Singapore Management University

Institutional Knowledge at Singapore Management University

Research Collection Lee Kong Chian School Of Business

Lee Kong Chian School of Business

1-2009

Idiosyncratic Risk and the Cross-Section of Expected Stock Returns

Fangjian FU
Singapore Management University, fjfu@smu.edu.sg

Follow this and additional works at: https://ink.library.smu.edu.sg/lkcsb_research

Part of the Finance and Financial Management Commons, and the Portfolio and Security Analysis Commons

Citation

Idiosyncratic Risk and the Cross-Section of Expected Stock Returns

Fangjian Fu[*]

Singapore Management University
Lee Kong Chian School of Business
50 Stamford Road, Singapore 178899
Tel: +65 6828 0244
Fax: +65 6828 0427
Email: fjfu@smu.edu.sg

This Version: May 2008

Abstract

Theories such as Merton (1987) predict a positive relation between idiosyncratic risk and expected return when investors do not diversify their portfolio. Ang, Hodrick, Xing, and Zhang (2006, Journal of Finance 61, 259-299) however find that monthly stock returns are negatively related to the one-month lagged idiosyncratic volatilities. I show that idiosyncratic volatilities are time-varying and thus their findings should not be used to imply the relation between idiosyncratic risk and expected return. Using the exponential GARCH models to estimate expected idiosyncratic volatilities, I find a significantly positive relation between the estimated conditional idiosyncratic volatilities and expected returns. Further evidence suggests that Ang et al.'s findings are largely explained by the return reversal of a subset of small stocks with high idiosyncratic volatilities.

---

1. Introduction

Modern portfolio theory suggests that investors hold a portfolio of stocks to diversify idiosyncratic risk. The capital asset pricing model (CAPM) builds on the portfolio theory and predicts that all investors hold the market portfolio in equilibrium. As a result, only systematic risk is priced in equilibrium and idiosyncratic risk is not. For various reasons, investors in reality however may not hold perfectly diversified portfolios. For instance, Goetzmann and Kumar (2004) show that, based on a sample of more than 62,000 household investors in the period of 1991 to 1996, more than 25 percent of the investor portfolios contain only one stock, over a half of the investor portfolios contain no more than three stocks and less than ten percent of the investor portfolios contain more than ten stocks. After examining the standard deviation of portfolio returns, Campbell, Lettau, Malkiel, and Xu (2001, p. 25) however suggest that "the number of randomly selected stocks needed to achieve relatively complete portfolio diversification" is about 50.

Various theories assuming under-diversification predict that idiosyncratic risk is positively related to the expected stock returns in the cross section. Among them are Levy (1978), Merton (1987), and Malkiel and Xu (2002). Under-diversified investors demand a return compensation for bearing idiosyncratic risk. But a recent paper by Ang, Hodrick, Xing, and Zhang (2006, AHXZ hereafter) finds that, in the cross-section of stocks, high idiosyncratic volatility in this month predicts abysmally low average returns in the next month, which they call "a substantive puzzle". Their study poses three important questions: (1) Do the findings imply that the relation between idiosyncratic risk and expected return is negative? (2) If not necessary, what is the true empirical relation? (3) If the true relation is not negative, how to explain their findings?

I attempt to answer these three questions in this paper. The sample of data includes stocks traded on the NYSE, the AMEX, and the NASDAQ during the period from July 1963 to December 2006. I first identify that idiosyncratic volatilities, unlike some firm characteristics, are very volatile over time. For an average individual stock, the standard deviation of its monthly idiosyncratic volatilities is 55 percent of the mean. In order to explain expected returns, the theoretically correct variable should be the expected idiosyncratic volatilities in the same period that the expected returns are measured. Since idiosyncratic volatilities are time-varying, the one-month lagged idiosyncratic volatility may not be an appropriate proxy for the expected idiosyncratic volatility of this month. Indeed, the average first-order autocorrelation of idiosyncratic volatility is only 0.33 in my sample. Dickey-Fuller tests further show that, for nine out of ten stocks, their idiosyncratic volatility does not follow a random walk process. These findings suggest that the negative relation between the lagged idiosyncratic volatility and average returns in AHXZ (2006) does not imply that the relation between idiosyncratic risk and expected return is negative. The lagged idiosyncratic volatility might not be a good estimate of expected idiosyncratic volatility.

In order to capture the time-varying property of idiosyncratic risk, I employ the Exponential Generalized AutoRegressive Conditional Heteroskedasticity (EGARCH) models and out-of-sample data to estimate expected idiosyncratic volatilities. I then run Fama-MacBeth regressions of monthly stock returns on the EGARCH estimates and other firm characteristics that are known to explain cross-sectional returns. I find that returns are positively related to the EGARCH-estimated conditional idiosyncratic volatilities. The positive relation is both statistically and economically significant. On average, a stock that has a conditional idiosyncratic volatility of one standard deviation

higher than the other stock earns a return of about 1 percent higher in a month. The coefficient of determination for the cross-sectional regressions also increases significantly after including the conditional idiosyncratic volatility. A zero-investment portfolio that is long in the 10% of the highest and short in the 10% of the lowest conditional idiosyncratic volatilities earns a positive return of 1.75% in a month. These findings support the theory prediction that idiosyncratic risk is positively related to expected returns.

However, AHXZ's (2006) finding that stocks with high idiosyncratic volatilities tend to have abnormally low returns in the subsequent month is still puzzling, especially given that the contemporaneous relation between return and volatility is significantly positive. I show that their results can largely be explained by the return reversal of stocks with high idiosyncratic volatilities. Specifically, stocks with high idiosyncratic volatilities are shown to have high contemporaneous returns. The positive abnormal returns tend to reverse, resulting in negative abnormal returns in the following month.[1] Moreover, these stocks are small in size. The 40% stocks with the highest idiosyncratic volatilities only contribute to 9% of the total market capitalization. Since transaction costs for small firms are notoriously high, and idiosyncratic risk

---

[1] In order to be consistent with AHXZ (2006) for a better comparison, I also use the Fama-French three-factor model as the benchmark model for expected returns. AHXZ's arguments are mainly based on the portfolio of the highest idiosyncratic volatility that yields a negative abnormal return in the following month.

increases holding costs and makes arbitrage more costly (Pontiff, 2006), it is dubious that the negative relation would present a true profitable opportunity.[2]

AHXZ's (2006) findings have attracted much attention recently. Bali and Cakici (2008) suggest that AHXZ's results are sensitive to (i) data frequency used to estimate idiosyncratic volatility, (ii) weighting scheme used to compute average portfolio returns, (iii) breakpoints utilized to sort stocks into quintile portfolios, and (iv) using a screen for size, price, and liquidity, and therefore are not robust. Huang, Liu, Rhee, and Zhang (2007) point out that AHXZ's results are driven by monthly stock return reversals. After controlling for the difference in the past-month returns, the negative relation between average return and the lagged idiosyncratic volatility disappears. Using a different method that more closely focuses on AHXZ's findings, I point to the same conclusion. Boyer, Mitton, and Vorkink (2007) suggest that idiosyncratic volatility is a good predictor of expected skewness – an explanatory variable of cross-sectional returns (Harvey and Siddique, 2000). The negative relation greatly reduces after controlling for expected skewness. Jiang, Xu, and Yao (2006) argue that high idiosyncratic volatility and low future returns are both related to a lack of information disclosure among firms with poor earnings prospects. Investors underreact to earnings information in idiosyncratic volatility.

In their more recent work, Ang, Hodrick, Xing, and Zhang (2008) find that the negative relation between average return and the lagged idiosyncratic volatility also

---

[2] In their more recent work, AHXZ (2008) admit that their documented negative relation implies "not necessarily a relation that involves expected volatility" but argue that it suggests a profitable trading strategy.

exists in other G7 countries. However, Brockman and Schutte (2007) follow my EGARCH method to estimate conditional idiosyncratic volatility and confirm that the relation between stock return and conditional idiosyncratic volatility is also positive in international data. Similarly Spiegel and Wang (2006) and Eiling (2006) adopt the EGARCH models to estimate conditional idiosyncratic volatility and both find the positive relation in the U.S. data. Spiegel and Wang also show that idiosyncratic volatility swamped liquidity in explaining the cross-sectional variation of average returns but not vice versa. Eiling shows that the idiosyncratic risk premium is related to hedging demand due to investors' non-tradable human capital. Chua, Goh, and Zhang (2007) model idiosyncratic volatility as an AR(2) process and decompose it into an expected and an unexpected components. Controlling for the unexpected idiosyncratic volatility, they also find the relation between expected return and expected idiosyncratic volatility is significantly positive.

The reminder of the paper proceeds as follows. In Section 2, I measure monthly idiosyncratic volatility, examine its time-series properties, and estimate the one-month-ahead conditional idiosyncratic volatilities using EGARCH models. In Section 3, I examine the cross-sectional relation between conditional idiosyncratic volatilities and expected returns. I replicate and explain AHXZ's findings in Section 4. Section 5 concludes.

2. Idiosyncratic volatility and its time-series property

The goal of the paper is to examine whether under-diversified investors are compensated for bearing idiosyncratic risk. From the theory perspective, the risk and return tradeoff should be contemporaneous. Investors earn returns for bearing the risk

in the same period. Therefore if idiosyncratic volatility, as a natural proxy for idiosyncratic risk, is priced, we expect to observe a positive empirical relation between expected return and expected idiosyncratic volatility. However, neither expected return nor expected idiosyncratic risk is observable. A conventional practice is to use the realized return as the dependent variable in cross-sectional regressions where the realized return is assumed to be the sum of the expected return and a random error.[3] The expected idiosyncratic volatility and other control variables are put on the right-hand-side of the regressions.

$$R_{it} = \gamma_{0t} + \gamma_{1t}E_{t-1}[IVOL_{it}] + \sum_{k=2}^{K}\gamma_{kt}E_{t-1}[X_{kit}] + \varepsilon_{it} \qquad i=1,2,\cdots,N_t, \ t=1,2,\cdots,T \qquad (1)$$

The dependent variable is the realized returns for stock i in period t. $E_{t-1}[\cdot]$ stands for the function of expectation conditional on the information set at t-1. $IVOL_{it}$ represents the idiosyncratic volatility of stock i during period t. $E_{t-1}[IVOL_{it}]$ is the expected idiosyncratic volatility for stock i at time t conditional on the information set at time t-1. X_{kit} represents other explanatory variables of cross-sectional returns. N_t is the total number of stocks at t and T is the total number of time periods. The null hypothesis is $\gamma_{1t} = 0$, that is, idiosyncratic risk is not priced. Existing theories assuming under-diversification such as Merton (1987) predicts that $\gamma_{1t} > 0$.

It is crucial to have a quality estimate of $E_{t-1}[IVOL_{it}]$ -- the expected idiosyncratic volatility. If idiosyncratic risk is highly persistent as following a random walk process, we can simply use the lagged value as an estimate of the expected value. In this case,

[3] See, for example, Fama and French (1992), Chordia, Subrahmanyam and Anshuman (2001), Easley, Hvidkjaer, and O'Hara (2002) among others.

idiosyncratic risk resembles some firm characteristics such as size and the market-to-book ratio of equity. Fama and French (1992), for example, use market capitalizations and book-to-market equity ratios of the current year to explain the cross-sectional variation of monthly returns in the next year. However, we have no prior reasons to presume high persistence in idiosyncratic risk. Idiosyncratic risk reflects firm-specific information that is volatile in its nature. Many factors could contribute to the time-varying nature of firm-specific information. For instance, disclosure of earnings information is periodical and infrequent; the supply and demand of certain firms are subject to seasonal variations; competitors' moves may also bring impact on the firm's profitability. I examine the time-series property of idiosyncratic volatility in this section. The results indeed suggest that idiosyncratic volatility varies over time substantially. Therefore, a quality estimate of conditional idiosyncratic volatility is demanded to draw an appropriate inference on the relation between idiosyncratic risk and expected return.

2.1. Estimation of idiosyncratic volatility

Idiosyncratic risk is defined as the risk that is unique to a specific firm, so it is also called firm-specific risk. By definition, idiosyncratic risk is independent of the common movement of the market. Following AHXZ (2006), I measure the idiosyncratic risk of an individual stock as follows. In every month, daily excess returns of individual stocks are regressed on the daily Fama-French three factors (Fama and French, 1993, 1996): (i) the excess return on a broad market portfolio $(R_m - r_f)$, (ii) the difference between the return on a portfolio of small stocks and the return on a portfolio of large stocks (*SMB*, small minus big), and (iii) the difference between the return on a portfolio

of high book-to-market stocks and the return on a portfolio of low book-to-market stocks (*HML*, high minus low),

$$R_{i\tau} - r_{\tau} = \alpha_{it} + b_{it}(R_{m\tau} - r_{\tau}) + s_{it}SMB_{\tau} + h_{it}HML_{\tau} + \varepsilon_{i\tau}. \qquad (2)$$

τ is the subscript for the day and t is the subscript for the month, $\tau \in t$, and b_i, s_i, and h_i are factor sensitivities or loadings. Daily stock returns are obtained from the CRSP. I include stocks traded on the NYSE, the AMEX or the Nasdaq during the period of July 1963 to December 2006. The daily factor data are downloaded from Kenneth R. French's website. [4] I perform a time-series regression for each stock in each month. The idiosyncratic volatility of a stock is computed as the standard deviation of the regression residuals. To reduce the impact of infrequent trading on idiosyncratic volatility estimates, I require a minimum of 15 trading days in a month for which the CRSP reports both a daily return and non-zero trading volume.[5] Moreover, I transform the standard deviation of daily return residuals to a monthly return residual by multiplying the daily standard deviation by the square root of the number of trading days in that month.[6] In the pooled sample of 2,946,521 firm-month observations, the mean monthly idiosyncratic volatility (*IVOL*) is 14.17% with a standard deviation of 13.91%.

2.2. Time-series property of idiosyncratic volatility

---

[4] The website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. I thank Kenneth French for making these data available.

[5] The trading days per month in my sample ranges from 15 to 23 days with a mean and median of 21 days. Only about 1% of firm-month observations have fewer than 19 trading days.

[6] A similar procedure is used by French, Schwert, and Stambaugh (1987) and Schwert (1989).

Table 1 presents the time-series property of individual stock idiosyncratic volatilities. I first compute the time-series statistics of idiosyncratic volatility for each firm and then summarize the mean statistics across about 26,000 firms in the sample. The time-series mean *IVOL* is on average 16.87% across stocks and the mean standard deviation is 9.94%. The mean coefficient of variation is 0.55, indicating that the standard deviation of *IVOL* for an average stock is 55% of its time-series mean. This suggests that individual stock idiosyncratic volatilities vary substantially over time. The last columns report the autocorrelations of *IVOLs*. The mean autocorrelation is 0.33 at the first lag and decays slowly. I also report the statistics of the changes in the logarithm of the idiosyncratic volatility ($Ln\left(\dfrac{IVOL_t}{IVOL_{t-1}}\right)$). The autocorrelation of this new variable is -0.42 at the first lag and close to zero at lags of higher orders. This evidence suggests that the first differences of $Ln(IVOL_t)$ for quite a few firms might follow a first-order moving average process.

AHXZ (2006) draw expected return implications on the basis of the observed relation between monthly stock returns and the one-month lagged *IVOL*. Their empirical methods implicitly assume that the time-series idiosyncratic volatility can be approximated by a random walk process. The first-order autocorrelation for a random walk process should be one, and the first differences of a random walk are a white noise and therefore the autocorrelation should be zero at all lags. The autocorrelation evidence in Table 1 suggests that the random walk hypothesis is not appropriate for a typical stock's idiosyncratic volatility process. To illustrate this point further, I run the following time-series regression for each stock,

$$IVOL_{i,t+1} - IVOL_{i,t} = \gamma_{0i} + \gamma_{1i}IVOL_{i,t} + \eta_i, \qquad t = 1, 2, \cdots T, \, i = 1, 2, \cdots, N. \tag{3}$$

The coefficient γ_{1i} should be indistinguishable from zero if the time-series of $IVOL_{it}$ follow a random walk. This is a standard unit-root test. For each time series of $IVOL$, I estimate the coefficient γ_{1i} and then compare its t-statistic with the Dickey-Fuller critical values for the unit-root tests. In Table 2, I report the cross-firm mean, median, the lower and upper quartiles of the γ_{1i} estimates and the associated t-statistics. The last column reports the percentage of firms for which the random walk hypothesis is rejected at the 1% level. For the purpose of regressions, I require firms to have at least 30 months of consecutive observations ($T_i \geq 30$ for every stock). [7] This requirement reduces the number of firms to 20,979. The mean γ_{1i} among these firms is -0.61 and the mean t-statistic of γ_{1i} is -6.81. According to the Dickey-Fuller critical values of t-statistics (Fuller, 1996), I reject the null hypothesis of a random walk in 90% of the firms. Examinations on $Ln(IVOL)$ yield very similar results, which are also reported in Table 2. The results suggest that it is not appropriate to describe a typical stock's idiosyncratic volatility process as a random walk.[8] Put differently, using this month's idiosyncratic volatility to approximate the value in the next month could introduce severe measurement errors. As

---

[7] The results are robust to setting *T>12* or *T>24.*

[8] Cross-sectional return studies often use the lagged variables of firm characteristics to explain future returns. For example, Fama and French (1992) use market capitalization and the book-to-market equity ratio of the current year to explain the cross-sectional returns in the next year. The underlying justification for this practice is that these firm characteristics are fairly persistent. I perform the same unit root test on the process of firm market capitalization and indeed am not able to reject the null hypothesis of random walk in 98% of the observations.

a result, AHXZ's findings should not be used to draw inference on the relation between idiosyncratic risk and expected return.

2.3. Estimation of expected idiosyncratic volatility

In order to examine the relation between expected return and expected idiosyncratic volatility, we need a better model to capture the time-varying property of idiosyncratic volatility. I resort to the exponential generalized autoregressive conditional heteroskedasticity (EGARCH) models to achieve this goal.

Engle (1982) proposes the autoregressive conditional heteroskedasticity (ARCH) model to represent a series with changing volatility. It proves to be an effective tool in modeling time-series behavior of many economic variables, especially financial market data. The ARCH model is attractive because the variance and the mean process are estimated jointly. Applying to stock market returns, it implicitly assumes that investors update their estimates of the mean and variance of returns each period using the newly revealed surprises in last period's returns. Bollerslev (1986) extends the ARCH model to GARCH, or the generalized autoregressive conditional heteroskedasticity model. The GARCH model provides a more flexible framework to capture the dynamic structure of conditional variances (volatilities). A step further, Nelson (1991) proposes an exponential GARCH (EGARCH) model to catch the asymmetric property of volatility, namely that the return volatility increases after a stock price drop. This phenomenon is also called "leverage effects" because the drop of stock price mechanically increases the leverage ratio and thus the risk of the firm.

GARCH models have been widely used to model the conditional volatility of returns. For example, French, Schwert, and Stambaugh (1987) model the market

volatility by a GARCH (1, 2) process and find that the market risk premium is positively related to the conditional market volatility. Bollerslev, Engle, and Wooldridge (1988) use a multivariate GARCH model to demonstrate time-varying risk premiums. GARCH models are of various types. My objective is to select a GARCH model that well describes the time-series idiosyncratic volatility of individual stock returns. Pioneered by Pagan and Schwert (1990), studies have suggested a number of approaches to compare alternative GARCH (and non-parametric) specifications. Pagan and Schwert fit a number of different models to monthly U.S. stock returns and find that Nelson's EGARCH model is the best in overall performance. Engle and Mustafa (1992) assess the specification of conditional variance models based on the observed prices for stock options. Specifically they use the option prices to compute the implied variances, which are then regarded as the benchmark for the estimates from various time-series models. They find that simple GARCH and EGARCH models perform the best among their selected time-series models. Emphasizing the importance of the asymmetry of the volatility response to news, Engle and Ng (1993) test the specifications of various volatility models using Lagrange Multiplier tests. They also conclude that Nelson's EGARCH specification does a good job in capturing the asymmetry of conditional volatilities. In addition, EGARCH models do not need to restrict parameter values to avoid negative variance as other ARCH and GARCH models do.

Weighing all evidence, I choose to model idiosyncratic volatilities by the EGARCH (p, q) model, in which $1 \le p \le 3, 1 \le q \le 3$. The explicit functional forms are as follows:

$$R_{it} - r_t = \alpha_i + \beta_i (R_{mt} - r_t) + s_i SMB_t + h_i HML_t + \varepsilon_{it}, \quad \varepsilon_{it} \sim N(0, \sigma_{it}^2), \tag{4}$$

$$\ln \sigma_{it}^2 = a_i + \sum_{l=1}^{p} b_{i,l} \ln \sigma_{i,t-l}^2 + \sum_{k=1}^{q} c_{i,k} \left\{ \theta \left(\frac{\varepsilon_{i,t-k}}{\sigma_{i,t-k}} \right) + \gamma \left[\left| \frac{\varepsilon_{i,t-k}}{\sigma_{i,t-k}} \right| - (2/\pi)^{1/2} \right] \right\}. \tag{5}$$

I describe the monthly return process by the Fama-French three-factor model as in equation (4). The conditional (on the information set at time t-1) distribution of residual ε_{it} is assumed to be normal with the mean of zero and the variance of σ_{it}^2. My objective is to estimate the conditional variance σ_{it}^2. It is a function of the past p-period of residual variance and q-period of return shocks as specified by equation (5). Permutation of these orders yields nine different EGARCH models: EGARCH (1, 1), EGARCH (1, 2), EGARCH (1, 3), EGARCH (2, 1), EGARCH (2, 2), EGARCH (2, 3), EGARCH (3, 1), EGARCH (3, 2), and EGARCH (3, 3). Each model is employed independently for each individual stock. Therefore, if a stock's idiosyncratic volatility process as of month t converges under all the nine models, I would have nine estimated conditional idiosyncratic volatilities at month $t+1$. The estimate generated by the model of the lowest Akaike Information Criterion (AIC) is chosen.[9] I also require firms to have at least 30 monthly returns to be eligible for estimation.

My EGARCH (p, q) model involves $p+q+3$ parameters. Using the full period data to estimate these parameters, though prevalent in early studies, incurs a look-ahead

---

[9] I also use the maximum log likelihood or the Schwartz information criterion to select the model of conditional idiosyncratic volatilities. The final results are not sensitive to these alternative selection criteria. The choice of the maximum number of lags to be three does not drive the final results either. The results also hold for alternative specification such as two or four.

bias.[10] To avoid this concern, I estimate EGARCH parameters by using an expanding window of data with a requirement of 30 minimum observations. In other words, the EGARCH parameters used to forecast conditional idiosyncratic volatility at month t are estimated on the basis of the data up through month t-1. This also applies to the model selection.

The estimated conditional idiosyncratic volatility, denoted by *E(IVOL),* will be used in the cross-sectional return tests in the next section. The mean *E(IVOL)* is 12.67% with a standard deviation of 10.91% in the pooled sample. The correlation between *IVOL* and *E(IVOL)* is 0.46 and is statistically significant at the 1% level. Empirical evidence confirms the importance to have more than one lag in estimating E(IVOL). Of all the estimates, only 26.67% are yielded by the EGARCH(1, q) models while 40% are generated by the EGARCH(3, q) models. In particular, EGARCH(1, 1) is the best-fitting model for the fewest number of firm-month observations (7.41%) and EGARCH (3, 1) is the best-fitting model for the most number of observations (16.58%).

3. Cross-Sectional Return Tests

3.1. Data and variables

In this section, I investigate the cross-sectional relation between average stock returns and the estimated conditional idiosyncratic volatilities. I examine stocks traded

---

[10] It is an empirical question how serious the look-ahead bias would be. To estimate conditional market volatility, French, Stambaugh, and Schwert (1987) use the full period data to estimate their GARCH model parameters. They show that assuming time-varying parameters do not change their results. I also find the same results by using the full period data to estimate EGARCH model parameters.

on the NYSE, the AMEX, and the Nasdaq during the period from July 1963 to December 2006 - 522 months in total. The data of monthly stock returns are obtained from the CRSP. Table 3 presents the variable descriptive statistics of the pooled sample. The mean monthly returns (*RET*) in my sample period are 1.18% and the mean excess return (*XRET*, raw return net of one-month T-bill rate) is 0.71%.[11] The mean idiosyncratic volatility is 14.17% and the mean expected idiosyncratic volatility is 12.67%.

The measure of systematic risk, *BETA*, is constructed as in Fama and French (1992). In each month, I use the previous 60 months of returns to estimate firm betas (β) by the market model. Stocks are assigned to 10x10 portfolios on the basis of size and β. This procedure rolls every month. I then compute the equal-weighted portfolio returns. For each size-β portfolio, I run the full-period time-series regression of the portfolio return on the current and the prior month's value-weighted market returns. The portfolio *BETA* is estimated as the sum of the slopes of these two market returns. The sum is meant to adjust for the effects of nonsynchronous trading (Dimson, 1979). Finally I allocate the *BETA* of a size-β portfolio to each stock in the portfolio. These are the *BETAs* to be used in the cross-sectional regressions of individual stock returns. The mean *BETA* is 1.22 and the median 1.17.

---

[11] In order to avoid the influence of some extremely high returns and possible data recording errors, I exclude 333 observations that have a monthly return greater than 300%. This consists of only 0.0001% of the whole sample (which has about 3 million firm-month observations).

Earlier studies document that firm size, the ratio of book-to-market equity, liquidity and its variance, and past returns have effects on cross-sectional returns.[12] I control for these variables in the following cross-sectional tests. Their summary statistics are also reported in Table 3. Firm size is measured by the market value of equity (*ME*) -- the product of monthly closing price and outstanding share numbers. Following Fama and French (1992), I construct book-to-market equity (*BE/ME*) as the fiscal-yearend book value of common equity divided by the calendar-yearend market value of equity. Due to the annual frequency of BE, this variable is updated yearly. In order to catch the momentum effects, I construct the variable *RET(-2, -7)*, the compound gross return from month *t*-7 to *t*-2 (inclusive) where *t* represents the month of expected return. The return of *t*-1 is excluded to avoid any spurious association between subsequent month returns caused by thin trading or the bid-ask spread effects. Jegadeesh (1990) shows that thin trading causes returns to exhibit first-order negative serial correlations. Following Chordia, Subrahmanyam and Anshuman (2001), liquidity is measured by the average turnover ratio of the previous 36 months (*TURN*) and the coefficient of variation of these 36 monthly turnovers (*CVTURN*).

3.2. Cross-sectional simple correlations

---

[12] See, for example, Fama and French (1992) for the effects of size and book-to-market equity, Jagadeesh and Titman (1993) for the effects of past returns, Amihud and Mendelson (1986), Brennan and Subrahmanyam (1996), and Amihud (2002) for the effects of liquidity, and Chordia, Subrahmanyam and Anshuman (2001) for the effects of the variance of liquidity on cross-sectional returns.

I investigate the bivariate relations between these variables. The correlation between return and idiosyncratic volatility can be regarded as a univariate test. Variables of *ME, BE/ME, TURN, CVTURN* are transformed to their natural logarithm because they are significantly skewed. I estimate the simple correlations between these variables in each month and then compute their time-series means. Table 4 presents the time-series mean correlation coefficients. The coefficients followed by * are significant at the 1% level based on their time-series standard error. The correlation between the monthly return and the contemporaneous idiosyncratic volatility is 0.135 and statistically significant at the 1% level (*t*-stat = 14.09). The correlation between return and the one-month lagged IVOL is however -0.016 with a *t*-stat of -2.79 (not reported in the Table). The contrasting results, in another way, suggest that the lagged IVOL might not be a good proxy for the expected value in next month. The correlation between return and the conditional idiosyncratic volatility is 0.09 and also statistically significant at the 1% level (*t*-stat =13.24). The univariate tests therefore imply a positive relation between return and idiosyncratic risk. Consistent with the findings in the literature, the returns are negatively related to size and liquidity, and are positively related to *BE/ME* and past returns. As documented in Fama and French (1992), the relation between return and *BETA* is flat. Conditional idiosyncratic volatilities are negatively related to size and the book-to-market equity ratio, and are positively related to *BETA* and the two liquidity variables. Small firms tend to have higher idiosyncratic volatilities than large firms; growth firms tend to have higher idiosyncratic volatilities than value firms; liquid firms tend to have higher idiosyncratic volatilities than illiquid firms. The correlation between IVOL and E(IVOL) is a significant 0.46.

3.3. Month-by-month cross-sectional regressions of individual stocks

I start the empirical analysis by replicating the main results documented by Fama and French (1992), because their paper is very influential in the literature of cross-sectional return studies. By regressing monthly stock returns on beta and various firm characteristics, Fama and French illustrate that size and BE/ME are two significant determinants for cross-sectional returns and that the relation between return and market beta is flat. Their sample period is from July 1963 to December 1990. My sample extends theirs by 16 years to December 2006. Following them, I use Fama and MacBeth (1973) regressions to control the cross-correlation in residuals. Specifically, for each month in the sample period, I run the following cross-sectional regression:

$$R_{it} = \gamma_{0t} + \sum_{k=1}^{K} \gamma_{kt} X_{kit} + \varepsilon_{it}, \qquad i = 1, 2, \cdots, N_t, \ t = 1, 2, \cdots, T, \tag{6}$$

where R_{it} is the realized return on stock i in month t. X_{kit} are the intended explanatory variables of cross-sectional expected returns such as beta, size, book-to-market ratio, and conditional idiosyncratic volatility. The disturbance term, ε_{it}, captures the deviation of the realized return from its expected value. N_t denotes the total number of stocks in month t, which can vary from month to month. The maximum month, T, equals 522 in this study. Equation (6) is essentially equivalent to Equation (1). The final estimate, $\hat{\gamma}_k$, and its variance are given by:

$$\hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T} \hat{\gamma}_{kt}, \tag{7}$$

$$Var(\hat{\gamma}_k) = \frac{\sum_{t=1}^{T} (\hat{\gamma}_{kt} - \hat{\gamma}_k)^2}{T(T-1)}. \tag{8}$$

In other words, the average slope is the time-series mean of the 522 monthly estimates from July 1963 to December 2006. The *t*-statistic is the average slope divided by its time-series standard error, which is the square root of the variance of $\hat{\gamma}_k$ divided by T ($\sqrt{\frac{Var(\hat{\gamma}_k)}{T}}$).

Following Fama and French (1992), I use size (the market capitalization, *ME*) in June to explain the returns of the following 12 months and use *BE/ME* of fiscal year *t* to explain the returns for the months from July of year *t*+1 to June of year *t*+2. The time gap between *BE/ME* and returns ensures that the information of *BE/ME* is available to the public prior to the returns. The estimation and application of *BETAs* also follow Fama and French. Table 5 presents the regression results. The first model replicates Fama and French's three major results. First, the relation between market beta and average stock return is flat. The average slope of *BETA* is not significantly different from zero. Second, size is negatively related to average returns in the cross-section. Small firms on average have higher returns than large firms. Third, *BE/ME* is positively related to average returns. Value firms tend to have higher returns than growth firms.

Cross-sectional return studies have evolved much since Fama and French (1992). Liquidity and momentum are probably the other two most important variables that have impact on cross-sectional returns. Amihud and Mendelson (1986) is among the first to propose a role for transaction costs in asset pricing, since rational investors select assets to maximize their expected return net of transaction costs. They measure liquidity by the bid-ask spread and find stocks with larger spreads are expected to have higher returns. The role played by liquidity is further supported by later studies including Brennan and Subrahmanyam (1996), Datar, Naik and Radcliffe (1998), Chordia,

Subrahmanyam and Anshuman (2001), Amihud (2002), and Pastor and Stambaugh (2003). Jegadeesh and Titman (1993) document that over an intermediate horizon of three to twelve months, past winners on average continue to outperform past losers, so that there is "momentum" in stock prices. In other words, past returns tend to predict future returns. The second model in Table 5 includes the liquidity and momentum variables and confirms these patterns. *TURN* is the average share turnover in the past 36 months, constructed in the same way as in Chordia et al. (2001). Chordia et al. also find that both the level and the volatility of trading activity are related to average returns in the cross-section. Following them, I compute the coefficient of variation of the previous 36 months' turnover (*CVTURN*). Easley, Hvidkjaer and O'Hara (2002) adopt the same measures to control for the effects of liquidity. In order to control for the momentum effects, I construct a past return variable, *RET(-2, -7)*, which is the compound gross return from month *t-7* to month *t-2* assuming the current month is *t*. The return of the immediate prior month (*t-1*) was excluded to avoid any spurious association between subsequent monthly returns caused by thin trading or bid-ask spread effects (Jegadeesh, 1990). Consistent with the previous studies, the coefficient estimates are positive for the past return variable and negative for the two liquidity variables.

Model 3, 4, and 5 in Table 5 yield striking evidence that expected idiosyncratic volatility is positively related to average returns in the cross-section. Model 3 is a univariate regression of return on *E(IVOL)*. Model 4 controls for size and BE/ME, and Model 5 in addition controls for past returns and two liquidity variables. The average slopes of *E(IVOL)* are positive and statistically significant in all the three models. The *t*-statistics are around 10. Moreover, the average *R*-squared increases substantially after including *E(IVOL)* in the regression. The effects of idiosyncratic risk on expected returns

are also economically significant. Since the average slope is over 0.10 the average standard deviation of IVOL is about 10%, a stock that has an IVOL of one standard deviation higher than the other stock would earn an average return of 1% higher in a month. Similarly, as *E(IVOL)* moves from the first quartile to the third quartile, the monthly expected return would increase by more than 1%.

In Model 6, I include the one-month lagged IVOL as an explanatory variable. The regression results qualitatively confirm AHXZ's findings. Monthly returns are negatively related to the lagged IVOL. However, the average slope is only -0.02, which casts doubt on how effective investors can make abnormal returns from the negative relation. Shorting stocks whose IVOL are around the third quartile and longing stocks whose IVOL are around the first quartile yield a monthly abnormal return of 0.2% before accounting for trading costs. Moreover, idiosyncratic volatilities change over time. Trading strategies betting on it need frequent rebalancing and thus are costly.

The last model, Model 7, examines the contemporaneous association between the return and the observed idiosyncratic volatility (estimated based on daily returns). The coefficient of $IVOL_t$ is 0.31 with a *t*-statistic of 20.56. There is a positive and significant association between the realized return and the contemporaneous idiosyncratic risk. From the theory perspective, we are not able to make inferences about expected returns from this regression because of the potential correlation between the unexpected return shock ($R_t - E(R_t)$) and the shock on idiosyncratic volatility ($IVOL_t - E(IVOL_t)$).[13] The

---

[13]The covariance between the realized return and the realized idiosyncratic volatility can be decomposed into four components as follows:

regression results, however, still serve as a reference for comparison. The results of Model 5 rather than Model 6 are close to the results of Model 7. This provides us some additional confidence on the positive relation between expected return and idiosyncratic risk.

One result is intriguing. The average slope of size changes sign after including *E(IVOL)* (or *IVOL*) in the regression. Controlling for conditional idiosyncratic volatility, large firms have higher average returns than small firms. This finding contrasts to the widely documented "size effect" that small firms have higher average returns than large firms, but supports one prediction of Merton's (1987) model that, all else equal, larger firms have higher expected returns. Merton explicitly points out that the findings of the "size effect" are due to the omitted controls for other factors such as idiosyncratic risk and investor base (Merton (1987), the last paragraph of page 496). My evidence lends direct support to Merton's prediction.

3.4. Return analysis of portfolios formed on E(IVOL)

The evidence from the Fama-MacBeth cross-sectional regressions suggests a positive relation between conditional idiosyncratic volatility and average stock returns. Next I examine the returns of portfolios formed on the sorting of E(IVOL). This is interesting because the portfolio-based approach produces easy-to-interpret returns on a

---

$$Cov(R_t, \ IVOL_t) = Cov[E(R_t) + \varepsilon_t, \ E(IVOL_t) + v_t]$$
$$= Cov[E(R_t), E(IVOL_t)] + Cov[E(R_t), v_t] + Cov[\varepsilon_t, E(IVOL_t)] + Cov[\varepsilon_t, v_t]$$

The first component is of our interest. The second and third components are zero by the definition of shocks (i.e., unrelated to the information set at t-1). The last component, the covariance between these two contemporaneous shocks, is however unknown.

feasible investment strategy. If individual stocks with high E(IVOL) have higher returns than stocks with low E(IVOL), a zero-investment portfolio that is long in high E(IVOL) stocks and short in low E(IVOL) stocks should earn a positive return.

The procedure of the portfolio-based approach is as follows. In each month, I sort E(IVOL) to form ten portfolios with equal number of stocks. The first portfolio contains the 10% of stocks that are expected to have the lowest idiosyncratic volatilities in the next month and the last portfolio consists of the 10% of stocks that are expected to have the highest idiosyncratic volatilities. Table 6 presents the descriptive statistics for these ten E(IVOL)-portfolios. The mean E(IVOL) increases from 3.19% for the first portfolio to 36.35% for the last portfolio. The ex-post measured idiosyncratic volatility (*IVOL*) also increases monotonically across these ten portfolios. The portfolio consisting of stocks with high E(IVOL) have higher returns than the portfolio consisting of low-E(IVOL) stocks. The value-weighted portfolio returns increase monotonically from 0.90% for the lowest-E(IVOL) portfolio to 2.65% for the highest-E(IVOL) portfolio. The average monthly return for the zero-investment portfolio is 1.75%. The equal-weighted portfolio returns display a similar pattern and the return spread between the highest and lowest E(IVOL) portfolios is even larger. This evidence confirms the positive relation between E(IVOL) and individual stock returns.

Next I run the time-series regressions of the value-weighted excess returns on the Fama-French three-factors for each portfolio. The last row of Table 6 reports the regression intercepts. The alpha is 0.03% for the lowest-E(IVOL) portfolio and 1.45% for the highest-E(IVOL) portfolio. A hedging portfolio longing portfolio 10 and shorting portfolio 1 yields a statistically significant monthly return of 1.42%. The Gibbons-Ross-Shanken (GRS, 1989) test statistic has a value of 5.92 and thus strongly rejects the null

23

hypothesis that all the intercepts jointly equal zero. This result contrasts sharply with the findings of AHXZ which are based on the lagged realized volatility and further confirms that firms with high expected idiosyncratic volatility have higher expected returns.

3.5. Robustness check

I emphasize the importance to have a quality estimate of E(IVOL) in estimating the relation between idiosyncratic risk and return and use EGARCH models to reach the goal. But theoretically speaking, this goal can also be reached by examining the empirical relation between idiosyncratic volatility and a quality estimate of expected returns. So far I use realized returns as a proxy for expected returns. Though realized returns are an unbiased estimate of expected returns, the estimate quality is rather poor (Elton, 1999). Elton therefore encourages "developing better measures of expected return and alternative ways of testing asset pricing theories that do not require using realized returns" (p.1200). One alternative measure of expected returns is the implied cost of capital (ICC), which is essentially the firm's internal rate of return that equates the present value of future dividends to the current stock price. This measure is increasingly used in the accounting and finance literature. For example, Pastor, Sinha, and Swaminathan (2007) show that the ICC estimated on the basis of earnings forecasts could be a useful proxy for expected stock returns. Based on this estimate, they find a positive intertemporal risk and return tradeoff at the market level.

Following Pastor et al., I estimate the ICC as a proxy for the expected return of individual stocks. Table 7 reports the simple correlations between the estimated ICC and idiosyncratic volatility (in Panel A) and the estimates from the Fama-MacBeth regressions by using ICC as the dependent variable (in Panel B). The results are

24

consistent with my earlier findings. Both the correlation coefficients and the regressions slopes of idiosyncratic volatility are positive and statistically significant. In addition, I also tried the expected returns estimated from the Value Line forecasts. These estimates are downloaded from Alon Brav's website.[14] Brav, Lehavy, and Michaely (2005) use them to test asset pricing models. I again find a positive relation between this Value Line forecasted expected return and idiosyncratic volatility.

I check the results robustness in many other ways. First, the dependent variable in the Fama-MacBeth regressions is the simple raw return. Simple returns are known to be skewed positively. In addition they have the lower bound of -100% but no upper bound. This makes the regression results potentially driven by outliers. Besides having imposed the upper bound of returns to be 300% for the reported results, I test the robustness by replacing simple returns by log returns (replace R_t by $\ln(1+R_t)$), which do not have a lower bound and are not skewed, and run the same Fama-MacBeth regressions. The cross-sectional relations between log returns and idiosyncratic volatility remain positive and statistically significant, though become weaker in magnitude. For example, controlling for the variables as in Model 5 of Table 5, the coefficient estimate is 0.09 with a *t*-statistic of 9.94 for E(IVOL) and 0.17 with a *t*-statistic of 13.15 for IVOL.

Second, I estimate the test-statistics in the month-by-month cross-sectional regressions by using the Generalized Least Squares (GLS) method suggested by Litzenberger and Ramaswamy (1979). Specifically, the GLS estimator, $\hat{\gamma}_k$, is the

---

weighted mean of the monthly estimates, where the weights are inversely proportional to the variances of the monthly estimates.[15]

$$\hat{\gamma}_k = \sum_{t=1}^{T} Z_{kt} \hat{\gamma}_{kt} \, , \qquad (9)$$

$$Var(\hat{\gamma}_k) = \sum_{t=1}^{T} Z_{kt} Var(\hat{\gamma}_{kt}) \, , \qquad (10)$$

$$Z_{kt} = [Var(\hat{\gamma}_{kt})]^{-1} \Big/ \sum_{t=1}^{T} [Var(\hat{\gamma}_{kt})]^{-1} \, . \qquad (11)$$

The results remain qualitatively intact.

Last but not least, my choice of EGARCH models is unlikely to be the exclusive factor that drives the results. The whole purpose to use EGARCH models is to improve the estimation of conditional idiosyncratic volatility. Any other models that serve this purpose would deliver same results of this paper and different results from the naïve model that uses the lagged IVOL as a proxy for E(IVOL). In fact, two contemporaneous studies that use different methods to estimate conditional idiosyncratic volatilities obtain similar results to my study.[16]

4. The relation between return and the lagged idiosyncratic volatility

---

[15] The Fama-MacBeth regression implicitly assumes equal weights across months. It can be regarded as a specific example of GLS in which the variances of monthly estimates are the same.

[16] Chua, Goh, and Zhang (2007) use an AR(2) model, and Diavatopoulos, Doran, and Peterson (2007) decompose implied volatility from option prices to estimate conditional idiosyncratic volatility. Both studies conclude the positive tradeoff between idiosyncratic risk and expected return.

I have shown that idiosyncratic volatilities of individual stocks change over time and the lagged IVOL is not an appropriate proxy for the expected IVOL. As a result, the negative relation between return and the one-month lagged IVOL documented by AHXZ (2006) should not be used to draw inference on the relation between idiosyncratic risk and expected return. Employing the EGARCH models to estimate the expected IVOL, I find a significantly positive relation between expected IVOL and expected return. However, AHXZ's findings of the negative relation are still puzzling, though Bali and Cakici (2008) suggest that their results are sensitive to the research methods. In this section I replicate AHXZ's results by strictly following their methods and then offer an empirical explanation.

My evidence suggests that AHXZ's findings are largely driven by the return reversal of stocks that have high idiosyncratic volatilities. High idiosyncratic volatilities are contemporaneous with high returns, which tend to reverse in the following month. As a result, the returns of high-IVOL stocks are abnormally low in the next month. In addition, the stocks that drive their results are small in size and even in aggregate have a negligible weight relative to the total market capitalization. The evidence casts doubt on the effectiveness of trading strategies suggested by AHXZ (2006, 2008), if transaction costs are seriously accounted for.

I start with replicating the main results in Table VI of AHXZ (2006). Their sample period is from July 1963 to December 2000. I extend it by six years to December 2006. My results, reported in Table 8, are very similar to theirs. The quintile portfolio of stocks with the highest IVOL has a Jensen's alpha of 1.22% lower than the portfolio of stocks with the lowest IVOL (1.19% in AHXZ (2006)). The procedure to get the Jensen's alphas is as follows. In each month, I divide the universe of stocks into quintiles on the basis of

their IVOL. Portfolio 1 (5) is the portfolio of the 20% of stocks with the lowest (highest) IVOL. I compute the value-weighted excess return (*VWXRET*) in the next month for each portfolio. The weight for each stock is its market capitalization in the previous month. As a result, I have a time-series of value-weighted excess returns for each portfolio. I then run a time-series regression of the *VWXRET* on the Fama-French three factors. The purpose is to estimate the intercept alpha -- the average excess return not explained by these three factors.

Three findings are worth mention. First, only two out of the five alphas are statistically significant. They are the two alphas for Portfolio 4 and 5 whose stocks have relatively high lagged IVOL (Hereafter Portfolio 4 and 5 are also called the high-IVOL portfolios). So precisely speaking, the two portfolios of stocks with high lagged idiosyncratic volatilities realize negative abnormal returns but the other three portfolios of stocks with relatively low idiosyncratic volatilities do not realize significant abnormal returns. Second, these 40% of stocks in the high-IVOL portfolios tend to be small firms and their total market capitalization is only 9% of the whole market. Third, if we read the return numbers literally, the patterns for RET(t), VWXRET(t), and even FF-3F alphas – the metric that AHXZ's conclusion is based on, are not monotonically increasing or decreasing across the IVOL portfolios. Therefore, AHXZ's findings are completely driven by these small stocks with high idiosyncratic volatilities. The question then becomes why these stocks earn low returns in the subsequent month.

In the last three columns of Table 8, I present the mean raw return (*RET(t-1)*) and value-weighted excess return (*VWXRET(t-1)*) that are contemporaneous to the IVOL. I find that these returns are monotonically increasing in the IVOL portfolios. Moreover, the alphas from the time-series regressions of *VWXRET(t-1)* on the Fama-French three

factors are significantly positive for Portfolio 4 and 5 and not different from zero for the other three portfolios. The positive abnormal returns in month *t-1* and the negative abnormal returns in month *t* for the high-IVOL portfolios are not likely coincidental. The negative abnormal returns in month *t* are, at least partly, caused by the reversal of the positive abnormal returns in month *t-1*.

Next I focus on the 40% "trouble-making" firms (i.e., stocks in Portfolio 4 and 5) and examine the impact of return reversal. I divide these firms into quintiles based on RET(*t-1*). Table 9 shows the return dispersions of these five portfolios. The mean RET(*t-1*) increases from -22.67% for the lowest RET(*t-1*) portfolio to 33.78% for the highest RET(*t-1*) portfolio. Interestingly, the mean raw return in the next month, RET(*t*), decreases from 3.35% to -0.21% monotonically across these five RET(*t-1*) portfolios. The portfolio excess return, both the equal-weighted and the value-weighted, show the same pattern. The alphas estimated from the time-series regressions confirm that the negative abnormal returns in month *t* concentrate in firms that have relatively high past returns (RET(*t-1*)). The evidence suggests that some stocks with high IVOL at month *t-1* earn positive abnormal returns in the same month and due to the return reversal, realize negative abnormal returns at month *t*.

The negative correlation of subsequent monthly returns has been documented in the literature for long time, e.g., Jegadeesh (1990). A recent study by Huang, Liu, Rhee, and Zhang (2007) also suggests that the return reversal in monthly returns explains the difference in results between AHXZ (2006) and Bali and Cakici (2008). Controlling for the difference in the past-month returns, there is no negative relation between average returns and the lagged idiosyncratic volatilities.

5. Conclusion

For various reasons investors in reality often do not hold perfectly-diversified portfolios. Theories assuming under-diversification of investor portfolios, such as Levy (1978) and Merton (1987), predict a positive relation between idiosyncratic risk and expected return. Ang, Hodrick, Xing, and Zhang (2006, 2008) however find that monthly stock returns are negatively related to the one-month lagged idiosyncratic volatilities. I show that idiosyncratic volatilities are time-varying and that the one-month lagged value is not a good proxy for the expected value. So AHXZ's findings should not be used to imply the relation between expected return and idiosyncratic risk. I use EGARCH models to estimate the expected idiosyncratic volatilities and find they are positively related to expected returns. The positive relation is both economically and statistically significant and also robust to different testing methods. This evidence supports the theories assuming under-diversification. Stocks that are expected to have high idiosyncratic risk earn high returns in the cross-section. I further show that AHXZ's findings are driven by a subset of small firms with high idiosyncratic volatilities. These firms have high returns in the month of high idiosyncratic volatility. The high returns reverse in the subsequent month and result in the findings of negative abnormal returns.

References

Amihud, Y., 2002. Illiquidity and stock returns: cross-section and time-series effects. Journal of Financial Markets 5, 31-56.

Amihud, Y., Mendelson H., 1986. Asset pricing and the bid-ask spread. Journal of Financial Economics 15, 223-249.

Ang, A., Hodrick, R., Xing Y., Zhang X., 2006. The cross-section of volatility and expected returns. Journal of Finance 61, 259-299.

Ang, A., Hodrick, R., Xing Y., Zhang X., 2008. High idiosyncratic volatility and low returns: International and further U.S. evidence. Journal of Financial Economics, forthcoming.

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

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31, 307-328.

Bollerslev, T., Engle, R., Wooldridge, J., 1988. A capital asset pricing model with time-varying covariances. Journal of Political Economy 96, 116-131.

Boyer, B., Mitton T., Vorkink, K., 2007. Idiosyncratic volatility and skewness: time-series relations and the cross-section of expected returns. Unpublished working paper. Brigham Young University.

Brav, A., Lehavy R., Michaely, R., 2005. Using expectations to test asset pricing models. Financial Management 34, 31-64.

Brennan, M., Subrahmanyam, A., 1996. Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441-464.

Brockman, P., Schutte, M., 2007. Is idiosyncratic volatility priced? The international evidence. Unpublished working paper. University of Missouri – Columbia.

Campbell, J., Lettau, M., Malkiel, B., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance 56, 1-43.

Chordia, T., Subrahmanyam, A., Anshuman, V., 2001. Trading activity and expected stock returns. Journal of Financial Economics 59, 3-32.

Chua, C., Goh, J., Zhang, Z., 2007. Idiosyncratic volatility matters for the cross-section of returns – in more ways than one! Unpublished working paper. Singapore Management University, Singapore.

Datar, V., Naik, N., Radcliffe, R., 1998. Liquidity and stock returns: an alternative test. Journal of Financial Markets 1, 203-219.

Diavatopoulos, D., Doran, J., Peterson, D., 2007, The information content in implied idiosyncratic volatility and the cross-section of stock returns: evidence from the option markets. Unpublished working paper, Florida State University.

Easley, D., Hvidkjaer, S., O'Hara, M., 2002. Is information risk a determinant of asset returns? Journal of Finance 57, 2185-2221.

Eiling, E., 2006. Can nontradable assets explain the apparent premium for idiosyncratic risk? The case of industry-specific human capital. Unpublished working paper. Tilburg University, Netherlands.

Engel, R., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50, 987-1007.

Engel, R., Mustafa, C. 1992. Implied ARCH models from options prices. Journal of Econometrics 52, 289-311.

Engel, R., Ng, V., 1993. Time varying volatility and the dynamic behavior of the term structure. Journal of Money, Credit and Banking 25, 336-349.

Elton, E., 1999. Expected return, realized return, and asset pricing tests. Journal of Finance 54, 1199-1220.

Fama, E., MacBeth, J., 1973. Risk, return and equilibrium: empirical tests. Journal of Political Economy 81, 607-636.

Fama, E., French, K., 1992. The cross-section of expected stock returns. Journal of Finance 48, 427-465.

Fama, E., French, K., 1993. Common risk factors in the returns on stocks and bonds. Journal of Financial Economics 33, 3-56.

Fama, E., French, K., 1996. Multifactor explanations of asset pricing anomalies. Journal of Finance 52, 55-84.

French, K., Schwert, G., Stambaugh, R., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3-29.

Fuller, W., 1996. Introduction to Statistical Time Series, John Wiley & Sons.

Gibbons, M., Ross, S., Shanken, J., 1989. A test of the efficiency of a given portfolio. Econometrica 57, 1121-1152.

Goetzmann, W., Kumar, A., 2004. Why do individual investors hold under-diversified portfolios? Unpublished working paper. Yale University.

Harvey, C., Siddique, A., 2000. Conditional skewness in asset pricing tests. Journal of Finance 55, 1263-1295.

Huang, W., Liu, Q., Rhee, G., Zhang, L., 2007. Another look at idiosyncratic risk and expected returns, Unpublished working paper, University of Hawaii at Manoa.

Jegadeesh, N., 1990. Evidence of predictable behavior of security returns. Journal of Finance 45, 881-898.

Jegadeesh, N., Titman, S., 1993. Returns to buying winners and selling losers: implications for stock market efficiency. Journal of Finance 48, 65-92.

Jiang G., Xu, D., Yao, T., 2006. The information content of idiosyncratic volatility. Unpublished working paper. University of Arizona.

Levy, H., 1978. Equilibrium in an imperfect market: a constraint on the number of securities in the portfolio. American Economic Review 68, 643-658.

Litzenberger, R., Ramaswamy, K., 1979. The effect of personal taxes and dividends on capital asset prices: theory and empirical evidence. Journal of Financial Economics 7, 163-195.

Malkiel, B., Xu, Y., 2002. Idiosyncratic risk and security returns. Unpublished working paper. University of Texas at Dallas.

Merton, R., 1987. A simple model of capital market equilibrium with incomplete information. Journal of Finance 42, 483-510.

Nelson, D., 1991. Conditional heteroskedasticity in asset returns: a new approach. Econometrica 59, 347-370.

Pagan, A., Schwert, G., 1990. Alternative models for conditional stock volatility. Journal of Econometrics 45, 267-290.

Pastor, L., Sinha, M., Swaminathan, B., 2007. Estimating the intertemporal risk-return tradeoff using the implied cost of capital. Journal of Finance, forthcoming.

Pastor, L., Stambaugh, R., 2003. Liquidity risk and expected stock returns. Journal of Political Economy 111, 642-685.

Pontiff, J., 2006. Costly arbitrage and the myth of idiosyncratic risk. Journal of Accounting and Economics 42, 35-52.

Schwert, G., 1989. Why does stock market volatility change over time? Journal of Finance 44, 1115-1153.

Spiegel, M., Wang, X., 2006. Cross-sectional variation in stock returns: liquidity and idiosyncratic risk. Unpublished working paper. Yale University.

Table 1
Time-series properties of idiosyncratic volatility

This table summarizes the time-series statistics of individual stock idiosyncratic volatilities. I compute first the time-series statistics of idiosyncratic volatility for each stock and then the mean statistics across all stocks. Stocks are traded in the NYSE, the AMEX, or the Nasdaq during July 1963 to December 2006. The idiosyncratic volatility (*IVOL*) is estimated as follows. In every month, excess daily returns of each individual stock are regressed on the daily Fama-French three factors: *RmRf, SMB,* and *HML.* The (monthly) idiosyncratic volatility of the stock is the product of the standard deviation of the regression residuals and the square root of the number of observations in the month.

	N	Mean	S.D.	C.V.	Skew	Autocorrelation at lags
						1	2	3	4	5	6	11	12	13
IVOL	26189	16.87	9.94	0.55	1.65	0.33	0.27	0.24	0.20	0.19	0.18	0.12	0.14	0.11
$Ln\left(\dfrac{IVOL_t}{IVOL_{t-1}}\right)$	26068	-0.004	0.54	366.24	-0.03	-0.42	-0.04	0.01	-0.02	-0.01	0.01	-0.02	0.03	-0.02

Table 2
Do monthly idiosyncratic volatilities follow a random walk process?

This table presents statistics of the estimations from the time-series regressions in which the changes in idiosyncratic volatility of an individual stock are regressed on the level of idiosyncratic volatility in the past month. The regression is intended to examine whether the time-series idiosyncratic volatilities of this individual stock follows a random walk. The reported statistics are the cross-sectional mean, median, the lower and the upper quartiles of the coefficient estimate γ_1 and its associated t-statistics. The t-statistics are compared with the Dickey-Fuller critical values* to examine whether the null hypothesis of a random walk is rejected. The last column reports the percentage of firms for which the random walk hypothesis is rejected at the 1% level. For the regression, I require firms to have at least 30 months of observations ($T_i \geq 30 \ \forall i$). The sample period is July 1963 to December 2006.

Variables	N	Mean	Median	Q1	Q3	RW rejected
Model: $IVOL_{i,t+1} - IVOL_{i,t} = \gamma_{0i} + \gamma_{1i} IVOL_{i,t} + \eta_i$, $i = 1, 2, \cdots, N$, $t = 1,2,\cdots,T_i$						
γ_1	20979	-0.61	-0.60	-0.76	-0.45	
$t(\gamma_1)$	20979	-6.81	-6.40	-8.43	-4.85	89.97%
Model: $LnIVOL_{i,t+1} - LnIVOL_{i,t} = \gamma_{0i} + \gamma_{1i} LnIVOL_{i,t} + \eta_i$, $i = 1, 2, \cdots, N$, $t = 1,2,\cdots,T_i$						
γ_1	20979	-0.56	-0.55	-0.70	-0.41	
$t(\gamma_1)$	20979	-6.38	-5.99	-7.86	-4.51	87.81%

*Dickey-Fuller critical t-statistics (from Fuller (1996))

Sample size	Critical t-statistics (1%)
25	-3.75
50	-3.59
100	-3.50
250	-3.45
500	-3.44

Table 3
Variable descriptive statistics for the pooled sample: July 1963 to December 2006

This table reports the pooled descriptive statistics of stocks that are traded in the NYSE, AMEX, or Nasdaq during July 1963 to December 2006. *RET* is the monthly raw return reported in percentage. *XRET* stands for the monthly excess return, which is the raw return net of the one-month T-bill rate. *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 monthly closing price and the number of outstanding shares in June. Book-to-market equity (*BE/ME*) is the fiscal-yearend book value of common equity divided by the calendar-yearend market value of equity. The idiosyncratic volatility (*IVOL*) is estimated as follows. In every month, excess daily returns of each individual stock are regressed on the Fama-French three factors: *RmRf, SMB,* and *HML*. The (monthly) idiosyncratic volatility of the stock is the product of the standard deviation of the regression residuals and the square root of the number of observations in the month. *E(IVOL)* is the one-month-ahead expected idiosyncratic volatility estimated by EGARCH models. *RET(-2, -7)* is the compound gross return from month *t*-7 to *t*-2. *TURN* is the average turnover and *CVTURN* is the coefficient of variation of turnovers in the past 36 months. Variables with skewness greater than 3.00 are reported as the natural logarithm. To avoid giving extreme observations heavy weight in the return regressions and potential data recording errors, the smallest and largest 0.5% of the observations in each month for *ME, BE/ME, IVOL, E(IVOL), Ret(-2, -7), TURN,* and *CVTURN* are set equal to the next smallest or largest values. Observations with monthly returns greater than 300% are deleted.

Variables	Mean	Std Dev	Median	Q1	Q3	Skew	N
RET (%)	1.18	16.86	0.15	-6.52	6.78	2.35	2,947,826
XRET(%)	0.71	16.87	-0.27	-6.99	6.34	2.35	2,947,826
Ln(1+RET) (%)	-0.12	16.09	0.15	-6.74	6.56	-0.54	2,947,826
IVOL	14.17	13.91	10.41	6.33	17.41	6.94	2,946,521
E(IVOL)	12.67	10.91	10.29	6.46	15.18	2.38	2,867,821
BETA	1.22	0.36	1.17	0.94	1.46	0.31	1,721,356
Ln(ME)	4.29	2.03	4.16	2.82	5.64	0.34	2,804,878
Ln(BE/ME)	-0.39	1.09	-0.35	-0.97	0.20	0.27	2,145,253
RET(-2, -7)	1.07	0.39	1.03	0.85	1.22	2.90	2,758,743
Ln(TURN) (%)	1.39	1.09	1.39	0.67	2.13	-0.08	2,041,658
Ln(CVTURN)	4.15	0.44	4.14	3.92	4.48	0.14	2,038,647

Table 4
Cross-sectional simple correlations

This table presents the time series means of the cross-sectional Pearson correlations. The variables relate to a sample of stocks traded in the NYSE, the AMEX, or the Nasdaq during July 1963 to December 2006. Variables are defined in Table 3. The correlation coefficients followed by * are significant at the 1% level based on their time-series standard error.

	Ln(1+RET)	IVOL	E(IVOL)	BETA	Ln(ME)	Ln(BE/ME)	RET(-2,-7)	Ln(TURN)	Ln(CVTURN)
RET	0.98*	0.14*	0.09*	-0.01	-0.01*	0.03*	0.02*	-0.02*	-0.00
Ln(1+RET)		0.05*	0.03*	-0.03*	0.02*	0.04*	0.04*	-0.03*	-0.02*
IVOL			0.46*	0.34*	-0.39*	-0.05*	-0.12*	0.16*	0.31*
E(IVOL)				0.35*	-0.34*	-0.11*	-0.04*	0.20*	0.30*
BETA					-0.34*	-0.04*	-0.03*	0.41*	0.23*
Ln(ME)						-0.21*	-0.03*	0.04*	-0.57*
Ln(BE/ME)							0.06*	-0.12*	0.16*
RET(-2, -7)								0.00	0.06*
Ln(TURN)									0.02*
Ln(CVTURN)									

Table 5
Fama-MacBeth regressions of stock returns on idiosyncratic volatility and firm characteristics

The table presents the time-series averages of the slopes in cross-sectional regressions using the standard Fama and MacBeth (1973) methodology. The t-statistic is the average slope divided by its time-series standard error. The sample period is July 1963 to December 2006. The dependent variable (RET_t) is the percentage monthly return. $E(IVOL_t)$ is the one-month-ahead expected idiosyncratic volatility estimated by an exponential GARCH model. $IVOL_{t-1}$ is the one-month lagged idiosyncratic volatility. BETA, ME, and BE/ME are estimated as in Fama and French (1992). TURN is the average turnover and CVTURN is the coefficient of variation of turnovers in the past 36 months. RET(-2, -7) is the compound gross return from month t-7 to t-2. To avoid giving extreme observations heavy weight in the regressions, the smallest and largest 0.5% of the explanatory variables (except BETA) are set equal to the next smallest and largest values. This has no effect on inferences. The last column reports the average R-squares of the cross-sectional regressions.

MODEL	BETA	Ln(ME)	Ln(BE/ME)	Ret(-2, -7)	Ln(TURN)	Ln(CVTURN)	$E(IVOL_t)$	$IVOL_{t-1}$	$IVOL_t$	$\overline{R^2}$ (%)
1	0.02	-0.12	0.23							3.82
	(0.08)	(-3.11)	(4.97)							
2	0.14	-0.17	0.19	0.64	-0.12	-0.44				5.73
	(0.93)	(-4.52)	(4.38)	(3.09)	(-2.05)	(-6.79)				
3							0.11			3.02
							(9.05)			
4		0.25	0.60				0.13			4.98
		(7.28)	(12.58)				(11.41)			
5		0.19	0.48	0.93	-0.48	-0.73	0.15			6.89
		(5.01)	(10.70)	(4.74)	(-7.34)	(-11.82)	(13.65)			
6		-0.21	0.18	0.67	-0.09	-0.39		-0.02		5.56
		(-5.76)	(4.04)	(3.36)	(-1.24)	(-6.48)		(-3.73)		
7		0.41	0.44	1.61	-0.55	-0.83			0.31	10.42
		(14.53)	(10.57)	(8.55)	(-8.54)	(-13.59)			(20.56)	

Table 6
Summary statistics for portfolios formed on conditional idiosyncratic volatility

Each month five portfolios are formed on *E(IVOL)*, the one-month-ahead expected idiosyncratic volatilities of individual stocks. *E(IVOL)* is estimated by an exponential GARCH model. The first portfolio (Low) consists of the 10% of stocks with the lowest *E(IVOL)* and the last portfolio (High) consists of the 10% of stocks with the highest *E(IVOL)*. Portfolios are updated monthly. The first two rows present, respectively, the time-series means of the value-weighted and the equal-weighted portfolio returns. Other rows show the pooled means/medians of variables within the particular portfolio. The medians instead of the means of *ME* and *BE/ME* are reported due to their substantial skewness. The last row reports the alphas (intercepts) from the time-series regressions of the value-weighted portfolio excess returns on the Fama-French three factors. The sample period is from July 1963 to December 2006.

Variables	Portfolios formed on *E(IVOL)*									
	Low	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**High**
Port. VWRET	0.90	0.96	0.97	0.98	1.00	1.02	1.17	1.18	1.28	2.65
Port. EWRET	0.54	0.77	0.79	0.80	0.78	0.82	0.85	0.91	1.41	5.33
E(IVOL)	3.19	5.17	6.52	7.80	9.19	10.78	12.73	15.34	19.58	36.35
IVOL	6.74	7.80	8.98	10.29	11.80	13.50	15.46	17.72	20.81	27.29
BETA	0.90	1.00	1.08	1.16	1.23	1.29	1.36	1.40	1.44	1.46
ME ($mil, med)	113.03	177.16	161.38	119.04	85.80	63.04	45.68	33.83	23.72	14.19
BE/ME (med)	0.90	0.78	0.75	0.74	0.73	0.71	0.68	0.64	0.59	0.52
FF Alphas	0.03	0.01	-0.02	-0.02	-0.05	-0.06	0.04	0.01	0.13	1.45

Table 7
Cross-sectional relations between the implied cost of capital and idiosyncratic volatility

Panel A of the table presents the simple correlations between the realized return, the estimated implied cost of capital (ICC), the realized idiosyncratic volatility, and the EGARCH-estimated expected idiosyncratic volatility. The correlation coefficients followed by * are significant at the 1% level based on their time-series standard error. Panel B reports the coefficient estimates and *t*-statistics (in parentheses) from the Fama-MacBeth regressions where the dependent variables are the estimated ICC. The estimation of ICC follows Pastor, Sinha, and Swaminathan (2007) and is based on analyst earnings forecasts. Due to the availability of analyst earnings forecasts, the sample period is from January 1980 to December 2006.

Panel A: Simple correlations

	ICC	IVOL	E(IVOL)
RET	0.06*	0.14*	0.07*
ICC		0.20*	0.17*
IVOL			0.54*
E(IVOL)			

Panel B: Fama-MacBeth regressions (Dependent variable: ICC)

MODEL	$Ln(ME)$	$Ln(BE/ME)$	$Ret(-2, -7)$	$Ln(TURN)$	$Ln(CVTURN)$	$E(IVOL_t)$	$IVOL_t$	$\overline{R^2}$ (%)
1	-0.05	0.03	-0.18	-0.01	-0.05	0.08		10.09
	(-31.94)	(6.52)	(-24.62)	(-1.60)	(-7.81)	(10.26)		
2	-0.06	0.03	-0.17	-0.02	-0.06		0.09	11.18
	(-31.42)	(6.14)	(-23.94)	(-3.92)	(-7.52)		(11.45)	

Table 8
Return dispersion of portfolios sorted by idiosyncratic volatility

This table illustrates differences in monthly percentage returns of portfolios sorted by idiosyncratic volatilities. In each month, I divide the universe of stocks into quintiles on the basis of their idiosyncratic volatility (*IVOL*). Portfolio 1 (5) is the portfolio of stocks with the lowest (highest) IVOL. The idiosyncratic volatility is estimated as follows. In every month, excess daily returns of each individual stock are regressed on the Fama-French three factors: *RmRf, SMB,* and *HML*. The (monthly) idiosyncratic volatility of the stock is the product of the standard deviation of the regression residuals and the square root of the number of observations in the month. The month that I form portfolios is indicated as *t-1* and the subsequent month *t*. N is the number of firm-month observations for the pooled sample. The numbers presented in other columns are means with *t*-statistics in brackets, if any. ME stands for market capitalization. RET is the raw return. VWXRET are the value-weighted excess returns for the portfolio, which are used to compute the FF-3F alphas in the time-series regressions. The sample period is July 1963 to December 2006.

IVOL Portfolio	N	IVOL (*t-1*)	ME(*t-1*) ($ mil)	MKT Share (%)	RET (*t*)	VWXRET (*t*)	FF-3F Alpha (t)	RET (t-1)	VWXRET (t-1)	FF-3F Alpha (t-1)
1 (Low)	574915	4.30	1885.04	43.00	1.10	0.52	0.074 (1.75)	0.44	0.39	-0.04 (-0.94)
2	574293	7.58	1451.81	33.08	1.34	0.57	0.034 (0.76)	0.55	0.63	0.05 (1.07)
3	574694	11.06	653.93	14.91	1.37	0.64	0.058 (0.83)	0.61	0.76	-0.02 (-0.13)
4	574707	16.17	294.70	6.72	1.19	0.29	-0.353 (-3.60)	0.77	0.79	0.26 (2.65)
5 (High)	574915	32.32	100.44	2.29	1.08	-0.40	-1.146 (-7.00)	4.11	1.66	0.85 (3.14)
5-1							-1.220 (-6.45)			0.89 (3.26)

Table 9
Return dispersion of high-IVOL stocks sorted by the one-month lagged return

This table examines the impact of return reversal on the high-IVOL stocks. At month *t-1*, I identify 40% of stocks that have the highest idiosyncratic volatilities and divide them into quintiles on the basis of their contemporaneous returns. Portfolio 1 (5) is the portfolio of stocks with the lowest (highest) RET(*t-1*). RET is the raw monthly percentage return. VWXRET (EWXRET) are the time-series mean value-weighted (equal-weighted) excess returns ($R_{pt} - r_{ft}$) for the portfolio, which are used to compute the FF-3F alphas in the time-series regressions. The sample period is July 1963 to December 2006.

Portfolio Sorted by RET(t-1)	N	RET(*t-1*)	RET(*t*)	EWXRET(*t*)	VWXRET(*t*)	IVOL(*t*)	ME(*t-1*) ($ mil)	FF-3F Alpha (EWXRET(*t*))	FF-3 Alpha (VWXRET(*t*))
1 (Low)	232405	-22.67	3.35	2.84	0.56	28.10	155.23	1.77 (6.55)	-0.33 (-1.52)
2	223492	-8.39	1.17	0.93	0.35	21.63	193.63	-0.07 (-0.46)	-0.41 (-2.67)
3	228808	0.00	0.90	0.67	0.09	20.32	193.69	-0.48 (-3.62)	-0.57 (-4.54)
4	233511	9.15	0.45	0.02	-0.06	19.35	228.15	-0.83 (-7.23)	-0.70 (-6.02)
5 (High)	231406	33.78	-0.21	-0.62	-0.10	21.15	216.79	-1.40 (-9.49)	-0.69 (-5.66)