The Long-Run Component of Idiosyncratic Volatility and Expected Stock Returns

Yunting Liu*

August 25, 2018

*Yunting Liu: Assistant Professor, School of Economics, Peking University, Beijing, China, 100871; yuntingliu@pku.edu.cn. This paper is a revised version of the Chapter 3 of my dissertation at the Johns Hopkins University. I'm especially grateful to Federico Bandi, Gregory Duffee and Jon Faust for their comments that have substantially improved this paper. The author acknowledges research support from School of Economics, Peking University. I have no conflicting interests to declare.

The Long-Run Component of Idiosyncratic Volatility and Expected Stock Returns

Abstract

This paper argues there exists a strong and robust negative relationship between expected long-run component of idiosyncratic volatility and expected stock returns. I present empirical evidence that idiosyncratic volatility of stock returns displays long memory properties. To capture this feature, I develop and estimate a parsimonious model for idiosyncratic volatility consisting of a long-run and short-run component. The long-run component is modeled to be more persistent and could contain a unit root. The short-run component is less persistent and has a zero mean. I find that there exists a robust negative relationship between expected long-run idiosyncratic volatility and expected stock returns. The negative relationship could be stronger and statistically more significant than using other measures of idiosyncratic volatility such as lagged realized idiosyncratic volatility used by Ang et al. (2006). The relationship is also robust after controlling for return reversals, unexpected idiosyncratic volatility.

Keywords: idiosyncratic volatility, long memory, short-run volatility, long-run volatility, cross-sectional stock returns

JEL Code: G12, G17

1 Introduction

Whether a stock's expected return depends on idiosyncratic volatility has been a central question in the asset pricing literature. Ang et al. (2006) (hereafter AHXZ) in a recent influential paper, present evidence that stocks with high realized idiosyncratic volatility have anomalously low returns in the subsequent month. This phenomenon is challenging because traditional asset pricing theories predict no relation between idiosyncratic volatility and expected returns when investors are well diversified and markets are complete and frictionless, or a positive relation when investors don't hold well diversified portfolios and face trading frictions (for example, Merton (1987)).

There is, however, considerable concern about the robustness of the finding by Ang et al. (2006). In particular, the measurement of idiosyncratic volatility is crucial to inferences about its relationship with expected returns. Ang et al. (2006) define the idiosyncratic volatility as the standard deviation of the residuals from the Fama and French (1993) (hereafter FF-3) model estimated with daily returns in the previous month. But lagged idiosyncratic volatility may not be a good proxy for expected idiosyncratic volatility. A number of papers construct measures of expected idiosyncratic volatility and find that the negative relationship between idiosyncratic volatility and expected returns may not be robust. Chua et al. (2010) and Fink et al. (2012) model the evolution of the idiosyncratic volatility as autoregressive moving average processes (ARMA) and Chua et al. (2010) find a positive relationship between expected idiosyncratic volatility and expected returns. Moreover, Huang et al. (2010) argues realized lagged realized idiosyncratic volatility is a biased estimator and it is not robust to controlling for return reversal. Fu (2009) construct measures of expected idiosyncratic volatility using exponential generalized autoregressive processes (EGARCH) and argues there exists a positive relationship between expected idiosyncratic volatility and expected returns.

However, existing models for estimating conditional idiosyncratic volatilities don't well capture the dynamics of idiosyncratic volatility in the long-run. I document empirical evidence that idiosyncratic volatility persists for long periods of time, at frequencies more than a year. it has been noted in the literature (Andersen et al. (2003), for example) that generalized autoregressive conditional heteroskedasticity (GARCH) and similar models don't well capture the persistence of stock return volatility in the long run. Thus, those models may produce biased proxies for the conditional idiosyncratic volatility.

To capture the long memory properties of idiosyncratic volatility, this paper develops a parsimonious model of idiosyncratic volatility featuring two components. One is a long-run component and could be modeled as containing a unit root. The other is a short-run one and is less persistent. This approach parsimoniously captures shocks to volatilities at different horizons. I find that there exists a strong and robust negative relationship between expected idiosyncratic volatility and expected stock returns. This relationship is robust after controlling for return reversals and unexpected idiosyncratic volatility. This analysis suggests that the expected long-run volatility may be the measure that investors compare about. Measures of expected idiosyncratic volatility without looking at the long-run persistence may produce a biased relationship between expected idiosyncratic volatility and expected stock returns.

Another reason why the long-run component is important for investors may be trading costs, which prevent them from actively adjusting portfolios. Thus, these investors might care about the conditional volatility of a stock's return over a horizon longer than a month. In this case, the long-run component is an even better measure of the conditional volatility. Yet another possibility is that the conditional volatility of a stock's idiosyncratic return is proxying for a common risk factor. If this is the case, we don't have a theory that says whether the common risk factor is the short-run component of volatility or the long-run component. It is worth investigating which component it might be.

I study two types of models for idiosyncratic volatility. In one type of the model, the long-run component is constrained to have a unit root and the short-run component is a white noise. The other type of the model doesn't impose these restrictions. Parameter estimates from the model using asset prices data suggest that these restrictions could be plausible assumptions about the idiosyncratic volatility process. The quantitative asset pricing implications are also similar. There is a significantly negative relationship between expected long-run volatility and expected returns. The short-run component is found to be very close to serially uncorrelated, which suggest that the expected short-run component is close to zero.

Therefore, this paper proposes that the expected long-run component of idiosyncratic volatility may be a better measure to capture the cross-sectional relationship between idiosyncratic volatility and expected stock returns. Given that it is the persistent component of idiosyncratic volatility that matters for the cross-section of stock returns, future research may be devoted to finding a risk-based explanation of the negative relationship between idiosyncratic volatility and expected stock returns (e.g, Guo and Savickas (2006)).

There have been several papers in the literature that demonstrate short-run and long-run components of volatility have different asset pricing implications. Lee and Engle (1993) show that aggregate stock market volatility is subject to shocks at different frequencies. Adrian and Rosenberg (2008) find that the short-run and long-run components of equity market volatility are differently priced risks and are important cross-sectional stock returns factors. Christoffersen et al. (2008) propose that short-run and long-run components of volatility have important implications for option pricing. In this paper, I decompose idiosyncratic volatilities at the firm level into short-run and long-run components and ask whether exposures to short-run and long-run volatilities have different asset pricing implications. Corsi (2009) proposes an additive cascade model of volatility defined over different time frequencies to capture the persistence of aggregate stock return volatility.

2 Estimating Idiosyncratic Volatilities

This section describes the data and methods used to estimate idiosyncratic volatilities.

2.1 Data

My dataset includes monthly and daily return data on stocks traded in the NYSE, AMEX, and NASDAQ return files from the Center for Research in Security Prices (CRSP). The accounting variables are from Compustat annual industrial files of income-statement and balance-sheet data.

The CRSP returns cover NYSE and AMEX stocks until 1973 when NASDAQ returns also come on line. The COMPUSTAT data are from 1962 to 2015. The 1962 start date reflects the fact that book value of common equity (COMPUSTAT item 60), is not generally available prior to 1962. More importantly, COMPUSTAT data from earlier years have a serious selection bias; the pre-1962 data are tilted toward big historically successful firms.

The procedures below are standard in the literature following Fama and French (1992). To ensure that the accounting variables are known before the returns they are used to explain, I match the accounting data for all fiscal year ends in calender year $t-1$ with the returns for July of year t to June of year $t+1$. The 6-month (minimum) gap between fiscal year end and the return tests is conservative. I use a firm's market equity at the end of December of year $t-1$ to compute its book-to-market ratio for year $t-1$.

2.2 Idiosyncratic Volatility Definition

Following Ang et al. (2006) and Bali and Cakici (2008), I concentrate on idiosyncratic volatility defined and measured relative to the Fama and French (1993) (FF-3) model. Specifically, I consider the following specification for each firm at each month

$$r^i_{t,d} = \alpha^i_t + \beta^i_{MKT} MKT_{t,d} + \beta^i_{SMB} SMB_{t,d} + \beta^i_{HML} HML_{t,d} + \sigma^i_t \epsilon^i_{t,d} \tag{1}$$

where for day d in month t, $r^i_{t,d}$ is stock i's excess return, $MKT_{t,d}$ is the market excess returns, $SMB_{t,d}$ and $HML_{t,d}$ capture size and book-to-market effects, respectively. The residuals $\eta^i_{t,d} \equiv \sigma^i_{t,d} \epsilon^i_{t,d}$ is the idiosyncratic risk for month t. I define the idiosyncratic volatility of stock returns for firm i in month t as

$$v^i_t = \sigma^i_t \sqrt{N_m} \tag{2}$$

where N_m is the number of trading days in month t for firm i. It is useful to note that the idiosyncratic volatility v^i_t is the daily standard deviation of residuals times the square root of the number trading days in that month. The inclusion of N_m transforms the daily return residuals into monthly residuals. This procedure can be can be seen at French et al. (1987) and Fu (2009).

2.3 Construction of Realized Idiosyncratic Volatility

Since the latent conditional volatility v^i_t cannot be directly observed, I use realized volatility: squared daily return residuals in month t obtained through the cross-sectional regression of equation (1) to measure the individual stock's idiosyncratic volatility for month t. Specifically,

$$IV^i_t \equiv \sqrt{\sum_{d=1}^{N_m} (\eta^i_{t,d})^2} \tag{3}$$

When I refer to idiosyncratic volatility in this paper, I mean idiosyncratic volatility relative to the FF-3 model. To improve the precision on idiosyncratic volatility measures, I require that firms should have at least 15 trading days in a month for which the CRSP reports a daily return. Moreover, I require that firms must have at least 30 months of estimates for the idiosyncratic volatility to be used for my empirical analysis.

2.4 Time Series Properties of Realized Idiosyncratic Volatility

Table 1 presents the time-series properties of the realized idiosyncratic volatility (IV).

[**Insert Table 1 about here**]

I first compute the time-series statistics of idiosyncratic volatility for each firm and then summarize the mean statistics across about 22,000 firms. The mean of idiosyncratic volatility is on average 15.54% across stocks and the mean standard deviation for IV is 9.21%. The skewness is 2.00 and kurtosis is 8.23 which suggests that the idiosyncratic volatility is positively skewed and fat-tailed. The autocorrelation for the realized idiosyncratic volatility is 0.39 at 1-month lag, 0.31 at 2-month lag, 0.21 at 5-month lag, 0.12 at lag 10, and 0.12 at one-year lag. The autocorrelation of 0.39 at one-month lag and 0.31 at two-month lag suggests that shocks to idiosyncratic volatility are not very persistent at short horizons (within a quarter). However, the autocorrelations decay slowly over longer periods, for example over a year after shocks. The autocorrelation of realized idiosyncratic with a lag period of 12 months is still more than 0.1. In comparison, the autocorrelation of an AR(1) process with first order autocorrelation of 0.39 would predict that the autocorrelation with 12 months lag is less than 0.2 basis point (One basis point equals to one hundredth of one percentage point).

This pattern of a relatively quick initial decline in the autocorrelation function followed by a slower decay is sufficient evidence to dismiss the simple ARMA model to model idiosyncratic volatility (for example, Barndorff-Nielsen and Shephard (2002)). It suggests that idiosyncratic volatility displays long memory properties: persistence of the volatility that lasts for long time periods (years). Corsi (2009) proposes that the long memory property of aggregate stock return volatility could be captured by a additive cascade model of volatility defined over different time periods. The empirical evidence in the section suggests that the idiosyncratic volatility of stock returns also displays long-memory property. Therefore, the realized idiosyncratic volatility could thus be better captured by a process with components of different persistences. I model the log of idiosyncratic volatility as the sum of a short-run and long-run component. The short-lived component is less persistent and has a large impact on the autocorrelations of idiosyncratic volatility at short horizons (over a quarter). There also exists a persistent long-run component, which dominates the autocorrelations at longer horizons (over a year and further).

2.5 Decomposing Idiosyncratic Volatility

To decompose the idiosyncratic volatility into short-run and long-run components, I model the idiosyncratic volatility v_t^i follows

$$\begin{aligned} \textbf{Idiosyncratic Volatility} &: \log v_t^i = s_t^i + l_t^i \qquad (4) \\ \textbf{Short-run Component} &: s_{t+1}^i = \rho_s^i s_t^i + \sigma_s^i \epsilon_{s,t}^i \\ \textbf{Long-run component} &: l_{t+1}^i = \phi_i + \rho_l^i l_t^i + \sigma_l^i \epsilon_{l,t}^i \end{aligned}$$

I refer this model as the short and long-run (SL) model hereafter. In equation (4), the log-volatility is the sum of two components, s_t and l_t. The short-run component s_t is a mean-reverting process with mean zero and shocks to the short-run component die off quickly over time. The long-run component l_t is a persistent component. I normalize the process such that $\rho_l > \rho_s$ and σ_s and σ_l are the volatility of shocks to the short-run and long-run components.. This restriction identifies the model as otherwise the two components can be interchangable. Adrian and Rosenberg (2008) consider a two components model for aggregate stock market volatility and find that the prices of risk are different for the short-run and long-run component. My paper differs from theirs in the focus on idiosyncratic volatility of stock returns.

For each firm, equation (4) is readily in the state space form and the unobserved short-run and long-run components can be directly estimated via Kalman filter. I consider $\hat{s}_t \equiv \mathbb{E}_{t-1}(s_t|y_1, y_2, ...y_{t-1})$ and $\hat{l}_t \equiv \mathbb{E}_{t-1}(l_{t+1}|y_1, y_2, ...y_t)$ as the expectation for the short-run and long-run components at time t based on information available at time $t-1$. The smoothed estimates $\tilde{s_t} = \mathbb{E}(s_t|y_1, y_2, ...y_T)$ and $\tilde{l_t} = \mathbb{E}(l_t|y_1, y_2, ...y_T)$, which use the whole sample information, may produce more precise estimates for the expectation of unobserved components s_t and l_t at each point in time. Therefore, full information set estimates are appropriate for asset pricing tests because of the gain in accuracy. However, in terms of the evaluation of trading strategies, incorporating future observations in the forecast leads to using substantial information beyond what investors are aware of. Therefore, I use both the filter estimates and smoothed estimates for evaluating trading strategies and asset pricing models based on idiosyncratic volatility, respectively.

2.6 A Permanent and Transitory Special Case

In my empirical work, I also investigate a special case of equation (4) where the long-run component is permanent and follows a random walk and the short-run component is purely transitory and follows a white noise. I refer this model as the permanent and transitory (PT) model.

$$\begin{aligned} \textbf{Idiosyncratic Volatility} &: \log v_t^i = s_t^i + l_t^i \\ \textbf{Short-run Component} &: s_{t+1}^i = \sigma_s^i \epsilon_{s,t}^i \\ \textbf{Long-run component} &: l_{t+1}^i = l_t^i + \sigma_l^i \epsilon_{l,t}^i \end{aligned} \tag{5}$$

Equation (5) may be viewed as a special case of (4) with the restriction that $\rho_s = 0$ and $\rho_l = 1$. The log-volatility is the sum of two components, s_t and l_t. The short-run component s_t is a white noise process and the long-run component l_t follows a random walk. Thus, changes to the long-run volatility could be permanent and are persistent over time. For each firm, equation (5) can also be estimated using Kalman filter. Since the long-run component l_t follows a random walk, the one-step-ahead conditional expectation of the long-run component $\mathbb{E}_t(l_{t+1}|y_1, y_2, ..., y_t) = \mathbb{E}_t(l_t|y_1, y_2, ..., y_t)$ and that of the short-run component $\mathbb{E}_t(s_{t+1}|y_1, y_2, ..., y_{t-1})$ trivially equals to zero. For the permanent and transitory (PT) model, I only consider the expectation of the long-run component: $\hat{l}_t = \mathbb{E}_{t-1}(l_t|y_1, y_2, ..., y_{t-1})$ and $\tilde{l}_t = \mathbb{E}(l_t|y_1, y_2, ..., y_T)$ and investigate their relationship with expected returns.

2.7 Parameter Estimates of the Idiosyncratic Volatility Model

In practice, the true conditional idiosyncratic volatility v_t cannot be directly observed. Consequently, the realized volatility IV_t is proxying the latent volatility subject to measurement errors. Because measurement errors are largely identically and independent distributed over time, it has little forecasting power for forming conditional expectations. In empirical studies, realized volatilities are usually treated as measuring latent volatilities without errors. (for example, Bollerslev and Zhou (2002), Chua et al. (2010)). In this paper, I take this approach as well. One reason for this treatment is it massively simplifies subsequent estimation and analysis.

Besides, the errors-in-variable problem in realized volatility tends to weaken the predictability of estimated conditional volatility on stock returns in this paper, if not having no effect. Taking the measurement error into full consideration may be useful for future research. Nevertheless, I don't think accounting for that would change the conclusion about the result of the paper that the there is a strong and negative relationship between the conditional volatility of the long-run component and expected returns.

Table 2 summarizes the parameter estimates for the short-run and long-run volatility (SL) model with equation (4) and the permanent transitory volatility (PT) model with equation (5). Both the SL and PT model are estimated using the maximum likelihood method. For the SL model, the mean AR(1) parameter for the short-run component is -0.06 while the median is 0.003. The long-run component is more persistent with mean AR(1) coefficient of 0.78 and median of 0.94. The mean volatility of shocks to the short-run component is 0.11 and the median is 0.09 . For the long-run component, the mean volatility is 0.07 and the median is 0.02.

[**Insert Table 2 about here**]

The permanent and transitory (PT) model can be viewed as a special case of the SL model with $\rho_s = 0$ and $\rho_l = 1$. Given that the median estimate of ρ_s is 0.003 and ρ_l is 0.94 from the SL model, the PT model can be a plausible model to capture the dynamics of idiosyncratic volatility. For the PT model, the only parameters to be estimated are the volatility of shocks to the short-run and long-run component. The mean volatility of shocks to the short-run component is 0.14 and the median is 0.11. As for the volatility of shocks to the long-run component, the mean is 0.05 and the median is 0.01. The magnitude of shocks is also largely similar to estimates from the SL model.

3 Pricing Idiosyncratic Volatility in the Cross-Section

This section considers the performance of portfolios formed by different measures of idiosyncratic volatilities and asks whether exposures to different volatilities are systematically important for expected stock returns.

3.1 Patterns in Average Returns for Idiosyncratic Volatility

I first consider value-weighted quintile portfolios formed every month by sorting stocks based on realized idiosyncratic volatility relative to the Fama and French (1993) model. Portfolios are formed every month, based on realized volatility computed using daily data over the previous month. Similar to findings of Ang et al. (2006), portfolios with high realized idiosyncratic volatility have low returns. Table 3 shows that average returns increases slightly from 0.92% per month to 0.98% going from the quintile 1 (low idiosyncratic volatility stock) to quintile 3. Then portfolios returns drop tremendously going to quintile 5. The portfolio with highest idiosyncratic volatility (quintile 5) has a surprisingly low average return with 0.09% per month. The difference of returns between the highest and lowest portfolio is as large as 0.83% per month, which is statistically significant with robust t-statistic of -2.98. The FF3-alpha for quintile 5, is -1.13% per month with a robust t-statistics of -7.85. Therefore, the difference in quintile portfolio returns can not be explained by our standard asset pricing model such as Fama and French (1993). The difference in average returns in Table 3 is strong evidence to support the negative relationship between expected return and idiosyncratic volatility.

[**Insert Table 3 about here**]

3.2 Portfolios Sorted by Short-Run and Long-Run Volatilities

Estimating the state-space model of (4) and (5) produces estimates for the expected short-run and long-run volatilities at the monthly frequency. I form value-weighted quintile portfolios sorting by the short-run volatility $\hat{s}_t$ or $\tilde{s}_t$ and long-run volatility $\hat{l}_t$ or $\tilde{l}_t$.

I first consider the performance of portfolios sorted by the filtered estimates of the expected long-run component. Portfolio returns displayed in Table 4 reveals substantial spread in average across quintile portfolios. Exposure to long-run volatilities is a robustly negative related with average returns. For both the SL model and PT model, the portfolio with highest long-run volatilities earns very low average returns. The average return for quintile 5 portfolio with the highest long-run idiosyncratic volatility is as low as -0.24% per month for the SL model and remarkably lower as -0.21% for the PT model. The difference in the average return between the portfolio with the lowest and highest long-run volatility is -1.33% per month for the SL model and -1.41% for the PT model. The difference in returns can also not be explained by the difference in the Fama-French model. The highest long-run volatility portfolio has FF3-alpha of -1.50% for the SL model and -1.54% for the PT model. The alphas are statistically significant with robust t-statistics of -6.53 and -7.56.

In particular, it is worth noting that this negative relationship between expected long-run idiosyncratic volatility and expected return is stronger than that from sorting on realized idiosyncratic idiosyncratic volatility. Forming quintile portfolios on lagged idiosyncratic volatility generate a return spread of -0.89% with a t-statistics of 2.98. Table 5 reports the portfolio performance by sorting stocks based on the smoothed estimates of the expected long-run component. Similar to findings in Table 4, high idiosyncratic volatility portfolios earn low returns. The quintile 5 portfolio with high smoothed estimates of the long-run component earn a return of -0.34% for the SL model and as low as -1.22% for the PT model. The return spread between quintile 1 and 5 portfolio for the PT model is remarkly large to -2.51% with a t-statistic of -4.76. This stronger relationship found in the PT may be gain of allowing for unit root to better capture the dynamics of idiosyncratic volatility. The potential existence of unit root further supports that there could be some very persistence variations in idiosyncratic volatility.

However in terms of the findings in Table 6, I don't find a robust significant relation between exposure to the short-run component and expected return. For the SL model, portfolios sorted by filtered estimates $\hat{s}_t$ don't reveal any robust return pattern. There is little spread in average returns across portfolios. For the smoothed estimates, the quintile 1 portfolio earns a return of 0.93% per month and the quintile 5 portfolio earns 1.37%, which suggests a positive relationship between the short-run component and expected returns.

4 Cross-Sectional Regressions

In this section, I investigate the cross-sectional relationship between average stock returns and the estimated conditional idiosyncratic volatilities. I employ Fama and MacBeth (1973) regressions of the cross-section of stock returns on idiosyncratic volatilities and other firms characteristics on a monthly basis and calculate the time-series averages of the coefficients. My goal is to test whether the coefficient on idiosyncratic volatility is significantly different

[**Insert Table 4, 5 and 6 about here**]

from zero in explaining cross-sectional stock returns.

Specifically, I run the following cross-sectional regressions each month for the SL and PT model:

$$R_{i,t+1} = X_{i,t}\beta_t + \gamma_{l,t}\hat{l}_t + \gamma_{s,t}\hat{s}_t + \epsilon_{i,t+1} \tag{6}$$

$$R_{i,t+1} = X_{i,t}\beta_t + \gamma_{l,t}\tilde{l}_t + \gamma_{s,t}\tilde{s}_t + \epsilon_{i,t+1} \tag{7}$$

where $X_{i,t} = [Beta_{i,t}, Ln(Size)_{i,t}, Ln(BE/ME)]$, $Beta_{i,t}$ is the estimate for stock i's market beta in month t. The term $Ln(Size)_{i,t}$ is the log of stock i's market capitalization at the end of month t, and $Ln(BE/ME)_{i,t}$ is the log of stock i's book-to-market ratio as of the end of month t based upon the last fiscal year's information.

To obtain estimates for $Beta_{i,t}$, the procedures are to follow Fama and French (1992). I divide all stocks traded in NYSE, AMEX, and NASDAQ into ten groups by their market capitalization. Within each size decile, stocks are sorted again by their pre-ranking betas into ten groups. The pre-ranking betas are estimated on previous 60 months returns for each firm. The one hundred portfolios are then rebalanced every month. For each portfolio, I estimate portfolio betas again using the full sample of returns on each of the 100 portfolios. The full-period beta estimates of betas are assigned to each stock in the 100 portfolios. Note that assigning the full-period portfolio betas to stocks does not mean that a stock's beta is not changing over time. When the portfolio is rebalanced every month, a stock can move across portfolios with changes in the market capitalization and in the beta estimates for the preceding 5 years.

Table 7 shows time-series averages of the coefficients from the month-by-month Fama-Macbeth (FM) regressions of the cross-section of stock returns on size, *Beta*, and book-to-market ratio and different measures of idiosyncratic volatility. The average coefficient on variables used to explain expected returns provides standard FM tests for determining which variables on average have explanatory power during the July 1963 to December 2015 period.

The regressions in Table 7 say that size and book-to-market are useful variables to explain the cross-section of stock returns. When regressing on size, *Beta* and book-to-market ratio, the average coefficient from the monthly regressions on size is -0.08 with a t-statistic of -1.50. The coefficient on book-to-market ratio is 0.19 with a t-statistic of 3.17. These significant regressions persist no matter what idiosyncratic volatility measures are put into the regression. In contrast to the power of size and book-to-market to explain average stock returns, market beta is unimportant to explain the cross-section of stock returns. When size and book-to-market are controlled for, the average coefficient for *Beta* is only -0.24 with a t-statistic of -0.87.

[**Insert Table 7 about here**]

Given that that size and book-to-market are useful benchmark variables to explain stock returns, the next step is to consider whether idiosyncratic volatility matters for the cross-section of stock returns. The average coefficient on the log of realized idiosyncratic volatility (IV) is -0.52 with a t-statistic of -4.83. The finding confirms a negative relationship between idiosyncratic volatility and expected return.

The next important regressions are to put the measure of expected short-run and long-run idiosyncratic volatility into Fama-Macbeth regressions. The regression using SL model is reported in Table 7. The regression results says that there exists a negative relationship between expected long-run volatility $\hat{l}_t$ and expected returns. The average coefficient is -0.93 with a significant t-statistic of -4.50. In comparison, the relationship between expected short-run volatility $\hat{s}_t$ and the cross-section of stock returns is positive with a coefficient of 2.96 and a t-statistic of 4.67. The FM test lends support to the finding in Section 3.2 of constructing portfolios by sorting on $\hat{s}_t$ and $\hat{l}_t$. Therefore, there exists a robustly significantly negative relationship between expected long-run volatility and expected returns.

It is also be useful to explain the finding using the PT model to measure expected short-run and long-run component. Table 8 provides the result. The average coefficient on the filtered estimates of expected long-run volatility $\hat{l}_t$ has a coefficient of -0.65 with a t-statistic of -3.65. The coefficient with the smoothed estimates is -0.88 with the significance level of -3.93.

Therefore, the section finds there isa significant negative relationship between the expected long-run component and average stock returns by using Fama-Macbeth regressions.

4.1 Controlling for Return Reversals

Stock returns display short-term reversals (Jegadeesh (1990) and Lehmann (1990)). Return reversal describes a phenomena that if a stock's previous month return is too high (low), it will tend to reverse the following month and earn a low (high) return. Following Huang et al. (2010), I use the returns of individual stocks in the prior month to control for return reversals. Therefore, the equation (7) is modified to allow for previous month's stock return

$$r_{t,d}^{i} = X_{i,t}\beta_t + \gamma_{l,t}\hat{l}_t + \gamma_{s,t}\hat{s}_t + \beta_{r,t-1}r_{t-1}^{i} + v_t^{i}\epsilon_{t,d}^{i} \tag{8}$$

Without previous month's stock return r_{t-1}^{i}, the relationship between idiosyncratic risk and expected stock returns may be negatively biased because the coefficient incorporates part of the return reversal that should have been captured by the stock returns of the previous month. Including return reversals into the Fama-Macbeth regression Table 7 shows that the coefficient on the log of realized volatility is reduced to a coefficient of -0.30 with a t-statistic of -2.64. This finding is in line with Huang et al. (2010) that part of the finding by Ang et al. (2006) can be explained by return reversals. And the coefficient on the lagged month return is statistically significant with a statistic of -11.50.

However, accounting for return reversals doesn't quite change the significance level on the coefficients of the expected long-run component and expected returns. For the SL and PT model, the statistical significance level is reduced from -4.50 to -3.67 for the filtered estimates and -6.52 to -3.96 for the smoothed estimates. Therefore, returns reversals are not the key driver of the relationship between the expected long-run component and expected returns.

4.2 Controlling for Unexpected Idiosyncratic Volatility

In line with the argument of French et al. (1987), the relationship between expected idiosyncratic volatility and expected returns in the above regression may be clouded by the relationship between unexpected idiosyncratic volatility and unexpected stock returns. To control for this effect, I add unexpected idiosyncratic volatility to the Fama-Macbeth regression. I define the unexpected idiosyncratic volatility as

$$\mu_t = \log v_t - \hat{s}_t - \hat{l}_t$$

where $\log v_t$ is the log of realized volatility at time t, $\hat{s}_t$ and $\hat{l}_t$ are volatility forecasts made at time $t-1$ for the short-run and long-run component at time t. When the unexpected idiosyncratic volatility is added to the regression, the relationship between expected idiosyncratic volatility and expected stock returns is robust and may become stronger, which further supports that there is a strong negative relationship between expected long-run component and average stock returns.

Table 7 includes results for the SL model. When the unexpected idiosyncratic volatility is added to the regression, the coefficient on the long-run component is slightly reduced for the filtered estimates but increases for the smoothed estimates.

Table 8 reports that the increase in the coefficient on expected long-run volatility is particular large for the smoothed estimates. The average coefficient increases from -0.86 to -1.68 with a significance level of -7.81. This finding echoes the portfolio analysis in Section 3. The return spread between low volatility and high volatility quintile portfolios is -2.51% for the smoothed estimates of the expected long-run volatility using the PT model.

5 Conclusion

The paper develops and estimates a model for idiosyncratic volatility and decomposes the volatility of idiosyncratic stock return into short-run and long-run components. The results of the paper suggest that there is a robust and significant negative relationship between expected long-run volatility and expected returns. The relationship is stronger than using lagged realized idiosyncratic volatility to measure expected stock returns. It suggest the expected long-run component may be a better measure to capture expected idiosyncratic volatility.

In contrast, expected short-run volatility may be positively to be related to expected returns. Fu (2009) and Huang et al. (2010) propose that returns reversals can largely account for the negative relationship between idiosyncratic volatility and expected returns. The findings of this paper remain robust when stock returns in the previous month are used to control for return reversals. This negative relationship is also robust controlling for unexpected idiosyncratic volatility.

Since the persistent component of idiosyncratic volatility is a negatively related to expected stock returns in the cross-section, the underlying mechanism driving the phenomena may be favorably risk-based. Idiosyncratic volatility related to irrational pricing may die off quickly. Stocks with persistent high expected idiosyncratic volatilities could earn lower returns because they are exposed less to fundamental risk factors and thus earn lower risk premium. The paper suggests more work could be devoted to developing a risk-based asset pricing model to explain the negative relationship between idiosyncratic volatility and expected returns.

References

Adrian, Tobias and Joshua Rosenberg, "Stock Returns and Volatility: Pricing the Short-Run and Long-Run Components of Market Risk," *The Journal of Finance*, 2008, *63* (6), 2997–3030.

Andersen, Torben G, Tim Bollerslev, Francis X Diebold, and Paul Labys, "Modeling and Forecasting Realized Volatility," *Econometrica*, 2003, *71* (2), 579–625.

Ang, Andrew, Robert J Hodrick, Yuhang Xing, and Xiaoyan Zhang, "The Cross-Section of Volatility and Expected Returns," *The Journal of Finance*, 2006, *61* (1), 259–299.

Bali, Turan G and Nusret Cakici, "Idiosyncratic Volatility and the Cross Section of Expected Returns," *Journal of Financial and Quantitative Analysis*, 2008, *43* (01), 29–58.

Barndorff-Nielsen, Ole E and Neil Shephard, "Econometric Analysis of Realized Volatility and Its Use in Estimating Stochastic Volatility Models," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2002, *64* (2), 253–280.

Bollerslev, Tim and Hao Zhou, "Estimating stochastic volatility diffusion using conditional moments of integrated volatility," *Journal of Econometrics*, 2002, *109* (1), 33–65.

Christoffersen, Peter, Kris Jacobs, Chayawat Ornthanalai, and Yintian Wang, "Option Valuation with Long-Run and Short-Run Volatility Components," *Journal of Financial Economics*, 2008, *90* (3), 272–297.

Chua, Choong Tze, Jeremy Goh, and Zhe Zhang, "Expected Volatility, Unexpected Volatility, and the Cross-Section of Stock Returns," *Journal of Financial Research*, 2010, *33* (2), 103–123.

Corsi, Fulvio, "A Simple Approximate Long-Memory Model of Realized Volatility," *Journal of Financial Econometrics*, 2009, *7* (2), 174–196.

Fama, Eugene F and James D MacBeth, "Risk, Return, and Equilibrium: Empirical Tests," *Journal of Political Economy*, 1973, *81* (3), 607–636.

_ **and Kenneth R French**, "The Cross-Section of Expected Stock Returns," *the Journal of Finance*, 1992, *47* (2), 427–465.

_ **and** _ , "Common Risk Factors in the Returns on Stocks and Bonds," *Journal of Financial Economics*, 1993, *33* (1), 3–56.

Fink, Jason D, Kristin E Fink, and Hui He, "Expected idiosyncratic volatility measures and expected returns," *Financial Management*, 2012, *41* (3), 519–553.

French, Kenneth R, G William Schwert, and Robert F Stambaugh, "Expected Stock Returns and Volatility," *Journal of Financial Economics*, 1987, *19* (1), 3–29.

Fu, Fangjian, "Idiosyncratic Risk and the Cross-Section of Expected Stock Returns," *Journal of Financial Economics*, 2009, *91* (1), 24–37.

Guo, Hui and Robert Savickas, "Idiosyncratic Volatility, Stock Market Volatility, and Expected Stock Returns," *Journal of Business and Economic Statistics*, 2006, *24* (1), 43–56.

Huang, Wei, Qianqiu Liu, S Ghon Rhee, and Liang Zhang, "Return Reversals, Idiosyncratic Risk, and Expected Returns," *The Review of Financial Studies*, 2010, *23* (1), 147–168.

Jegadeesh, Narasimhan, "Evidence of Predictable Behavior of Security Returns," *The Journal of Finance*, 1990, *45* (3), 881–898.

Lee, Gary GJ and Robert F Engle, "A Permanent and Transitory Component Model of Stock Return Volatility," 1993.

Lehmann, Bruce N, "Fads, Martingales, and Market Efficiency," *The Quarterly Journal of Economics*, 1990, *105* (1), 1–28.

Merton, Robert C, "A Simple Model of Capital Market Equilibrium with Incomplete Information," *The Journal of Finance*, 1987, *42* (3), 483–510.

Tables and Figures

Table 1: Time Series Properties of Idiosyncratic Volatility

Panel A: Some Summary Statistics of Idiosyncratic Volatility						
Mean	Std.Dev.	Skewness	Kurtosis			
15.54	9.21	2.00	8.23			
Panel B: Autocorrelations of Idiosyncratic Volatility						
ACF(1)	ACF(2)	ACF(3)	ACF(4)	ACF(5)	ACF(10)	ACF(12)
0.39	0.31	0.28	0.23	0.21	0.12	0.12

This table summarizes the time-series statistics for idiosyncratic volatility. I first compute the statistics for each stock and the average the statistics across all stocks. The sample period is January 1963 to December 2015. The ACF stands for estimated autocorrelations at different lags. The unit of the mean and standard deviation is percentage point.

Table 2: Parameter Estimates for Idiosyncratic Volatility Model

Panel A: The Short and Long Run Volatility (SL) Model				
Variables	ρ_s	ρ_l	σ_s	σ_l
Mean	-0.06	0.78	0.11	0.07
Median	0.003	0.94	0.09	0.02
Panel B: The Permanent and Transitory Volatility Model				
Variables	σ_s	σ_l		
Mean	0.14	0.05		
Median	0.11	0.01		

This table summarizes the properties of parameter estimates for the short-run and long-run idiosyncratic volatility process. I first compute the parameter estimates for each stock and then construct the mean and median statistics across all stocks. The sample period is January 1963 to December 2015.

Table 3: Portfolios Sorted by Idiosyncratic Volatility

Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	0.92	3.76	42.8%	0.09
				[2.12]
2	0.96	4.70	31.9%	0.02
				[0.51]
3	0.98	5.76	15.4%	-0.03
				[-0.56]
4	0.76	7.16	7.2%	-0.33
				[-3.51]
5 (high)	0.09	8.52	2.5%	-1.13
				[-7.85]
5-1	-0.83	6.79		-1.61
	[-2.98]			[-9.33]

I form value-weighted quintile portfolios every month by sorting stocks based on idiosyncratic volatility relative to the Fama and French (1993) model. Portfolios are formed every month, based on volatility computed using daily data over the previous month. Portfolio 1 (5) is the portfolio of stocks with the lowest (highest) volatilities. The statistics in the columns labeled Mean and Std.Dev. are measured in monthly percentage terms and apply to total, not excess returns. The row $5 - 1$ refers to the difference in monthly returns between portfolio 5 and portfolio 1. The alpha column report Jensen's alpha with respect to the Fama-French (1993) three-factor model. Robust Newey-West (1987) t-statistics are reported in square brackets. The sample period is January 1963 to December 2015.

Table 4: Portfolios Sorted by the Filtered Estimates of Expected Long-Run Volatility

Panel A: Short and Long-Run Volatility (SL) Model				
Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	1.11	4.47	42.2%	0.19
				[1.77]
2	1.04	5.18	31.1%	0.05
				[0.60]
3	0.85	7.17	16.1%	-0.15
				[-1.08]
4	0.64	8.88	8.1%	-0.40
				[-2.29]
5 (high)	-0.24	10.37	2.6%	-1.50
				[-6.53]
5-1	-1.33	8.65		-2.08
	[-3.34]			[-7.70]

Panel B: Permanent and Transitory Volatility (PT) Model				
Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	1.09	4.28	41.9%	0.18
				[1.86]
2	1.07	5.17	31.1%	0.05
				[0.58]
3	0.83	7.12	16.2%	-0.18
				[-1.32]
4	0.63	9.30	8.1%	-0.50
				[-2.70]
5 (high)	-0.21	10.13	2.6%	-1.54
				[-7.56]
5-1	-1.41	8.74		-2.20
	[-3.72]			[-8.70]

Portfolios are formed every month based on expected volatility of the filtered estimtates $\hat{l}_t$. Portfolio 1 (5) is the portfolio of stocks with the lowest (highest) volatilities. The statistics in the columns labeled Mean and Std.Dev. are measured in monthly percentage terms and apply to total, not excess returns. The row $5 - 1$ refers to the difference in monthly returns between portfolio 5 and portfolio 1. The alpha column report Jensen's alpha with respect to the Fama-French (1993) three factor model. Robust Newey-West (1987) t-statistics are reported in square brackets. The sample period is January 1963 to December 2015.

Table 5: Portfolios Sorted by the Smoothed Estimates of Expected Long-Run Volatility

Panel A: Short and Long-Run Volatility (SL) Model				
Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	1.09	4.01	41.9%	0.23
				[2.99]
2	1.13	5.20	31.5%	0.01
				[1.09]
3	0.83	7.23	15.8%	-0.23
				[-1.72]
4	0.24	9.46	8.2%	-0.98
				[-4.46]
5 (high)	-0.34	12.69	2.6%	-1.75
				[-5.05]
5-1	-1.55	11.50		-2.40
	[-2.91]			[-6.42]
Panel B: Permanent and Transitory Volatility (PT) Model				
Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	1.13	4.06	42.4%	0.26
				[3.22]
2	1.08	5.19	31.2%	0.05
				[0.56]
3	0.84	7.30	16.1%	-0.23
				[-1.70]
4	0.19	9.90	8.0%	-0.98
				[-4.06]
5 (high)	-1.22	12.16	2.4%	-2.68
				[-8.20]
5-1	-2.51	10.96		-3.37
	[-4.76]			[-9.43]

Portfolios are formed every month based on expected volatility from the smoothed estimates $\tilde{l}_t$. Portfolio 1 (5) is the portfolio of stocks with the lowest (highest) volatilities. The statistics in the columns labeled Mean and Std.Dev. are measured in monthly percentage terms and apply to total, not excess returns. The row $5-1$ refers to the difference in monthly returns between portfolio 5 and portfolio 1. The alpha column report Jensen's alpha with respect to the Fama-French (1993) three factor model. Robust Newey-West (1987) t-statistics are reported in square brackets. The sample period is January 1963 to December 2015.

Table 6: Portfolios Sorted by Estimates of Expected Short-Run Volatility

Panel A: The Filtered Estimates of Expected Short-Run Volatility				
Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	1.01	5.11	17.2%	0.04
				[0.38]
2	0.91	5.10	22.6%	-0.12
				[-1.77]
3	1.08	5.29	20.1%	0.09
				[0.87]
4	1.02	5.08	23.0%	0.03
				[0.35]
5 (high)	1.06	5.46	17.0%	0.06
				[0.55]
5-1	0.04	3.10		
	[0.30]			
Panel B: The Smoothed Estimates of Expected Short-Run Volatility				
Rank	Mean	Std. Dev.	% Mkt Share	FF-3 alpha
1 (low)	0.93	4.95	58.8%	-0.02
				[-0.34]
2	1.01	4.79	13.5%	0.07
				[0.68]
3	1.01	4.93	11.7%	0.02
				[0.29]
4	1.19	6.37	9.5%	0.06
				[0.38]
5 (high)	1.37	8.38	6.5%	-0.01
				[-0.07]
5-1	0.43	5.80		
	[-1.68]			

Portfolios are formed every month based on expected volatility of $\hat{s}_t$ or $\tilde{s}_t$. Portfolio 1 (5) is the portfolio of stocks with the lowest (highest) volatilities. The statistics in the columns labeled Mean and Std.Dev. are measured in monthly percentage terms and apply to total, not excess returns. The row $5-1$ refers to the difference in monthly returns between portfolio 5 and portfolio 1. The alpha column report Jensen's alpha with respect to the Fama-French (1993) three-factor model. Robust Newey-West (1987) t-statistics are reported in square brackets. The sample period is January 1963 to December 2015.

Table 7: Relationship between Idiosyncratic Risk and Expected Returns: Cross-Sectional Evidence

Panel A: Short and Long-Run Volatility (SL) Model

Beta	log(ME)	log(BE/ME)	$\log v_t$	$\hat{s}_t$	$\hat{l}_t$	$\tilde{s}_t$	$\tilde{l}_t$	$Ret(-1)$	μ_t
-0.24	-0.08	0.19							
[-0.87]	[-1.50]	[3.17]							
0.00	-0.17	0.14	-0.52						
[0.01]	[-3.91]	[2.50]	[-4.83]						
-0.04	-0.10	0.20	-0.30					-0.06	
[-0.13]	[-2.53]	[3.58]	[-2.64]					[-11.50]	
-0.27	-0.12	0.22		2.96	-0.93				
[-0.97]	[-2.38]	[3.63]		[4.67]	[-4.50]				
-0.21	-0.08	0.25		3.75	-0.77			-0.06	
[-0.71]	[-1.65]	[4.30]		[5.84]	[-3.67]			[-12.10]	
-0.06	-0.11	0.21		4.01	-0.66			-0.06	-0.29
[-0.21]	[-2.72	[3.77]		[6.12]	[-3.35]			[-11.7]	[-2.48]
0.68	-0.32	0.01				2.34	-1.67		
[2.49]	[-7.28]	[0.32]				[12.48]	[-6.52]		
0.52	-0.23	0.07				2.12	-1.39	-0.07	
[1.58]	[-4.33]	[1.21]				[7.95]	[-3.96]	[-10.17]	
0.39	-0.23	0.09				2.62	-2.14	-0.06	0.91
[1.14]	[-4.06]	[1.42]				[7.10]	[-4.34]	[-6.56]	[2.26]

Stocks are assigned the post-ranking β of the size-β portfolios they are in every month. BE is the book value of common equity plus balance-sheet deferred taxes. The accounting ratios BE/ME are measured using market equity ME in December of year $t-1$. Firm size $Ln(ME)$ is measures in June of year t. In the regressions, these are values of the explanatory variables for individual stocks are matched with CRSP returns for the months from July of year t to June of year $t+1$. The gap between the accounting data and the returns ensures that the accounting data are available prior to returns. The average coefficient is the time-series average of monthly regression coefficients for July 1963 to December 2015, and the t-statistics is the average coefficient divided by its time-series standard error. The t-statistic is reported in brackets. For the PT model, the numbers in parentheses are the time series median of coefficients in the FM regression.

Table 8: Relationship between Idiosyncratic Risk and Expected Returns: Cross-Sectional Evidence

Panel B: Permanent and Transitory Volatility (PT) Model						
Beta	ln(ME)	ln(BE/ME)	$\hat{l}_t$	$\tilde{l}_t$	$Ret(-1)$	μ_t
0.05	-0.20	0.13	-0.65			
[0.19]	[-4.73]	[2.33]	[-3.68]			
0.10	-0.27	0.10		-0.88		
[0.40]	[-6.40]	[1.81]		[-3.93]		
0.05	-0.14	0.18	-0.49		-0.06	
[0.19]	[-3.44]	[3.31]	[-2.78]		[-12.06]	
0.06	-0.15	0.18	-0.58		-0.06	0.58
[0.22]	[-3.95]	[3.23]	[-3.08]		[-12.1]	[4.32]
0.13	-0.22	0.14		-0.86	-0.06	
[0.54]	[-5.55]	[2.58]		[-3.81]	[-13.39]	
0.53	-0.31	0.07		-1.68	0.05	9.08
[2.20]	[-7.89]	[1.40]		[-7.87]	[-8.79]	[36.02]

Stocks are assigned the post-ranking β of the size-β portfolios they are in every month. BE is the book value of common equity plus balance-sheet deferred taxes. The accounting ratios BE/ME are measured using market equity ME in December of year $t-1$. Firm size $Ln(ME)$ is measures in June of year t. In the regressions, these are values of the explanatory variables for individual stocks are matched with CRSP returns for the months from July of year t to June of year $t+1$. The gap between the accounting data and the returns ensures that the accounting data are available prior to returns. The average coefficient is the time-series average of monthly regression coefficients for July 1963 to December 2015, and the t-statistics is the average coefficient divided by its time-series standard error. The t-statistic is reported in brackets. For the PT model, the numbers in parentheses are the time series median of coefficients in the FM regression.