t-1} + \upsilon_{i,t} \end{aligned} \tag{2.6}$$

The results in Panel B of Table 2.4 indicate that the lagged spread and lagged percentage of zero returns do not explain a significant portion of the cross-sectional variation in the estimate of forecasted idiosyncratic volatility (R^2=9.28%).

The results in Panel A show that the microstructure variables explain a significant portion of the cross-sectional variation in contemporaneous IVOL. But we know that contemporaneous IVOL is highly correlated with FIVOL. So we might expect that the forecasted IVOL estimate in month t is related to the spread and the percentage of zero returns in t. However, similar regressions of $FIVOL_t$ on the microstructure variables at t have an even lower value of R^2. These results indicate that $FIVOL$ is a cleaner measure of idiosyncratic volatility than the AHXZ (2006) measure.

2.3.3 Relation between returns, idiosyncratic volatility and the liquidity components

The previous results indicate that the forecasted idiosyncratic volatility is an important factor in explaining returns. Moreover, Table 2.4 indicates that the proportional spread and the percentage of zero returns explain a substantial portion of the realized idiosyncratic volatility, but not of the forecasted volatility. Nonetheless, to have a better knowledge of the interaction between idiosyncratic volatility and the microstructure variables, a model that includes the liquidity components in the return equation is considered.

Jegadeesh (1990) documents that the first-order autocorrelation in monthly stock returns for individual firms is negative and highly significant. Huang et al.(2009) show that the AHXZ (2006) result is driven by an omitted variable bias because the authors do not explicitly control for the return reversals. They report that the relationship between R_t and $IVOL_{t-1}$ is no longer significant after accounting for return reversals. Hence, I include the previous month's return to investigate whether the positive relation between FIVOL and expected returns holds true after controlling for return reversals. I estimate different specification models nested in the following cross-sectional regression

$$\begin{aligned} R_{i,t} = \alpha_{0t} + \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ FIVOL_{i,t} + \beta_{5t}\ IVOL_{i,t-1} \\ + \alpha_{1t}\ \%Zeros_{i,t} + \alpha_{2t}\ Spread_{i,t} + \alpha_{3t}\ R_{i,t-1} + e_{i,t} \end{aligned} \tag{2.7}$$

Table 2.5 reports the time-series averages of the slopes in these regressions. A comparison of models 3 and 7 elucidates that the inclusion of *Spread* and $\%Zeros$ strengthens the explanatory power of the lagged idiosyncratic volatility (the AHXZ measure). The t-statistic for the coefficient on $IVOL_{t-1}$ increases from -1.58 in model 3 to -2.55 in model 7. This shows that the negative relation between lagged $IVOL$ and returns is not caused by an omitted variable bias due to the exclusion of these microstructure variables. I also find that the percentage of zero returns is not significant in explaining returns in any of these models.

Next, I study the relation between $FIVOL_t$ and expected returns after controlling for *Spread* and *%Zeros*. Model 5 shows that the forecasted idiosyncratic volatility is still positively related to the cross-section of average returns. The relationship between spread and expected returns is not significant in this model. I find that the microstructure variables lose their explanatory power when FIVOL is included in the regression [1].

To check the robustness of these results, I divide the sample into two sub-periods, from January 1983 to December 1994 and from January 1995 to December 2006. HL2011 report a decline in the percentage of zero returns since the change in tick size in 1997 and a marked decline in bid-ask spreads after decimalization in 2001. If microstructure variables are related to expected returns, we would expect to find a significant difference between the results during the two sub-periods. However, I find that the results (not reported) are similar across these two sub-periods.

Models 5 and 7 show that both realized $IVOL$ and $FIVOL$ have a significant relation with expected returns in the cross-section. Models 12 and 13 include both of these measures of idiosyncratic risk simultaneously. Fama-MacBeth regressions of the expected returns on $IVOL_{t-1}$ and $FIVOL_t$ along with the other control variables show that both of these measures are statistically significant. However their effects on returns act in opposite directions. The parameter estimates on $IVOL_{t-1}$ and $FIVOL_t$ are bigger when they are included in the same regression. The parameter estimate on FIVOL increases from 0 .14 (t=8.05) in model 1 to 0.17(t=10.98) in model 12. Thus an increase in FIVOL results in an increase in expected returns. The coefficient on lagged IVOL changes from -0.02 (t=-1.58) in model 3 to -0.06(t=-7.00). The results show that the significance of lagged IVOL increases considerably when it is used in conjunction with FIVOL. Thus both FIVOL and lagged IVOL

[1] Pastor and Stambaugh (2003) find that a stock's liquidity beta, which represents the sensitivity of stock returns to innovations in aggregate liquidity, plays an important role in predicting returns. Stocks with higher sensitivity to aggregate liquidity shocks, have higher expected returns. To examine whether the high expected returns earned by high FIVOL stocks can be attributed to the premium demanded by low liquidity betas, I estimate idiosyncratic risk based on a five-factor model that includes the momentum factor and the Pastor and Stambaugh (2003) aggregate liquidity factor. The estimated conditional idiosyncratic volatility from this model is denoted by $FIVOL\_5factor$. The post-ranking alphas of value-weighted portfolios formed on the basis of $FIVOL\_5factor$ are reported in Table A1 which shows that the positive return on the high minus low FIVOL portfolio is not a result of the pricing of systematic liquidity risk. Fama-MacBeth methodology is also employed to study the role of $FIVOL\_5factor$ in explaining returns. The results are reported in Table A2. The parameter estimate on $FIVOL\_5factor$ is positive and significant. Overall, the evidence strongly rejects the hypothesis that the pricing ability of forecasted idiosyncratic volatility is attributable to the aggregate liquidity risk premium.

have significant pricing ability. Previous literature has treated the AHXZ (2006) and the Fu (2009) methodology of computing idiosyncratic volatility as two competing methods. However, this test shows that IVOL and FIVOL are two distinct measures that convey different information and affect expected returns in different ways.

GH2012 show that after controlling for penny stocks and January returns, the negative relation in AHXZ (2006) holds true over return horizons up to two years. Thus it would be interesting to explore whether the positive relation between $FIVOL$ and returns is valid for a longer horizon. Models 2, 6 and 10 in Table 2.5 indicate that $FIVOL_{t-1}$ does not play a significant role in explaining the expected returns in month t. This indicates that the unexpected change in $FIVOL$ is positively related to the next month's returns and this change is driving the strong positive relation between $FIVOL_t$ and returns. These results imply that the main factor that causes the pricing ability of $FIVOL_t$ is the contribution of the return shocks between $t-2$ and $t-1$ to the conditional variance estimate in the EGARCH model.

To substantiate the premise that the new information between $t-2$ and $t-1$ causes the positive relation between forecasted idiosyncratic volatility and returns, I decompose $FIVOL_t$ into $FIVOL_{t-1}$ (which is dependent on information till $t-2$) and the innovation, $Innov$, which embodies the return shocks between $t-2$ and $t-1$. I find that $Innov$ is positively related to returns. This shows that though $IVOL_{t-1}$ and $Innov_t$ are based on the same time period, i.e., $t-2$ to $t-1$, they have opposite effects on the cross-section of expected returns. This implies that the estimates of $IVOL$ and $FIVOL$ do not capture the same information. The AHXZ measure of idiosyncratic volatility used in this paper is calculated over daily data and is based on the volatility of daily returns over the past one month. On the other hand, the forecasted idiosyncratic volatility, $FIVOL$, is calculated over monthly returns and is dominated by the innovation, $Innov$, which is a squared monthly return shock. I find that the risk premium associated with $Innov$ induces the strong positive relation in the Fu (2009) results.

2.3.4 Controlling for January and penny stocks

Tax-loss selling in the month of December causes a drop in prices that results in high returns in the month of January. GH2012 show that these high January returns conceal the true relationship between $IVOL$ and the cross-section of expected returns. They report that an equally weighted portfolio consisting of stocks that belong to the highest $IVOL$ quintile earns a negative return when the January returns are excluded. It would be interesting to investigate whether the positive relation between $FIVOL$ or $Innov$ and expected returns is driven by the January effect. Table A3 reports the results of the Fama-Macbeth regression given in Equation 2.7. I find that the positive relation is robust to controlling for January returns. I find that the exclusion of January weakens the explanatory power of the microstructure variables. The explanatory power of $FIVOL_t$ attenuates once we control for January, but it still remains highly significant.

GH2012 highlight the role of penny stocks in concealing the true relation between idiosyncratic volatility and returns. They propose that the high illiquidity of these stocks adds noise and biases the IVOL rankings and measured returns. Following their method, I exclude stocks whose prices are less than $5 at the end of the portfolio formation month. I also exclude the January returns and notice a substantial change in the results. These results are shown in Table 2.6.

I find that in Model 1, the parameter estimate on $FIVOL_t$ drastically reduces from 0.14 in Table 2.5 to 0.04 and the t-statistic reduces to 2.61. The coefficient on $FIVOL_{t-1}$ becomes more negative and highly significant (t-statistic=-2.84). Following the exclusion of January and penny stocks, the parameter estimate on spread becomes negative, but still remains insignificant. However, the striking result in this table is that the coefficient estimate on $Innov$ still remains positive and highly significant. The drastic changes in the coefficient estimates that is observed for $FIVOL_t$ and $IVOL_{t-1}$ is not observed for $Innov$. This shows that the pricing ability of $Innov$ is not driven by a subset of stocks that has special characteristics or by the returns in a specific month.

2.4 Changes in idiosyncratic volatility and its effect on returns

The results in Tables 2.5 and 2.6 suggest that $Innov$, which is the contribution of the information available between $t-2$ and $t-1$ to the $FIVOL_t$ estimate, plays a crucial role in driving the positive relation between forecasted idiosyncratic volatility and expected returns. $Innov_t$ represents the difference between $FIVOL_{t-1}$ and $FIVOL_t$ and hence it denotes the change in $FIVOL$ between the two successive months. Thus it would be interesting to study whether the change in a firm's FIVOL quintile ranking explains the high returns earned by stocks in high FIVOL portfolios.

Saryal (2009) shows that the change in a firm's realized idiosyncratic volatility ranking can explain AHXZ's puzzling result. She finds that the firms that move from a low IVOL quintile to a higher IVOL quintile earn very high positive returns. The firms that move from a high IVOL quintile to a lower IVOL quintile earn negative returns. For stocks that have a highly persistent level of IVOL and remain in the same IVOL quintile, a positive relation exists between IVOL and future returns. However, she uses the realized IVOL in her paper and the changes in IVOL quintile ranking are only available ex-post.

In this section, I study the changes in a firm's FIVOL quintile ranking between two successive periods. For each month t, I sort stocks on the basis of their $FIVOL_t$. I define a variable called $Migrate$ which indicates the movement of stocks from one FIVOL quintile to another, and $Migrate(t)$ equals a firm's FIVOL quintile rank at t minus its FIVOL quintile rank at $t-1$. Accordingly, $Migrate$=4 indicates that the firm was in the lowest FIVOL quintile in month $t-1$ and is in the highest FIVOL quintile in month t. Similarly, $Migrate$=-4 indicates a jump from the highest FIVOL quintile to the lowest FIVOL quintile. $Migrate$=0 indicates that the firm is in the same FIVOL quintile in month $t-1$ and t.

Table 2.7 shows the distribution of the $Migrate$ variable. We see that 82.8% of all Quintile 1 firms stay in the same quintile, whereas 71.8% of all Quintile 5 firms stay in the same quintile. On an average across all quintiles, 68.7% of stocks remain in the same FIVOL quintile between successive time periods.

I form 25 equal-weighted portfolios based on the FIVOL quintile in month $t-1$ and t. Next, I study the firms that are in the highest quintile portfolio in month t. Table 2.7 shows that only 28.2% of the firms in the highest FIVOL quintile in month t have moved from lower quintile portfolios. However, these stocks earn significant positive returns. Unadjusted monthly returns vary across the *Migrate* portfolios. The *Migrate*=4 portfolio earns an average monthly return of 6.66%, whereas the *Migrate*=0 portfolio, which is composed of firms that are persistently in the highest FIVOL quintile, earns an average return of 1.64%. The positive CAPM and FF-3 factor alphas of the stocks in the highest FIVOL quintile are an outcome of the high alphas earned by the stocks that move from the lower FIVOL quintiles to Quintile 5. The *Migrate*=4 portfolio earns a CAPM alpha of 5.62% (t=6.76) and a FF-3 factor alpha of 5.39% (t=6.67). The CAPM alpha of the *Migrate*=0 portfolio is 0.67% and is insignificant, and the FF-3 factor alpha of this portfolio is 0.82%. This shows that the positive returns for stocks in the highest FIVOL quintile are driven by the high positive returns earned by stocks that have moved from a lower FIVOL quintile to this quintile.

The results are different when we consider the firms that are in the lowest quintile portfolio in month t. A transition from a higher FIVOL quintile to the lowest FIVOL quintile indicates a decrease in uncertainty about the firm. Table 2.7 shows that only 17.2% of the firms in the lowest FIVOL quintile in month t have moved from a higher quintile. The stocks in the lowest FIVOL quintile have a greater tendency to remain in the same quintile in the successive period than the stocks in the highest FIVOL quintile. In other words, persistent FIVOL quintile ranking is more common in the stocks in the lowest FIVOL quintile than in the stocks in the higher FIVOL quintiles. Table 2.9 reports the returns earned by the portfolios of stocks that are in the lowest FIVOL quintile in month t. The *Migrate*=-4 portfolio, which consists of firms that have moved from the highest FIVOL quintile in month $t-1$ to the lowest FIVOL quintile in month t, earns a return of -1.98% in month t. The CAPM alpha of this portfolio is -2.64% (t=-9.14) and the FF-3 factor alpha is -2.71% (t=-10.66). On the contrary, *Migrate*=0 portfolio, which is composed of stocks that are in the lowest FIVOL quintile in month $t-1$ and t, earns an insignificant

FF-3 factor alpha. Table 2.9 shows that the overall negative returns earned by stocks in the lowest FIVOL quintile is a consequence of the significant negative returns earned by the stocks that move from a higher FIVOL quintile to this quintile.

In general, the results in this section document that the pricing ability of forecasted idiosyncratic volatility is driven by stocks that have a transition in their FIVOL quintile ranking between month $t-1$ and t.

2.4.1 Changes in IVOL ranking and expected returns

In this section, I study the relationship between changes in IVOL ranking and expected returns. In each month t, I sort stocks into quintiles based on $IVOL_t$. I define a variable called $Migrate\_IVOL$ which indicates the movement of a firm from one IVOL quintile to another, and $Migrate\_IVOL(t)$ equals the firm's IVOL quintile rank at t minus its IVOL quintile rank at $t-1$. Similar to the previous section, I form 25 portfolios based on IVOL ranking in month $t-1$ and t. Table 2.10 reports the returns for stocks that have migrated from the highest IVOL quintile to lower quintiles. The $Migrate\_IVOL$=-4 portfolio consists of stocks that belong to Quintile 5 in month $t-1$ and Quintile 1 in month t. These firms earn a negative return of -0.33% (t=-2.68) and a FF-3 factor alpha of -0.55% (t=-4.78). I find that all portfolios with negative $Migrate\_IVOL$ earn significant negative returns. Thus the stocks that fall from the highest IVOL quintile in month $t-1$ to lower IVOL quintiles in month t earn significant negative returns. The $Migrate\_IVOL$=0 portfolio, composed of stocks that consistently belong to the highest IVOL quintile, earns a positive return of 2.62%.

Table 2.11 reports the returns for the portfolios of stocks that have migrated from the lowest IVOL quintile to higher IVOL quintiles. I find that the $Migrate\_IVOL$=4 portfolio earns a positive FF-3 factor alpha of 4.79% (t=5.55). Thus the stocks that move from the lowest IVOL quintile to the highest IVOL quintile earn very high returns. On the other hand, the stocks that consistently remain in the lowest IVOL quintile earn insignificant returns. These tables show that the pricing of the negative transitions in IVOL quintile ranking give rise to the overall negative relation between $IVOL_{t-1}$ and the cross-section of expected returns.

2.4.2 Explanatory power of migration dummy variables

The previous section shows that the lagged realized IVOL and the forecasted IVOL are driven by different components. Saryal (2009) shows that the movement of stocks from a low IVOL quintile to a high IVOL quintile is often accompanied by a large positive return. So when we relate the large positive returns to the IVOL of the preceding month (the lagged realized IVOL in AHXZ) , the low IVOL stocks seem to earn high returns in the cross-section.

When stocks are ranked by FIVOL, the stocks that have a jump in FIVOL and move from the lower FIVOL quintiles to the highest quintile earn high positive returns. A shock to the idiosyncratic return may cause a jump in FIVOL, which results in this positive return. It is this change in FIVOL that results in the overall positive relationship in the Fama-MacBeth based tests reported in Table 2.3. To study whether the results are driven solely by the stocks that move from the lower FIVOL quintiles to the higher quintiles, I employ cross-sectional regressions with dummy variables that indicate changes in FIVOL and IVOL rankings. The cross-sectional regression is described by

$$\begin{aligned} R_{i,t} = \alpha_{0t} + \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ MIG\_FIVOL_{i,t} * FIVOL_{i,t} \\ + \beta_{5t}\ FIVOL_{i,t} + \beta_{6t}\ MIG\_IVOL_{i,t} * IVOL_{i,t-1} + \beta_{7t}\ IVOL_{i,t-1} + \beta_{8t}\ R_{i,t-1} + e_{i,t} \end{aligned} \tag{2.8}$$

where $R_{i,t}$ is the return to stock i in month t, MIG_FIVOL is the FIVOL migration dummy that takes a value of 1 if the stock moves from a high FIVOL quintile to a low FIVOL quintile or remains in the same FIVOL quintile between month $t-1$ and t, and zero otherwise. MIG_IVOL is the IVOL migration dummy that takes a value of 1 if the stock moves from a high IVOL quintile at $t-1$ to a low IVOL quintile at t,and 0 otherwise. The results are shown in Table 2.12.

The specific contribution of stocks, which have a fall or no change in their FIVOL quintile ranking, to the overall relationship between FIVOL and expected returns can be identified by the coefficient estimates on the interaction term $MIG\_FIVOL*FIVOL$. The coefficient on $FIVOL_t$ gives the relation between FIVOL and expected returns for the other subset of stocks, which have MIG_FIVOL=0 . A coefficient of 0.19 (t=11.61) shows that for these stocks that move from a low FIVOL quintile at $t-1$ to a high FIVOL quintile at t, FIVOL is positively related to the cross-section of expected returns. A coefficient of -0.06 (t=-7.19) on $MIG\_FIVOL*FIVOL$ shows that for stocks which have a high to low FIVOL transition or no transition, the relationship between FIVOL and expected returns is less positive than the other stocks.

Similarly, the contribution of the stocks, which have a fall in their IVOL quintile ranking, to the overall relationship between IVOL and expected returns can be identified by the coefficient estimates on $MIG\_IVOL*IVOL$. The coefficient of -0.01 (t=0.73) shows that for stocks that have a rise or no change IVOL ranking, there is no relationship between IVOL and expected returns. The negative and highly significant coefficient of -0.13 on $MIG\_IVOL*IVOL$ indicates that the negative relation between realized IVOL and future returns is completely driven by the high to low transition in IVOL quintile ranking.

The results in Table 2.12 show that the transitions in idiosyncratic volatility ranking play a pivotal role in the relationship between idiosyncratic risk and expected returns.

2.5 Information Content of idiosyncratic Volatility

The results in Tables 2.8, 2.9 and 2.12 show that the positive relation between FIVOL and future stock returns is a result of the changes in FIVOL quintile ranking from one month to another. Section 2.3.3 shows that the idiosyncratic return shocks in the most recent month drive this relation. This section seeks to throw light on the underlying reason behind these return shocks. I investigate whether earnings surprises can explain the relation between FIVOL and expected returns.

Earnings are an important element of capital markets and may drive the changes in idiosyncratic volatility. It has been shown that stock prices respond to unanticipated changes in earnings, and there is a significant correlation between earnings surprises and future stock returns. In an efficient market, the information from a firm's current earnings should be quickly incorporated into its stock price. However, Ball and Brown (1968), Foster et al. (1984) and Bernard and Thomas (1989) show that stock prices continue to drift in the direction of an earnings surprise for three quarters. This concept of post-earnings-announcement-drift (PEAD) is consistent with the behavioral models in which prices react slowly to public news. The exisiting literature has shown that stock prices for individual firms react positively to earnings news but require several quarters to fully reflect the information contained in the earnings.

The stocks with highest unexpected earnings outperform the stocks with the lowest unexpected earnings, with the abnormal returns concentrated around earnings announcements. Frazzini and Lamont (2007) show that stocks earn higher returns during months when earnings are announced than during non-announcement months. Barber, George, Lehavy and Trueman (2013) document that there is a spike in IVOL during the announcement window that may be caused by the firm-specific information disclosed through earnings. They contend that the uncertainty over the nature of information to be revealed drives the higher returns in the earnings announcement months. Thus I also include an announcement dummy, which accounts for the returns due to earnings announcement in a certain month.

Standardized unexpected earnings measure the information content of quarterly earnings. Unexpected earnings for a company in quarter q is the difference between the most recently announced earnings and expected earnings which is given by the earnings in the same quarter of the previous year. The standardized unexpected earnings (SUE) for a stock is given by the unexpected earnings divided by the standard deviation of the quarterly unexpected earnings over the last two years.

It is interesting to study whether the relation between FIVOL and expected returns is driven by earnings surprises. I seek to answer this question using two approaches. First, I use Fama-MacBeth cross-sectional regressions to study the effect of positive and negative earnings surprises on the relation between FIVOL and expected returns. Next, I examine the relation between the transitions in FIVOL ranking and expected returns after controlling for SUE.

2.5.1 Fama-MacBeth regressions with dummy variables

In this section, I study the relation between FIVOL and average returns specifically for the stocks with extreme SUE and use dummy variables to account for positive, negative and negligible earnings surprises. The main cross-sectional regression specification I work with is given by

$$\begin{aligned} R_{i,t} = \alpha_{0t} &+ \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ FIVOL_{i,t} \\ &+ \beta_{7t}\ HIGH\_SUE_{i,t-1} * FIVOL_{i,t} + \beta_{8t}\ LOW\_SUE_{i,t-1} * FIVOL_{i,t} \\ &+ \beta_{9t}\ POS\_SUE_{i,t-1} * FIVOL_{i,t} + \beta_{10t}\ NEG\_SUE_{i,t-1} * FIVOL_{i,t} \\ &+ \beta_{11t}\ ANNOUN_{i,t-1} + e_{i,t} \end{aligned} \tag{2.9}$$

where $R_{i,t}$ is the return to stock i in month t, $HIGH\_SUE_{i,t-1}(LOW\_SUE_{i,t-1})$ is a dummy variable that equals one if the stock i is among the top (bottom) 20% of stocks in month $t-1$ when ranked by the absolute value of standardized unexpected earnings. Similarly, $POS\_SUE_{i,t-1}$ ($NEG\_SUE_{i,t-1}$) equals one if the stock i is among the top (bottom) 20% of stocks in month $t-1$ when ranked by the actual value of standardized unexpected earnings. $ANNON_{i,t-1}$ equals one if there is an earnings announcement in month $t-1$.

Positive as well as negative earnings surprises can be interpreted as information shocks, which may drive the relation between expected returns and FIVOL. To control for both positive and negative news, I use the absolute value of SUE in Model 1 in Table 2.13. The relation between FIVOL and expected returns for the stocks belonging to the highest quintile based on the absolute value of SUE is given by $(\beta_4 + \beta_7)$ and the relation between FIVOL and expected returns for the stocks belonging to the three middle quintiles based on absolute(SUE) is given by β_4. Thus the positive coefficient of 0.05 (t-statistic=4.8) on the interaction term shows that for stocks with high absolute SUE, the positive relation between FIVOL and expected returns is stronger.

On the other hand, for stocks belonging to the lowest absolute SUE quintile, the negative coefficient on $LOW\_SUE * FIVOL$ indicates that the relation between FIVOL and returns is weaker or less positive than for the other firms. The coefficient on this interaction term highlights the effect of no news and hence no information shocks on the relation between FIVOL and expected returns. The difference between the coefficients on these two interaction terms shows that extreme earnings surprises drive a part of the positive relation between FIVOL and returns. The coefficient of 0.09 (t=7.59) on $FIVOL$ shows that the relationship between FIVOL and returns is significant even after we exclude the stocks with extreme earnings surprises and no earnings surprises. These results show that though standardized unexpected earnings drive a part of the relation between FIVOL and expected returns, they do not completely explain the positive relationship.

In Model 2, I separate the effects of negative and positive earnings surprises and use dummies $NEG\_SUE_{i,t-1}$ and $POS\_SUE_{i,t-1}$ to determine the effect of negative and positive news on the pricing ability of FIVOL. I also include the $ANNOUN$ dummy to account for an earnings announcement in the previous month. The interaction term, $POS\_SUE * FIVOL$ has a coefficient of 0.15 (t=14.41), which shows that stocks belonging to the highest SUE quintile have a stronger positive relation between FIVOL and expected returns. This may be attributable to the PEAD phenomenon, which results in positive returns following positive earnings surprises. On the other hand, the coefficient on $NEG\_SUE * FIVOL$ suggests that stocks belonging to the lowest SUE quintile have a more negative relation between FIVOL and expected returns than the other stocks. I also find that stocks that belong to the lowest absolute(SUE) quintile, which represents stocks that have little or no earnings surprises, have a more negative relation between FIVOL and returns than other stocks. The coefficient of 0.11 (t=8.38) on $FIVOL$ in Model 2 highlights that FIVOL is positively related to expected returns even after excluding the effects of stocks that have extreme positive, extreme negative and no earnings surprises. This shows that the pricing ability of FIVOL is not completely induced by stocks with extreme earnings surprises.

2.5.2 Standardized Earnings Surprises and Migration

The results from Table 2.13 with POS_SUE and NEG_SUE dummies show that the stocks in the highest SUE quintile have a stronger positive relation between FIVOL and expected returns and stocks in the lowest SUE quintile have a less positive relation between FIVOL and expected returns, in the cross-section. It has been shown in Section 2.4 that the transitions in FIVOL quintile ranking, represented by the $Migrate$ variable, are the main reason behind the positive relationship between FIVOL and the cross-section of expected returns. In this section, I focus on the stocks that belong to the highest and the lowest SUE quintiles.

Panel A of Table 2.14 gives the distribution of the average raw returns, the CAPM alphas and the FF-3 factor alphas with respect to $Migrate$ for stocks belonging to the highest SUE quintile. A FF-3 factor alpha of 7.30% is earned by a portfolio composed of stocks that belong to the highest SUE quintile at $t-1$ and move from the lowest FIVOL quintile at $t-1$ to the highest FIVOL quintile at t. This is greater than the FF-3 factor alpha of 5.39% that is earned by the portfolio of all stocks that move from the lowest FIVOL quintile to the highest FIVOL quintile. The FF-3 factor alpha for stocks in the highest SUE quintile that move from the highest FIVOL quintile to the lowest FIVOL quintile is -1.46% . This is again higher than the FF-3 factor alpha for all stocks with $Migrate$=-4 given in Table 2.9. These differences in abnormal returns are attributable to the positive relation between positive earnings surprises and future returns. However, this table shows that even after controlling for the level of SUE, the transitions in FIVOL quintile ranking are still priced.

Panel B of Table 2.14 gives the returns for stocks in the lowest SUE quintile. I find that the abnormal return for the portfolio of stocks that move from the highest FIVOL quintile to the lowest FIVOL quintile is -2.38%. Thus the negative return earned by the portfolio of stocks that have a decrease in FIVOL ranking is lower for stocks with negative earnings surprises than for stocks with positive earnings surprises. Similarly, the positive return earned by the portfolio of stocks with $Migrate$=4 is higher for stocks with positive earnings surprises than for stocks with negative earnings surprises.

However, the relation between $Migrate$ and expected returns is similar to that in Tables 2.8 and 2.9. The results in Table 2.14 show that even after controlling for the level of standardized unexpected earnings, the transitions in FIVOL ranking still drive the relation between idiosyncratic volatility and the cross-section of expected returns.

Table 2.15 gives the returns for stocks in the lowest quintile ranked by the absolute value of SUE. These stocks have lower information shocks related to earnings surprises and if the relation between FIVOL and returns was completely driven by earnings surprises, then we would expect to find no relation between $Migrate$ and expected returns for stocks in this quintile. However, I find that the FF-3 factor alpha earned by the $Migrate$=4 portfolio is 4.64% and significant at the 5% level. This is smaller than the FF-3 factor alpha of 7.30% that was earned by this $Migrate$ portfolio for stocks in the highest positive SUE quintile. This shows that earnings surprises explain a part of the returns earned by the portfolio of stocks that have extreme transitions in FIVOL ranking. However, these results document that even for the subset of stocks that have minimal or no earnings surprises, there are substantial differences between the returns earned by the different $Migrate$ portfolios. Therefore, even in the absence of unexpected earnings, transitions in FIVOL ranking still result in the pricing ability of FIVOL.

2.6 Conclusion

In this paper, I study the relation between idiosyncratic volatility and expected returns. I document that the pricing ability of forecasted idiosyncratic volatility is not dependent on the embedded liquidity costs. Moreover, the microstructure variables, embodied by the bid-ask spread and the percentage of zero returns, lose their explanatory power when used in conjunction with FIVOL.

Earlier papers have regarded the AHXZ (2006) and the Fu (2009) measures of idiosyncratic volatility as substitutes or as competing measures. However, I find that the inclusion of FIVOL increases the significance of the relation between lagged IVOL and expected returns. This indicates that the two measures do not convey the same information.

I show that the positive relation between FIVOL and the cross-section of expected returns does not hold for return horizons longer than one month. I also find that the high positive returns earned by stocks that move from a low FIVOL quintile to a higher FIVOL quintile drive the positive relation between FIVOL and expected returns. Cross-sectional regressions using dummy variables for transitions in FIVOL quintiles demonstrate the importance of these low to high transitions in FIVOL.

I also study the effect of earnings surprises on the relation between FIVOL and expected returns. I find that even after controlling for the level of standardized unexpected earnings, the low to high transitions in FIVOL ranking still drive the positive relation between FIVOL and expected returns.

Theoretical models in the existing literature, such as Levy (1978) and Merton (1987), assert that the cross-sectional differences in levels of idiosyncratic volatility are priced. However, in this paper, I show that the transitions drive the differences in expected returns.

Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X., 2006. The crosssection of volatility and expected returns. The Journal of Finance, 61(1), pp.259-299.

Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X., 2009. High idiosyncratic volatility and low returns: International and further US evidence. Journal of Financial Economics, 91(1), pp.1-23.

Avramov, D., Chordia, T. and Goyal, A., 2006. Liquidity and autocorrelations in individual stock returns. The Journal of Finance, 61(5), pp.2365-2394.

Bali, T.G. and Cakici, N., 2008. Idiosyncratic volatility and the cross section of expected returns. Journal of Financial and Quantitative Analysis, 43(01), pp.29-58.

Bali, T.G., Cakici, N. and Whitelaw, R.F., 2011. Maxing out: Stocks as lotteries and the cross-section of expected returns. Journal of Financial Economics, 99(2), pp.427-446.

Ball, R. and Brown, P., 1968. An empirical evaluation of accounting income numbers. Journal of Accounting Research, pp.159-178.

Banz, R.W., 1981. The relationship between return and market value of common stocks. Journal of financial economics, 9(1), pp.3-18.

Barber, B.M., De George, E.T., Lehavy, R. and Trueman, B., 2013. The earnings announcement premium around the globe. Journal of Financial Economics, 108(1), pp.118-138.

Berk, J.B., 1995. A critique of size-related anomalies. Review of Financial Studies, 8(2), pp.275-286.

Bernard, V.L. and Thomas, J.K., 1989. Post-earnings-announcement drift: delayed price response or risk premium?. Journal of Accounting research, pp.1-36.

Bernard, V.L. and Thomas, J.K., 1990. Evidence that stock prices do not fully reflect the implications of current earnings for future earnings. Journal of Accounting and Economics, 13(4), pp.305-340.

Blume, M.E. and Stambaugh, R.F., 1983. Biases in computed returns: An application to the size effect. Journal of Financial Economics, 12(3), pp.387-404.

Campbell, J.Y., Grossman, S.J. and Wang, J., 1992. Trading volume and serial correlation in stock returns (No. w4193). National Bureau of Economic Research.

Fama, E.F. and French, K.R., 1992. The crosssection of expected stock returns. the Journal of Finance, 47(2), pp.427-465.

Fama, E.F. and French, K.R., 1993. Common risk factors in the returns on stocks and bonds. Journal of financial economics, 33(1), pp.3-56.

Fama, E.F. and MacBeth, J.D., 1973. Risk, return, and equilibrium: Empirical tests. The Journal of Political Economy, pp.607-636. Vancouver.

Foster, G., Olsen, C. and Shevlin, T., 1984. Earnings releases, anomalies, and the behavior of security returns. Accounting Review, pp.574-603.

Frazzini, A. and Lamont, O.A., 2007. The earnings announcement premium and trading volume. NBER working paper.

Fu, F., 2009. Idiosyncratic risk and the cross-section of expected stock returns. Journal of

Financial Economics, 91(1), pp.24-37.

George, T.J. and Hwang, C.Y., 2004. The 52week high and momentum investing. The Journal of Finance, 59(5), pp.2145-2176.

George, T. and Hwang, C.Y., 2012. Analyst Coverage and Two Volatility Puzzles in the Cross Section of Returns. Unpublished working paper, University of Houston.

Goetzmann, W.N. and Kumar, A., 2005. Why do individual investors hold under-diversified portfolios? Unpublished working paper, Yale School of Management.

Han, Y. and Lesmond, D., 2011. Liquidity biases and the pricing of cross-sectional idiosyncratic volatility. Review of Financial Studies, 24(5), pp.1590-1629.

Han, Y., Hu, T. and Lesmond, D.A., 2015. Liquidity Biases and the Pricing of Cross-Sectional Idiosyncratic Volatility Around the World. Journal of Financial and Quantitative Analysis, 50(06), pp.1269-1292.

Huang, W., Liu, Q., Rhee, S.G. and Zhang, L., 2010. Return Reversals, Idiosyncratic Risk, and Expected Returns. Review of Financial Studies, 23(1).

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. The Journal of Finance, 45(3), pp.881-898.

Jegadeesh, N. and Titman, S., 1993. Returns to buying winners and selling losers: Implications for stock market efficiency. The Journal of Finance, 48(1), pp.65-91.

Levy, H., 1978. Equilibrium in an Imperfect Market: A Constraint on the Number of Securities in the Portfolio. The American Economic Review, 68(4), pp.643-658.

Lintner, J., 1965. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. The review of economics and statistics, pp.13-37.

Malkiel, B., and Xu, Y., 2004. Idiosyncratic risk and security returns, Unpublished Working Paper. Princeton University.

Merton, R.C., 1987. A simple model of capital market equilibrium with incomplete information. The Journal of Finance, 42(3), pp.483-510.

Pastor, L. and Stambaugh, R.F., 2003. Liquidity Risk and Expected Stock Returns. Journal of Political Economy, 111(3), pp.642-685.

Saryal, F.S., 2009. Rethinking idiosyncratic volatility: Is it really a puzzle. Unpublished Working paper. University of Toronto.

Spiegel, M., and Wang, X., 2005. Cross-sectional variation in stock returns: Liquidity and idiosyncratic risk. Unpublished working paper, Yale School of Management.

Table 2.1 Summary statistics

This table presents the summary statistics. *Return* is the monthly raw return reported as a percentage. $BETA$, ME and BE/ME are estimated as in Fama and French (1992). $BETA$ is the portfolio beta estimated from the full period using 100 size and pre-ranking beta portfolios. The market value of equity, ME, is the product of the monthly closing price and the number of shares outstanding in June. The book-to-market equity, BE/ME is defined by the ratio of the book value of equity in the month of June to the market value of equity in the month of December. For estimating the idiosyncratic volatility, $IVOL$, the excess daily returns of each individual stock are regressed on the Fama-French three factors: $R_m - R_f$, SMB, and HML on a monthly basis. The monthly idiosyncratic volatility of the stock is defined as the standard deviation of the regression residuals. $FIVOL$ is the one-month ahead forecasted idiosyncratic volatility, estimated by an EGARCH model. The daily proportional spread is measured by the ask quote minus the bid quote divided by the quote midpoint. *Spread* is the average of the daily proportional spread over each month. The percentage of zero returns, $\%Zeros$ is the fraction of trading days in a month that experience no price movement from the prior-end-of-day price estimated using CRSP daily stock returns.

Variables	Mean	Std. dev.	Median	Q1	Q3	Skew
$Return(\%)$	1.44	19.09	0	-7.1	7.76	5.56
$IVOL$	14.55	14.36	9.78	5.82	16.98	8.22
$FIVOL$	13.95	7.74	10.05	6.32	15.92	10.44
$BETA$	1.32	0.34	1.26	1.06	1.56	0.2
ME	1177.71	8299.05	70.92	17.33	367.83	25.67
BE/ME	3.073	52.39	0.71	0.38	1.27	65.06
$Ln(ME)$	4.75	2.15	4.57	3.17	6.2	0.29
$Ln(BE/ME)$	-0.45	1.11	-0.45	-1.04	0.08	0.98
$Spread$	0.04	0.14	0.02	0.01	0.04	179.84
$\%Zeros$	0.16	0.15	0.13	0.05	0.24	1.23

Table 2.2 Cross-sectional correlations

This table presents the time-series averages of the cross-sectional Pearson correlation coefficients. $LIVOL$ is the lagged IVOL. The other variables are defined in Table I. The correlation coefficients followed by *** are significant at the 1% level based on their time-series standard error.

	Returns	*IVOL*	*FIVOL*	*LIVOL*	*ln(ME)*	*ln(BE/ME)*	*Spread*	*% Zeros*
Returns	1.00***	0.14***	0.13***	0.02***	-0.02***	0.02***	0.02***	0.01***
IVOL		1.00***	0.51***	0.67***	-0.42***	-0.02***	0.30***	0.15***
FIVOL			1.00***	0.50***	-0.30***	-0.09***	0.08***	0.02***
LIVOL				1.00***	-0.39***	-0.02***	0.27***	0.20***
ln(ME)					1.00***	-0.30***	-0.25***	-0.46***
ln(BE/ME)						1.00***	0.06***	0.13***
Spread							1.00***	0.18***

Table 2.3 Fama-MacBeth regressions of stock returns on idiosyncratic volatility and other firm characteristics

The table shows the time-series means of the slopes in cross-sectional regressions using the Fama and MacBeth(1973) methodology. For each month t, cross-sectional regressions of the following form are estimated

$$\begin{aligned} R_{i,t} = \alpha_0 + \beta_{1t}\ BETA_{i,t-1} + \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ FIVOL_{i,t} \\ + \beta_{5t}\ IVOL_{i,t} + \beta_{6t}\ IVOL_{i,t-1} + e_{i,t} \end{aligned}$$

The dependent variable ($R_{i,t}$) is the percentage monthly return between $t-1$ and t. $FIVOL_t$ is the one-month-ahead expected idiosyncratic volatility estimated by an EGARCH model. $IVOL_{t-1}$ is the one-month lagged idiosyncratic volatility. $BETA$, ME, and BE/ME are estimated as in Fama and French (1992) .The last column reports the average R^2 of the cross-sectional regressions. Newey and West t-statistics are indicated in parentheses. *, ** and *** denote significance at the 10%, 5% and 1% level respectively.

Model	$BETA$	$ln(ME)$	$ln(BE/ME)$	$FIVOL_t$	$IVOL_{t-1}$	$IVOL_t$	R^2
1	-0.08	-0.14**	0.15*				3.04%
	(-0.25)	(-2.42)	(2.09)				
2				0.12***			2.83%
				(7.01)			
3						0.18***	5.30%
						(8.86)	
4					-0.01		1.70%
					(-0.71)		
5		0.18***	0.41***	0.14***			4.17%
		(4.14)	(6.07)	(8.05)			
6		-0.13**	0.14		-0.02		2.94%
		(-2.55)	(1.84)		(-1.58)		
7		0.43***	0.41***			0.23***	7.66%
		(9.21)	(7.36)			(11.43)	

Table 2.4 Idiosyncratic volatility and liquidity regressions

This table reports the Fama-MacBeth regression results of idiosyncratic volatility on the influences of microstructure variables, embodied by the proportional bid-ask spread and the percentage of zero returns. To evaluate the relation of each microstructure variable with idiosyncratic volatility, a number of separate regression specifications are used. To account for the first- and the second-order influence of spreads on the IVOL estimate, I include the spread and the squared spread in the model. Since the percentage of zero returns may be a proxy for spreads, an interaction term which accounts for the joint effect of spreads and zeros returns is included. In Panel A, the idiosyncratic volatility estimate is $IVOL_{it}$, the standard deviation of residuals from a time-series regression of stock returns on the Fama-French three factors. The microstructure variables and the $IVOL_{it}$ estimate are contemporaneous. In panel B, the idiosyncratic volatility estimate is $FIVOL_{it}$, the one-month ahead forecasted idiosyncratic volatility estimated by an EGARCH model. These results document the relation between $FIVOL_{it}$ and liquidity variables at $t-1$. Newey and West(1987) t-statistics are indicated in parentheses. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: Relation between $IVOL_{it}$ and liquidity variables at t

Model	*Intercept*	*Spread*	*%Zeros*	*Squared Spread*	*%Zeros ∗ Spread*	R^2 (%)
1	7.84***	221.06***	-10.21***	-266.12***	-34.76***	38.99
	(41.68)	(28.07)	(-15.51)	(-8.07)	(-3.32)	
2	8.74***	151.23***	-6.03***		-42.55***	32.2
	(47.45)	(17.79)	(-10.84)		(-3.04)	
3	8.39***	213.63***	-13.05***	-289.3***		36.64
	(45.76)	(30.51)	(-14.87)	(-9.56)		
4	9.33***	131.64***	-7.12***			26.9
	(45.88)	(21.74)	(-7.47)			
5	7.2***	188.2***		-239.5***		33.37
	(42.3)	(31.02)		(-9.4)		
6	8.92***	120.23***				24.07
	(53.41)	(22.12)				

Panel B: Relation between $FIVOL_{it}$ and liquidity variables at $t-1$

Model	*Intercept*	*Spread*	*%Zeros*	*Squaredspread*	*%Zeros ∗ spread*	R^2
1	9.32***	141.95***	-2.45	-312.85***	24.42	9.28
	(18.29)	(7.6)	(-1.15)	(-8.11)	(0.82)	
2	9.78***	135.94***	-3.33	-289.19***		8.78
	(22.1)	(8.96)	(-1.10)	(-8.34)		
3	9.26***	128.84***		-276.97***		7.9
	(27.79)	(7.87)		(-8.09)		

Table 2.5 Fama-MacBeth regressions of stock returns on idiosyncratic volatility, liquidity and other control variables

This table documents the time-series means of the slopes in cross-sectional regressions using the Fama and MacBeth (1973) methodology. The dependent variable, $R_{(t-1,t)}$ is the percentage monthly return. *Innov* is the contribution of the information between time $t-2$ and $t-1$ to $FIVOL_t$. $R_{(t-2,t-1)}$ is the previous month's returns which controls for the return reversals. The other variables are defined in earlier tables. Newey and West t-statistics are indicated in parentheses. ** and *** denote significance at the 5% and 1% level respectively.

Model	$ln(ME)$	$ln(BE/ME)$	$FIVOL_t$	$FIVOL_{t-1}$	$IVOL_{t-1}$	$Innov$	$Spread$	$\%Zeros$	$R_{(t-2,t-1)}$	$R^2(\%)$
1	0.18***	0.41***	0.14***							4.17
	(4.14)	(6.07)	(8.05)							
2	-0.13**	0.16**		-0.01						2.6
	(-2.54)	(2.11)		(-0.75)						
3	-0.13**	0.14			-0.02					2.94
	(-2.55)	(1.84)			(-1.58)					
4	-0.12	0.16				0.04***				2.38
	(-1.7)	(1.86)				(4.53)				
5	0.19***	0.39***	0.15***				2.79	-0.02		4.82
	(4.01)	(5.54)	(8.39)				(1.14)	(-0.05)		
6	-0.07	0.15**		-0.01			6.67**	-0.54		3.22
	(-1.49)	(2.06)		(-0.9)			(2.79)	(-1.15)		
7	-0.12**	0.11			-0.03**		7.49***	-0.62		3.58
	(-2.33)	(1.37)			(-2.55)		(2.89)	(-1.28)		
8	-0.06	0.16				0.04***	6.18**	-0.5		3.05
	(-0.93)	(1.87)				(4.47)	(2.5)	(-0.98)		
9	0.17***	0.41***	0.14***						-0.04***	5.16
	(3.7)	(6.05)	(7.57)						(-8.39)	
10	-0.11**	0.17**		-0.01					-0.04***	3.36
	(-2.23)	(2.23)		(-0.36)					(-7.98)	
11	-0.11	0.16				0.04***			-0.04***	3.16
	(-1.61)	(1.94)				(4.53)			(-7.24)	
12	0.09**	0.37***	0.17***		-0.06***					5.05
	(2.00)	(5.49)	(10.98)		(-7.00)					
13	0.11**	0.36***	0.17***		-0.07***		7.98***	-0.99**	-0.03***	6.16
	(2.02)	(5.98)	(11.05)		(-7.21)		(2.98)	(-2.00)	(-6.74)	

Table 2.6 Fama-MacBeth regressions of stock returns on idiosyncratic volatility, liquidity and other control variables (Excluding January returns and penny stocks)

This table documents the time-series means of the slopes in cross-sectional regressions using the Fama and MacBeth (1973) methodology.January returns and penny stocks (price less than \$5) are excluded from the sample. The dependent variable, $R_{(t-1,t)}$ is the percentage monthly return. $Innov$ is the contribution of the information between time $t-2$ and $t-1$ to $FIVOL_t$. $R_{(t-2,t-1)}$ is the previous month's returns which controls for the return reversals. The other variables are defined in earlier tables. Newey and West t-statistics are indicated in parentheses. ** and *** denote significance at the 5% and 1% level respectively.

Model	$ln(ME)$	$ln(BE/ME)$	$FIVOL_t$	$FIVOL_{t-1}$	$IVOL_{t-1}$	$Innov$	$Spreads$	$\%Zeros$	$R_{(t-2,t-1)}$	$R^2(\%)$
1	0.09**	0.22***	0.04**						-0.01**	3.88
	(2.43)	(3.29)	(2.61)						(-2.66)	
2	-0.01	0.12		-0.04**					-0.01**	3.39
	(-0.26)	(1.62)		(-2.84)					(-2.17)	
3	-0.03	0.1			-0.06***				-0.01	3.45
	(-0.83)	(1.3)			(-4.25)				(-1.02)	
4	0.04	0.16				0.03***			-0.01	2.95
	(0.86)	(1.94)				(5.07)			(-1.82)	
5	0.05	0.22***	0.04**				-6.13	-0.21	-0.01**	4.46
	(1.21)	(3.44)	(2.58)				(-1.78)	(-0.44)	(-2.55)	
6	-0.06	0.12		-0.04***			-6.42	-0.75	-0.01**	3.99
	(-1.40)	(1.76)		(-3.20)			(-1.83)	(-1.43)	(-2.04)	
7	-0.08	0.11			-0.06***		-3.72	-1.18**	-0.01	4.08
	(-1.83)	(1.47)			(-4.60)		(-0.90)	(-2.54)	(-0.92)	
8	-0.01	0.17**				0.03***	-6.6	-0.52	-0.01	3.65
	(-0.18)	(2.16)				(4.94)	(-1.74)	(-0.89)	(-1.79)	

Table 2.7 Distribution of FIVOL Migration

For each month t, all firms are sorted into quintiles based on $FIVOL_t$. $Migrate_t$ for a stock is defined by its FIVOL quintile rank at t − FIVOL quintile rank at $t-1$. $Migrate$=4 for a stock that was in Quintile 1 at $t-1$ and is in Quintile 5 at t. This table reports the percentage of stocks that are in Quintile i at $t-1$ (based on $FIVOL_{t-1}$) and in Quintile j at t (based on $FIVOL_t$)

		$FIVOL$ Quintile at t				
		1	2	3	4	5
$FIVOL$ quintile at $t-1$	1	82.8	12.36	2.45	1.09	1.2
	2	12.71	68.37	13.93	3.11	1.93
	3	2.24	14.95	61.35	16.08	5.43
	4	1.03	2.61	17.64	59.09	19.63
	5	1.14	1.71	4.63	20.62	71.8

Table 2.8 Alphas of stocks in the highest FIVOL quintile

Quintile portfolios are formed for each month t based on $FIVOL_t$. $Migrate$ is defined by quintile rank at t − quintile rank at $t-1$. This table shows the simple unadjusted returns of the various $Migrate$ portfolios for stocks which belong to the highest FIVOL quintile at t. $Migrate$=4 indicates a stock that was in Quintile 1 at $t-1$ and is in Quintile 5 at t. The CAPM alpha and the Fama-French 3 factor alphas are also documented. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

$Migrate$	0	1	2	3	4
Average ret	1.64**	2.63***	4.05***	5.76***	6.66***
	(2.48)	(4.31)	(5.7)	(6.84)	(6.67)
CAPM Alpha	0.67	1.70***	3.18***	4.85***	5.62***
	(1.43)	(3.78)	(5.23)	(6.37)	(6.76)
FF3 Alpha	0.82**	1.80***	3.16***	4.83***	5.39***
	(2.3)	(5.34)	(7.04)	(7.46)	(6.67)

Table 2.9 Alphas of stocks in the lowest FIVOL quintile

Quintile portfolios are formed for each month t based on $FIVOL_t$. $Migrate$ is defined by quintile rank at $t-$ quintile rank at $t-1$. This table shows the simple unadjusted returns of the various $Migrate$ portfolios for stocks which belong to the lowest FIVOL quintile at t. $Migrate$=-4 indicates a stock that was in Quintile 5 at $t-1$ and is in Quintile 1 at t. The CAPM alpha and the Fama French 3 factor alphas are also documented. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

$Migrate$	-4	-3	-2	-1	0
Average ret	-1.98***	-0.52*	0.06	0.58***	0.65***
	(-4.63)	(-1.82)	(0.24)	(2.69)	(3.86)
CAPM Alpha	-2.64***	-1.05***	-0.42**	0.15	0.32**
	(-9.14)	(-5.34)	(-2.56)	(0.94)	(2.54)
FF3 Alpha	-2.71***	-1.20***	-0.66***	-0.16	0.07
	(-10.66)	(-7.57)	(-5.22)	(-1.52)	(0.76)

Table 2.10 Alphas of stocks that have migrated from the highest IVOL quintile to lower quintiles

Quintile portfolios are formed every month t based on IVOL$_t$. $Migrate\_IVOL$ is defined by quintile rank at t − quintile rank at $t-1$. This table shows the simple unadjusted returns of the various $Migrate\_IVOL$ portfolios that have migrated from the highest IVOL quintile to lower IVOL quintiles. $Migrate\_IVOL$=-4 indicates that the stock was in Quintile 5 at $t-1$ and is in Quintile 1 at t.$Migrate\_IVOL$=0 indicates that the stock was in Quintile 5 at $t-1$ and is still in Quintile 5 at t.The CAPM alpha and the Fama French 3 factor alphas are also documented. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

$Migrate\_IVOL$	-4	-3	-2	-1	0
Average ret	-0.33***	-1.97***	-2.59***	-2.08***	2.62***
	(-2.68)	(-7.34)	(-8.16)	(-5.21)	(3.29)
CAPM Alpha	-0.46***	-2.32***	-3.14***	-2.80***	1.62***
	(-3.63)	(-9.34)	(-13.58)	(-11.18)	(2.7)
FF3 Alpha	-0.55***	-2.50***	-3.16***	-2.73***	1.72***
	(-4.78)	(-11.38)	(-15.81)	(-15.58)	(3.73)

Table 2.11 Alphas of stocks that have migrated from the lowest IVOL quintile to higher quintiles

Quintile portfolios are formed every month t based on IVOL$_t$. $Migrate\_IVOL$ is defined by quintile rank at t − quintile rank at $t-1$. This table shows the simple unadjusted returns of the various $Migrate\_IVOL$ portfolios that have migrated from the lowest IVOL quintile to higher IVOL quintiles. $Migrate\_IVOL$=4 indicates that the stock was in Quintile 1 at $t-1$ and is in Quintile 5 at t. $Migrate\_IVOL$=0 indicates that the stock was in Quintile 1 at $t-1$ and is still in Quintile 1 at t.The CAPM alpha and the Fama French 3 factor alphas are also documented. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

$Migrate\_IVOL$	0	1	2	3	4
Average ret	0.35***	0.91***	1.44***	1.42***	6.41***
	(2.62)	(3.77)	(4.11)	(2.72)	(5.31)
CAPM Alpha	0.1	0.44**	0.89***	0.86*	5.73***
	(0.85)	(2.43)	(3.24)	(1.84)	(5.02)
FF3 Alpha	-0.1	0.08	0.45**	0.25	4.79***
	(-1.18)	(0.71)	(2.41)	(0.73)	(5.55)

Table 2.12 Explanatory power of Migration Dummy Variables

This table documents the time-series means of the slopes in cross-sectional regressions using the Fama-MacBeth methodology. Quintile portfolios are formed every month t, based on $IVOL_t$ and $FIVOL_t$. Then for each month t, cross-sectional regressions of the following forms are estimated.

$$R_{i,t} = \alpha_{0t} + \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ MIG\_FIVOL_{i,t} * FIVOL_{i,t} + \beta_{5t}\ FIVOL_{i,t} + \beta_{6t}\ MIG\_IVOL_{i,t} * IVOL_{i,t-1} \\ +\beta_{7t}\ IVOL_{i,t-1} + \beta_{8t}\ R_{i,t-1} + e_{i,t}$$

where $R_{i,t}$ is the return to stock i in month t, $MIG\_FIVOL_{i,t}$ is the FIVOL migration dummy that takes a value of 1 if the stock moves from a high FIVOL quintile at $t-1$ to a low FIVOL quintile at t or remains in the same FIVOL quintile, and zero otherwise. $MIG\_IVOL_{i,t}$ is the IVOL migration dummy that takes a value of 1 if the stock moves from a high IVOL quintile at $t-1$ to a low IVOL quintile at t, and 0 otherwise. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

Model	$ln(ME)$	$ln(BE/ME)$	MIG_FIVOL*$FIVOL_t$	$FIVOL_t$	MIG_IVOL* IVOL	$IVOL_t$	R_{t-1}
1	0.11**	0.41***	-0.06***	0.19 ***	-0.13***	-0.01	-0.04***
	(2.62)	(7.46)	(-7.19)	(11.61)	(-14.07)	(-0.73)	(-9.36)

Table 2.13 Fama-MacBeth regressions of stock returns on idiosyncratic volatility and dummies for standardized unexpected earnings

This table documents the time-series means of the slopes in cross-sectional regressions using the Fama-MacBeth methodology. Quintile portfolios are formed every month t based on FIVOL$_t$. Then for each month, cross-sectional regressions of the following forms are estimated.

$$\begin{aligned} R_{i,t} = \alpha_{0t} &+ \beta_{2t}\ ln(ME)_{i,t-1} + \beta_{3t}\ ln(BE/ME)_{i,t-1} + \beta_{4t}\ FIVOL_{i,t} + \beta_{7t}\ HIGH\_SUE_{i,t-1} * FIVOL_{i,t} \\ &+ \beta_{8t}\ LOW\_SUE_{i,t-1} * FIVOL_{i,t} + \beta_{9t}\ POS\_SUE_{i,t-1} * FIVOL_{i,t} + \beta_{10t}\ NEG\_SUE_{i,t-1} * FIVOL_{i,t} + \beta_{11t}\ ANNOUN_{i,t-1} + e_{i,t} \end{aligned} \tag{2.10}$$

where $R_{i,t}$ is the return to stock i in month t, $HIGH\_SUE_{i,t-1}(LOW\_SUE_{i,t-1})$ is a dummy variable that takes the value of 1 if the absolute value of standardized unexpected earnings for stock i is ranked in the top (bottom) 20% of stocks in month $t-1$, and zero otherwise. Similarly, $POS\_SUE_{i,t-1}$ ($NEG\_SUE_{i,t-1}$) is a dummy variable that takes the value of 1 if the standardized unexpected earnings for stock i is ranked in the top (bottom) 20% of stocks in month $t-1$, and zero otherwise. $ANNOUN_{i,t-1}$ is the announcement dummy that takes a value of 1 if there is an earnings announcement in month $t-1$, and zero otherwise. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

Model	$ln(ME)$	$ln(BE/ME)$	$FIVOL$	$HIGH\_SUE$* $FIVOL$	LOW_SUE* $FIVOL$	POS_SUE* FIVOL	NEG_SUE* FIVOL	$ANNOUN$	R^2(%)
1	0.01	0.13**	0.09***	0.05***	-0.04***				3.37
	(0.94)	(2.64)	(7.59)	(4.8)	(-4.2)				
2	0.02	0.14**	0.11***	0.08***	-0.07***	0.15***	-0.07***	0.04***	6.67
	(1.84)	(2.42)	(8.38)	(3.46)	(-4.01)	(14.41)	(-6.08)	(7.33)	

Table 2.14 Pricing ability of transitions in FIVOL quintile ranking after controlling for the level of standardized unexpected earnings

This table shows the returns of the various *Migrate* portfolios for the stocks which belong to the top (or bottom) SUE quintile . Quintile portfolios are formed every month t based on $FIVOL_t$ for all stocks in the sample. *Migrate* is defined by FIVOL Quintile rank at t − FIVOL Quintile rank at $t-1$. *Migrate*=-4 indicates that the stock was in Quintile 5 at $t-1$ and is in Quintile 1 at t. *Migrate*=0 indicates that the stock remains in the same quintile in $t-1$ and t. *Migrate*=4 indicates that the stock was in Quintile 1 at $t-1$ and is in Quintile 5 at t. Panel A reports the results for stocks that belong to the highest SUE quintile. Panel B reports the results for stocks that belong to the lowest SUE quintile. The CAPM alpha and the Fama French 3 factor alphas are also documented. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

Panel A: Stocks belonging to the highest SUE quintile

Migrate	**-4**	**-3**	**-2**	**-1**	**0**	**1**	**2**	**3**	**4**
Average ret	-0.85 (-1.60)	-0.34 (-0.99)	-0.06 (-0.21)	0.63* (1.97)	1.77*** (5.15)	2.38*** (6.3)	3.90*** (8.00)	5.64*** (6.89)	9.06*** (5.74)
CAPM Alpha	-1.49*** (-3.42)	-0.95*** (-3.72)	-0.68*** (-3.82)	-0.05 (-0.26)	1.09*** (5.12)	1.66*** (6.39)	3.17*** (8.46)	4.88*** (6.14)	8.08*** (5.45)
FF3 Alpha	-1.46*** (-3.40)	-1.06*** (-4.76)	-0.79*** (-6.55)	-0.19* (-1.86)	0.95*** (8.06)	1.52*** (9.52)	2.92*** (10.24)	4.61*** (6.41)	7.30*** (4.81)

Panel B: Stocks belonging to lowest SUE quintile

Migrate	-4	-3	-2	-1	0	1	2	3	4
Average ret	-1.46*** (-2.77)	-1.65*** (-4.39)	-1.38*** (-4.05)	-0.73** (-2.13)	0.48 (1.32)	0.73* (1.91)	1.50*** (2.68)	3.51*** (4.2)	7.04*** (3.63)
CAPM Alpha	-2.25*** (-5.47)	-2.26*** (-7.98)	-1.98*** (-8.87)	-1.38*** (-7.11)	-0.23 (-1.03)	0.02 (0.1)	0.79* (1.86)	2.87*** (3.63)	6.08*** (3.37)
FF3 Alpha	-2.38*** (-6.55)	-2.32*** (-9.60)	-2.09*** (-12.74)	-1.53*** (-12.47)	-0.35** (-2.28)	-0.15 (-0.91)	0.58 (1.51)	2.87*** (3.61)	5.95*** (3.25)

Table 2.15 Pricing ability of transitions in FIVOL quintile ranking for stocks with no earnings surprises

This table shows the returns of the various $Migrate$ portfolios for the stocks which belong to the lowest absolute (SUE) quintile. Quintile portfolios are formed every month t based on $FIVOL_t$ for all stocks in the sample. $Migrate$ is defined by FIVOL Quintile rank at t − FIVOL Quintile rank at $t-1$. $Migrate$=-4 indicates that the stock was in Quintile 5 at $t-1$ and is in Quintile 1 at t. $Migrate$=0 indicates that the stock remains in the same quintile in $t-1$ and t. $Migrate$=4 indicates that the stock was in Quintile 1 at $t-1$ and is in Quintile 5 at t. The CAPM alpha and the Fama French 3 factor alphas are also documented. Newey-West t-statistics are indicated in parenthesis. ** and *** denote significance at the 5% and 1% level respectively.

$Migrate$	-4	-3	-2	-1	0	1	2	3	4
Average ret	-1.61***	-1.18***	-0.58*	-0.18	1.14***	1.51***	2.19***	4.80***	6.17***
	(-3.12)	(-3.72)	(-1.90)	(-0.61)	(3.35)	(4.57)	(5.23)	(5.84)	(3.63)
CAPM Alpha	-2.14***	-1.68***	-1.15***	-0.79***	0.51**	0.89***	1.53***	4.12***	5.50***
	(-4.69)	(-7.80)	(-5.52)	(-4.29)	(2.32)	(3.95)	(4.78)	(5.79)	(3.11)
FF3 Alpha	-2.28***	-1.79***	-1.37***	-1.01***	0.32**	0.66***	1.29***	4.09***	4.64**
	(-5.30)	(-8.74)	(-9.49)	(-9.71)	(2.32)	(4.38)	(5.59)	(5.63)	(2.53)

Table A1 Returns of portfolios sorted by FIVOL based on a five-factor model

I estimate idiosyncratic risk based on a five-factor model that includes the momentum and the aggregate liquidity factor along with the three Fama-French factors. The forecasted idiosyncratic volatility from an EGARCH model, based on this five-factor model, is the $FIVOL\_5factor$. This table documents the simple unadjusted returns, the CAPM alphas and the Fama French 3 factor alphas when stocks are sorted into quintiles based on their $FIVOL\_5factor$. Quintile 5 stocks have the highest $FIVOL\_5factor$ whereas Quintile 1 stocks have the lowest $FIVOL\_5factor$. Newey and West t-statistics are indicated in parentheses. ** and *** denote significance at the 5% and 1% level respectively.

Quintile	1	2	3	4	5
Average ret	0.46***	0.69***	0.57*	0.29	2.30***
	(2.61)	(2.77)	(1.81)	(0.73)	(3.5)
CAPM Alpha	0.09	0.16	-0.08	-0.48**	1.33***
	(0.76)	(0.99)	(-0.46)	(-2.08)	(2.8)
FF3 Alpha	-0.14	-0.14	-0.33***	-0.59***	1.44***
	(-1.41)	(-1.30)	(-3.38)	(-5.03)	(4.13)

Table A2 Fama-Macbeth regressions of stock returns on forecasted idiosyncratic volatility based on a five-factor model

This table documents the time-series means of the slopes in cross-sectional regressions using the Fama and MacBeth(1973) methodology. $FIVOL\_5factor$ is the FIVOL based on the model including the Fama-French factors, the momentum factor and the Pastor Stambaugh liquidity factor. The other variables are defined in earlier tables. Newey and West t-statistics are indicated in parentheses. ** and *** denote significance at the 5% and 1% level respectively.

Model	$ln(ME)$	$ln(BE/ME)$	$FIVOL\_5factor$	$Spreads$	$\%Zeros$
1	0.17***	0.38***	0.16***		
	(3.39)	(5.54)	(7.35)		
2	0.19***	0.38***	0.16***	2.98	-0.27
	(3.65)	(5.68)	(7.5)	(1.47)	(-0.58)

Table A3 Fama-MacBeth regressions of stock returns on idiosyncratic volatility, liquidity and other control variables (Controlling for January)

This table documents the time-series means of the slopes in cross-sectional regressions using the Fama and MacBeth(1973) methodology. January returns are excluded from this sample. The dependent variable, $R_{t-1,t}$ is the percentage monthly return. Newey and West t-statistics are indicated in parentheses. ** and *** denote significance at the 5% and 1% level respectively.

Model	$ln(ME)$	$ln(BE/ME)$	$FIVOL_t$	$FIVOL_{t-1}$	$IVOL_{t-1}$	$Innov$	$Spread$	$\%Zeros$	$R_{(t-2,t-1)}$	$R^2(\%)$
1	0.26***	0.42***	0.12***						-0.03***	4.19
	(5.72)	(5.69)	(6.96)						(-6.56)	
2	-0.02	0.19**		-0.02					-0.03***	2.98
	(-0.37)	(2.4)		(-1.85)					(-6.28)	
3	-0.03	0.18**			-0.03**				-0.03***	3.21
	(-0.68)	(2.16)			(-2.47)				(-5.93)	
4	0.02	0.22**				0.04***			-0.03***	2.84
	(0.31)	(2.34)				(4.43)			(-5.45)	
5	0.24***	0.42***	0.12***				0.85	-0.77	-0.03***	4.8
	(5.53)	(5.83)	(7.14)				(0.33)	(-1.69)	(-6.50)	
6	-0.02	0.2**		-0.02			4.28	-1.22**	-0.03***	3.57
	(-0.33)	(2.48)		(-1.93)			(1.73)	(-2.52)	(-6.26)	
7	-0.05	0.17**			-0.04***		5.38	-1.33**	-0.03***	3.74
	(-1.05)	(2.18)			(-3.16)		(1.96)	(-2.74)	(-5.57)	
8	0.02	0.22**				0.04***	3.75	-1.14**	-0.03***	3.45
	(0.32)	(2.43)				(4.34)	(1.45)	(-2.21)	(-5.57)	