Asia-Pacific Journal of Business (아태비즈니스연구)
Vol. 15, No. 1, March 2024 (pp.1-14)
https://doi.org/10.32599/apjb.15.1.202403.1

APJB
ISSN 2233-5900 (Print)
ISSN 2384-3934 (Online)

Is The Idiosyncratic Volatility Puzzle Driven By A Missing Factor?*

Hanjun Kim[a], Bumjean Sohn[b]

[a]David Eccles School of Business, University of Utah, U.S.A.
[b]Korea University Business School, Korea University, South Korea

Received 29 February 2024, Revised 20 March 2024, Accepted 24 March 2024

Abstract

Purpose - We investigate whether a potential missing pricing factor plays a significant role in the idiosyncratic volatility puzzle.
Design/methodology/approach - We theoretically show how a missing pricing factor can affect the idiosyncratic volatility puzzle, and also show how to get around the problem empirically. We adopt the Fama-French five factor model for the estimation of the idiosyncratic risk and use randomly constructed portfolios as test assets.
Findings - We find that a missing factor does not drive the idiosyncratic volatility puzzle. Thus, we conclude that the idiosyncratic volatility does affect the risk premium of its stock.
Research implications or Originality - The Fama-French five factor model does a pretty good job in explaining the risk premiums of stocks, and it can be used to reliably estimate idiosyncratic risk of stocks.

Keywords: Cross-Section of Equity Returns, Idiosyncratic Risk, Idiosyncratic Volatility Puzzle, Missing Factor, Random Portfolio
JEL Classifications: G12, G14

Ⅰ. Introduction

In traditional asset pricing theories, investors are compensated for bearing only the systematic risk because the idiosyncratic risk can easily be diversified away. In other words, only the systematic risk should be priced, and the idiosyncratic risk should not. However, assumptions in the traditional theories, such as the perfect market and the complete information, usually don't reflect the real world. Investors might require a risk premium for bearing the idiosyncratic risk because investors in the real world can't easily diversify it. Merton (1987) argues that, if investors have incomplete information (*i.e.*, investors don't have the full information of available securities), they can't hold a well-diversified portfolio and the idiosyncratic risk can be positively priced. Similarly, Levy (1978) finds that, in an imperfect market, market frictions

* This work was supported by Korea University Business School Research Fund of 2023.
[a] First Author, E-mail: Hanjun.Kim@eccles.utah.edu
[b] Corresponding Author, E-mail: sohnb@korea.ac.kr

force investors to hold an undiversified portfolio and thus idiosyncratic risk can be positively priced.

Motivated by the above theories about the idiosyncratic risk, many papers try to find the empirical relationship between the idiosyncratic risk and stock returns. Ang et al. (2006) find that the idiosyncratic volatility has a negative relationship with stock returns. Because theories suggest that idiosyncratic risks should not be priced or be positively priced in an imperfect market, this result presents a puzzle for future research and is named the idiosyncratic volatility puzzle. Many papers try to explain the negative relationship presented in Ang et al. (2006). Fu (2009) and Huang et al. (2009) argue that return reversal is the main cause of the idiosyncratic volatility puzzle. They show that idiosyncratic risk is positively priced after controlling the return reversal. Han and Lesmond (2011) suggest liquidity as an explanation for the idiosyncratic volatility puzzle. They argue that the estimation of idiosyncratic volatility largely depends on the effect of liquidity cost, and, after controlling it, they find that idiosyncratic volatility has no significant relationship with the stock returns. Bali and Cakici (2008) indicate that methodological differences in the estimation of idiosyncratic volatility are critical in the presence and significance of the idiosyncratic volatility puzzle. They find that the existence and direction of the relationship between idiosyncratic volatility and stock returns are determined by data frequency used to estimate idiosyncratic volatility, weighting schemes, and criteria of screening the data. Stambaugh et al. (2015) argue that the idiosyncratic volatility puzzle results from the combination of arbitrage risk and arbitrage asymmetry. Although idiosyncratic risks deter the arbitrage opportunities (arbitrage risk) in both overpriced and underpriced stocks, underpriced stocks are more eliminated because buying is easier than selling (arbitrage asymmetry). Hence, they suggest that remaining overpriced stocks, which have a negative relationship with idiosyncratic volatility, drive the idiosyncratic volatility puzzle. Avramov et al. (2013) find that firms with low credit ratings are the main causes of the idiosyncratic volatility puzzle.

Although idiosyncratic risk is a well-known theoretical concept in the field of finance, idiosyncratic volatility doesn't have a precise definition for empirical analysis, and it is strongly model-dependent. Researchers should choose the appropriate model for measuring idiosyncratic volatility and it is conventional to define idiosyncratic volatility as standard deviation of residuals relative to the Fama-French three factor (henceforth, FF-3) model. However, because every asset pricing model, including the FF-3 model, is not a perfect model, there must be missing pricing factors, and therefore the estimate of idiosyncratic volatility is composed of not only the pure idiosyncratic risk but also the components from the missing factors. If the components from the missing factors have a significant relationship with stock returns, it might mislead the empirical analysis of previous papers. For example, the components from missing factors can be negatively related to stock returns and the results from Ang et al. (2006) might be largely driven by that relationship. Then, the idiosyncratic volatility puzzle is no longer a puzzle. Hence, one needs to check whether this problem with the missing factors has a significant role in the idiosyncratic volatility puzzle or not.

Therefore, in this paper, we carefully test the relationship between stock returns and the components from missing factors. First, we confirm that the idiosyncratic volatility puzzle presented in Ang et al. (2006) still exists in our sample period. We extend the sample period, and, for robustness, use the Fama-French five factor (henceforth, FF-5) model to measure the idiosyncratic volatility. Second, to eliminate the pure idiosyncratic risk in the test portfolios, we randomly construct the portfolios with 100 individual stocks. Since the pure idiosyncratic risk is diversified away in the portfolio level, the estimated idiosyncratic volatility of these portfolios should only contain the components from the missing factors. Then, we examine the relationship between the estimated idiosyncratic volatility and the returns of these well-diversified portfolios. We find that there is no significant relationship between idiosyncratic volatility and returns of randomly constructed portfolios. Thus, the results suggest that omitted factors are not a significant problem in the idiosyncratic volatility puzzle and that the idiosyncratic risk is indeed priced.

The paper is organized as follows. In Section II, we provide the theoretical motivation for the effect of a missing pricing factor in estimating the idiosyncratic volatility. In Section III, we empirically test the significance of the effect of the missing factor in the context of the idiosyncratic volatility puzzle. Section IV concludes.

II. Theoretical Motivation

We show that the portfolio sorting by the estimated idiosyncratic volatility is equivalent to the portfolio sorting by the beta or the absolute value of the beta to the missing pricing factor under certain conditions. For the sake of simplicity, a two-factor model is considered as follows.

$$r_{i,t}^{e} = a_i + \beta_{i,1} F_{1,t} + \beta_{i,2} F_{2,t} + \epsilon_{i,t} \tag{1}$$

where $r_{i,t}^{e}$ is the excess return of an individual stock i. Suppose that the second factor $F_{2,t}$ is unknown to researchers and that the idiosyncratic volatility is measured by the corresponding empirical model as below.

$$r_{i,t}^{e} = a_i + \beta_{i,1} F_{1,t} + e_{i,t} \tag{2}$$

Note that the error term in equation (2) is different from the error term in equation (1). More precisely, the error term in equation (2) is the sum of the last two terms in equation (1).

$$e_{i,t} = \beta_{i,2} F_{2,t} + \epsilon_{i,t} \tag{3}$$

With the assumption that factors ($F_{1,t}$ & $F_{2,t}$) and error term ($\epsilon_{i,t}$) are orthogonal, taking conditional variance of the both sides in equation (3) yields the following result.

$$Var_{t-1}(e_{i,t}) = \beta_{i,t}^{2} Var_{t-1}(F_{2,t}) + Var_{t-1}(\epsilon_{i,t}) \tag{4}$$

It is straightforward to see that, if the idiosyncratic volatility is calculated from the empirical model with the missing factor ($F_{2,t}$), the measured idiosyncratic risk is the contaminated by the term from the missing factor, $\beta_{i,2}^{2} Var_{t-1}(F_{2,t})$. Thus, if we sort individual stocks based on the value of $\sqrt{Var_{t-1}(e_{i,t})}$ (This is how Ang et al. (2006) sort their portfolios), the resulting portfolios might not represent their varying exposure to the idiosyncratic risk $Var_{t-1}(\epsilon_{i,t})$ because $\beta_{i,2}^2 Var_{t-1}(F_{2,t})$ might dominate. In the case where $Var_{t-1}(\epsilon_{i,t})$ is small enough compared to $\beta_{i,2}^2 Var_{t-1}(F_{2,t})$, we have the following:

$$\sqrt{Var_{t-1}(e_{i,t})} \approx |\beta_{i,2}| STD_{t-1}(F_{2,t})$$

Note that $STD_{t-1}(F_{2,t})$ is common across all assets. Thus, in this case, portfolio sorting by the idiosyncratic risk would be equivalent to that by the absolute value of the betas for the missing pricing factor. Furthermore, if all the assets have the same signed betas (all positive or negative), the sorting would be equivalent to that by the beta for the missing pricing factor.

Previous papers about the idiosyncratic volatility puzzle implicitly assume that the estimated idiosyncratic volatility correctly measures the pure idiosyncratic risk. However, it is possible that the term with missing factor dominates the value of estimated idiosyncratic volatility. Hence, it is important to test whether the effect of a missing factor has a significant role in the idiosyncratic volatility puzzle.

To examine the effect of a missing pricing factor, we need to separate the term with the missing factor and the pure idiosyncratic risk. We construct test portfolios by randomly selecting stocks to eliminate pure idiosyncratic risk from these portfolios. Then, the estimated idiosyncratic volatility of the well-diversified test portfolios would have only the component from the missing factor.

Consider a well-diversified portfolio P of the individual assets in equation (1).

$$r_{p,t}^e = a_p + \beta_{p,1} F_{1,t} + \beta_{p,2} F_{2,t} + \epsilon_{p,t} \tag{5}$$

where $r_{p,t}^e$ is the excess return of the portfolio P. Note that $\epsilon_{p,t}$ is zero if the portfolio is well-diversified. Suppose that second factor ($F_{2,t}$) is unknown like before and that the idiosyncratic volatility is estimated by just a one-factor model.

$$r_{p,t}^e = a_p + \beta_{p,1} F_{1,t} + e_{p,t} \tag{6}$$

Then, as in equation (3), the error term in equation (6) can be expressed as the sum of the last two terms in equation (5).

$$e_{p,t} = \beta_{p,2} F_{2,t} + \epsilon_{p,t} \tag{7}$$

The conditional variance of the error term yields the following.

$$Var_{t-1}(e_{p,t}) = \beta_{p,2}^2 Var_{t-1}(F_{2,t}) + Var_{t-1}(\epsilon_{p,t}) \tag{8}$$

However, if the portfolio is well-diversified, $\epsilon_{p,t}$ is zero and thus the pure idiosyncratic risk term in equation (8), $Var_{t-1}(\epsilon_{p,t})$, is also zero. Hence, equation (8) can be simplified as follows.

$$Var_{t-1}(e_{p,t}) = \beta_{p,2}^2 Var_{t-1}(F_{2,t}) \tag{9}$$

Now that the measured idiosyncratic risk is solely represented by the term from a missing factor, one can test whether this incorrectly measured idiosyncratic risk, which in fact represents

$|\beta_{p,2}|$, has some significant relationship with returns of portfolio or not. If the relationship between return and idiosyncratic volatility in this portfolio level is the same as the negative relationship in Ang et al. (2006), one might say that the effect of a missing factor dominates the idiosyncratic volatility puzzle. On the other hand, if there is no significant relationship in the portfolio level, then one might say that the effect of a missing factor has no important role in the idiosyncratic volatility puzzle.

Ⅲ. Empirical Analysis

1. Presence of Idiosyncratic Volatility Puzzle

In this section, we precisely follow the method of Ang et al. (2006) and check whether the idiosyncratic volatility puzzle still exists in our sample. Because we study the effects of a missing factor, it is important to include all factors that are known to be priced. Hence, we calculate the idiosyncratic volatility relative to the FF-5 model rather than the FF-3 model to capture the effects of (truly) missing factors. Idiosyncratic volatility is defined as standard deviation of residuals relative to the FF-5 model. For each month, by using each month's daily data, we calculate the idiosyncratic volatility for each firm.

$$\begin{aligned} r_{i,t}^{e} &= \alpha_i + \beta_{i,m} r_{m,t}^{e} + \beta_{i,smb} SMB_t + \beta_{i,hml} HML_t \\ &\quad + \beta_{i,\mathrm{rmw}} \mathrm{RMW}_t + \beta_{\mathrm{i,cma}} \mathrm{CMA}_t + \mathrm{e}_{\mathrm{i,t}} \end{aligned} \tag{10}$$

And, monthly idiosyncratic volatility is defined as $\sqrt{Var(e_{i,t})} \times \sqrt{trading\,days}$.

Table 1. Quintile Portfolios Sorted by Idiosyncratic Volatility

Quintile	RET (%)	Stdev (%)	SIZE ($mil)	CAPM_alpha (%)	FF3_alpha (%)	FF5_alpha (%)
1	0.93	3.57	24440.04	0.13	0.10	0.01
	(6.72)***			(2.94)***	(2.85)***	(0.36)
2	0.97	4.51	26840.60	0.04	0.02	-0.03
	(5.53)***			(1.13)	(0.55)	(-0.97)
3	1.04	5.52	14905.55	0.00	-0.01	0.08
	(4.84)***			(0.02)	(-0.09)	(1.51)
4	0.79	6.87	6870.67	-0.35	-0.35	-0.09
	(2.97)***			(-2.98)***	(-4.00)***	(-1.18)
5	0.25	8.27	2194.21	-0.96	-1.01	-0.56
	(0.77)			(-4.92)***	(-6.85)***	(-4.37)***
5-1	-0.69			-1.09	-1.11	-0.57
	(-2.66)***			(-4.80)***	(-6.62)***	(-4.00)***

To examine the relationship between return and (lagged) idiosyncratic volatility, for each month, we sort individual stocks into quintiles based on (lagged) idiosyncratic volatility[1). Then, we construct a value-weighted portfolio for each quintile. Portfolio 1 (5) is the portfolio with the lowest (highest) idiosyncratic volatility. Results are reported in Table 1 and they are very similar to Ang et al. (2006). Portfolios with higher idiosyncratic volatility have a smaller size and a larger standard deviation of returns. Most importantly, a portfolio with the highest idiosyncratic volatility has significantly lower mean return and alphas (CAPM, FF-3, and FF-5) than a portfolio with the lowest idiosyncratic volatility. Hence, the idiosyncratic risk is negatively priced and idiosyncratic volatility puzzle still exists in our sample.

However, as we discussed in Section II, the results in Table 1 might be distorted by the term with a missing factor. The component from the missing factor might drive the results in Table 1. If this is the case, Table 1 reveals that the missing factor, not (pure) idiosyncratic risk, is priced among the test portfolios. Thus, we carefully test whether the effect of a missing factor has a significant role in the context of the idiosyncratic volatility puzzle by constructing a well-diversified portfolio.

2. Randomly Constructed Portfolio and Idiosyncratic Volatility

We construct random portfolios to eliminate the pure idiosyncratic risk component in the estimated idiosyncratic volatility based on theoretical background in Section II. The reason why we do not use some conventional test portfolios such as 25 size-BM portfolios or 25 size-momentum portfolios is that such portfolios are sorted by specific risk characteristics and idiosyncratic risk might not be well-diversified away because of such risk characteristics. Since individual stocks in a specific test portfolio share the same risk characteristic such as size or BM, the idiosyncratic shocks to those individual stocks can be highly correlated and even be amplified in portfolio level. The idiosyncratic risk would be diversified away in the portfolio level only if the idiosyncratic shocks are not significantly correlated. Hence, we choose to randomly construct the portfolios to surely make well-diversified portfolios. The details are as follows.

First, we need to determine the number of individual stocks in each random portfolio. According to Campbell et al. (2001), an investor needs more and more stocks to hold diversified portfolios because idiosyncratic volatility significantly increased over time. In other words, the number of stocks to reduce the volatility of a portfolio to any given level has increased over time. For instance, in the period from 1974 to 1985, an investor needs 20 stocks to reduce annualized excess standard deviation to about five percent[2). However, in the 1986 to 1997 subsample, an investor needs at least 50 stocks to achieve the same level of excess standard deviation. Therefore, conservatively, I randomly choose 100 individual stocks from the stock universe to construct the well-diversified portfolio.

Second, we need to think about the holding period of these random portfolios and what

1) This assumes that lagged idiosyncratic volatility is a proxy for ex ante idiosyncratic volatility. For other measure, Fu (2009) uses E-GARCH model to estimate ex ante idiosyncratic volatility. However, Jin (2013) shows that, as a proxy for ex ante idiosyncratic volatility, E-GARCH model doesn't perform as well as simple lagged idiosyncratic volatility.

2) It is defined as the difference between the standard deviation of portfolio and the standard deviation of an equally weighted index.

to do with rebalancing of the portfolios. For typical Fama-French test portfolios, each test portfolio once constructed would be held for a year and then rebalanced to keep the same risk characteristic for the portfolio. However, the issues with the random portfolios stem from that they do not represent any specific risk characteristic (they are intentionally constructed to avoid those features) and that the betas of the portfolios would still vary over time because the betas of the individual stocks in these random portfolios are known to vary over time. In addition, since new stocks are added every year and many of the existing stocks are delisted, we need to regularly update the information about the new stock universe by reconstructing random portfolios.

Table 2. Number of Listed Firms in Formation Dates

Period	Date of Formation	The number of stocks
1	196307	2,059
2	196807	2,176
3	197307	5,484
4	197807	4,825
5	198307	5,696
6	198807	7,023
7	199307	7,231
8	199807	8,940
9	200307	6,714
10	200807	6,903
11	201307	6,646

Due to these issues, we cannot maintain the same composition of the random portfolios indefinitely, but we cannot rebalance the random portfolios by re-sorting the individual stocks with a certain criteria to keep their risk characteristics. We decide to follow the practice in Fama-MacBeth regression in that we assume that the betas of the random portfolios would be stable over five year period. We hold the random portfolios for five years, and we create a new set of random portfolios. The five year holding period will be non-overlapping and the returns over the period will be treated separately from the returns over the next five sample period. Therefore, in our sample, there are 11 periods, which span from Aug 1963 to July 2018, with a length of 5-year each. We end our sample period in 2018 because the extended period of another 5 years from 2018 to 2023 would include the COVID-19 pandemic period in which the stock market showed irregular and extreme behavior.

Finally, we need to choose the number of random portfolios to be constructed in each of five- year period. Table 2 shows the number of stocks in the dataset in each formation date of random portfolios. According to Table 2, in our sample period, from Aug 1963 to July 2018, there are about 2,000 stocks in 1963 and about 6,600 stocks in 2018. Since making too many random portfolios can yield a high correlation among randomly selected portfolios, we construct 20 random portfolios in each of the 5 year period[3]. We construct the random portfolios

3) When we construct 100 random portfolios for each period, we obtain similar results for empirical test in later

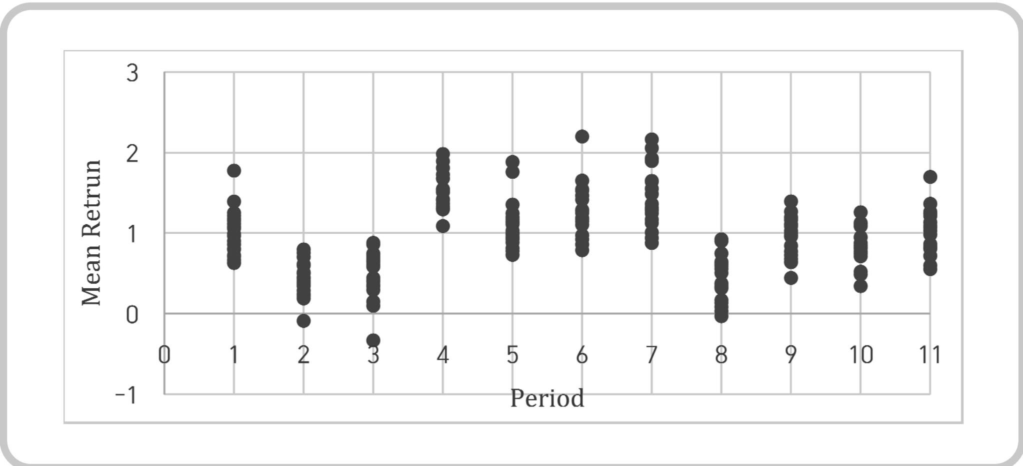
by sampling with replacement, so each of 20 random portfolios is constructed based on the same stock universe.[4)]

In sum, the random portfolio formation of this paper is as follows. From August 1963 to July 2018, in every 5 years (1 period), 20 portfolios are randomly constructed with 100 individual stocks in each portfolio. Then, the composition in each of 20 portfolios are fixed for 5 years and, during the 5 years of holding period, we calculate the value-weighted and equally weighted returns of 20 portfolios, so we obtain the monthly time-series return data. In our final dataset, we have 11 holding periods and 20 portfolios in each holding period, and each portfolio has 60 monthly data points or about 1,250 daily data points in each period. In Figure 1, the mean returns of 20 portfolios for each period are presented. Panel A and B of Figure 1 show the plot of value-weighted mean returns and equally-weighted mean returns, respectively. Compared to value-weighted returns in Panel A, equally-weighted returns in Panel B tend to be more clustered in each period.

Because individual stocks in each random portfolio are held for 5 years, it is important to consider the returns of stocks that are delisted from the CRSP database. In most cases, CRSP provides delisting returns for delisted firms and we use delisting returns whenever they are available. If delisting returns are not available, then we follow the method of Shumway (1997) and Campbell (2008). Shumway (1997) suggests that, if firms are delisted with performance-related reasons, then missing CRSP delisting returns should be replaced by -0.3, which is the median value of performance-related delisting returns.[5)] Moreover, according to Campbell (2008), if firms are delisted with some other reasons, then missing CRSP delisting returns are replaced by the returns of the previous month.

Fig. 1. Summary Statisttics of Randomly Selected Portfolios

Panel A. Value-Weighted Random Portfolios

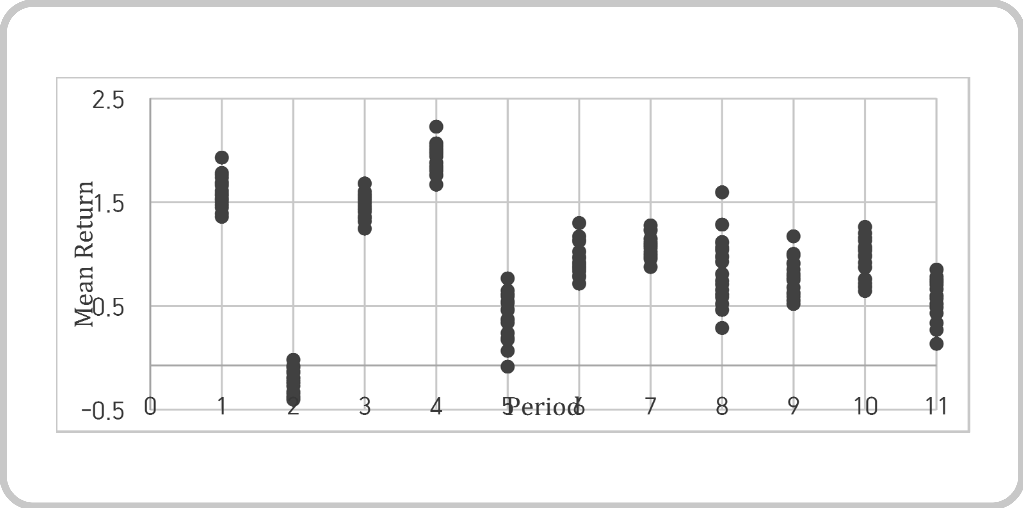


section.

4) When we use the method of random sampling without replacement, the results are similar to those of this paper.

5) Performance-related delisting codes are 500, and 520-599.

Panel B. Equally-Weighted Random Portfolios



Now that we have return series of the random portfolios, idiosyncratic volatilities of the random portfolios are estimated by two methods. First, idiosyncratic volatility is estimated as in Ang et al. (2006) but with the FF-5 factor model. In each month, by using daily return data of the previous month, idiosyncratic volatility is measured relative to the FF-5 model. Second, idiosyncratic volatility is defined relative to the same FF-5 model but it is estimated with the entire 5-year daily data of one period. Hence, each of the 220 random portfolios has single idiosyncratic volatility estimates for each period. This measure of the idiosyncratic volatility is more reliable than the 5-year average of monthly idiosyncratic volatility series. The usage of this second measure of the idiosyncratic volatility is explained in the later section.

3. The Relation between Return and Idiosyncratic Volatility at a Porfolio Level

Since idiosyncratic volatility of each random portfolio is solely composed of the component from the missing factor, the relation between the returns of portfolio and the component from the missing factor can be directly examined by investigating the relationship between the portfolios' return and idiosyncratic volatility.

In each month, 20 portfolios are ordered based on the idiosyncratic volatility and each of portfolios is labeled based on its idiosyncratic volatility. Since each of 20 portfolios is already a portfolio, we do not construct quintile or decile portfolios of the portfolios. Portfolio 1 (20) is the portfolio with the lowest (highest) idiosyncratic volatility. We also construct a long-short portfolio that takes a long position in portfolio 20 and a short position in portfolio 1 (labeled as 20-1). Including the long-short portfolio, we obtain 21 time-series of returns of portfolios ordered based on the idiosyncratic volatility. Since 20 portfolios are fixed for 5 years (1 period) and newly constructed, tests are conducted 11 times for 11 periods. Table 3 shows the results for the portfolio 20-1, a long-short portfolio in each period.

Table 3. Returns of Randomly Selected Portfolios Sorted by Idiosyncratic Volatility

Panel A. Value-Weighted 20 Portfolios

Period	Portfolio	RET	RET_T-stat	Stdev	Alpha	Alpha_Tstat
1		0.17	0.56	2.37	0.12	0.35
2		-0.26	-0.80	2.52	-0.21	-0.60
3		0.38	0.66	4.39	0.88	1.42
4		-0.39	-0.98	3.09	-0.39	-0.90
5		-0.16	-0.43	2.85	-0.11	-0.22
6	20-1	-0.48	-0.96	3.88	-0.71	-1.29
7		-0.58	-1.09	4.15	-1.25	-1.88*
8		-0.43	-0.39	8.68	1.25	1.14
9		-0.16	-0.37	3.37	-0.31	-0.65
10		-0.13	-0.36	2.76	-0.16	-0.40
11		0.74	2.05**	2.80	0.43	1.08

Panel B. Equally-Weighted 20 Portfolios

Period	Portfolio	RET	RET_T-stat	Stdev	Alpha	Alpha_Tstat
1		0.22	0.91	1.88	-0.07	-0.27
2		0.03	0.20	1.31	0.07	0.37
3		0.16	0.53	2.37	-0.03	-0.09
4		0.28	1.12	1.98	0.15	0.53
5		-0.28	-0.87	2.56	-0.19	-0.45
6	20-1	-0.50	-1.38	2.83	-0.57	-1.33
7		0.10	0.45	1.81	0.25	0.81
8		0.18	0.33	4.20	0.33	0.60
9		-0.27	-1.25	1.74	-0.26	-1.15
10		-0.59	-1.43	3.31	-0.55	-1.24
11		-0.17	-0.72	1.82	-0.23	-0.91

Panel A of Table 3 shows that mean returns of the (value-weighted) long-short portfolio are not significantly different from zero except for the period 11 although most of them have negative signs. The mean return of the portfolio with the highest idiosyncratic volatility is not significantly different from the mean return of the portfolio with the lowest idiosyncratic volatility. Also, in almost all periods, Jensen's alpha of portfolio 20-1 is not significantly different from zero.

Panel B of Table 3 reports similar results. Both of the mean returns and alphas are not significantly different from zero. These results suggest that the estimated idiosyncratic volatility, which is solely represented by the component from the missing factor, is not statistically related with the stock returns. Hence, the results suggest that, in the idiosyncratic volatility puzzle from Ang et al. (2006), a missing factor has no important role.

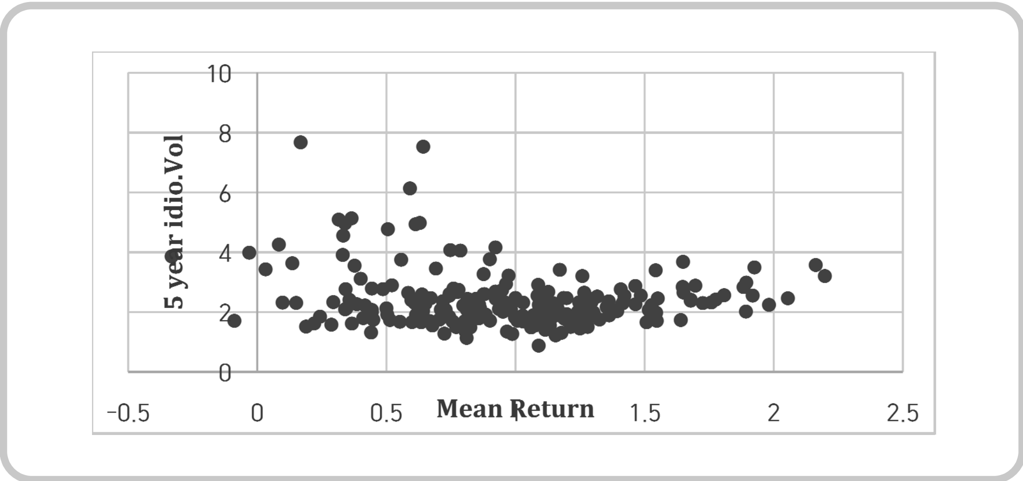
4. All-in-One Cross-Sectional Relationship between Mean Return and Idiosyncratic Volatility (Missing Factor)

Since 20 portfolios in each period are constructed from 100 individual stocks and share the same period of time, their mean returns may not significantly differ from one another.

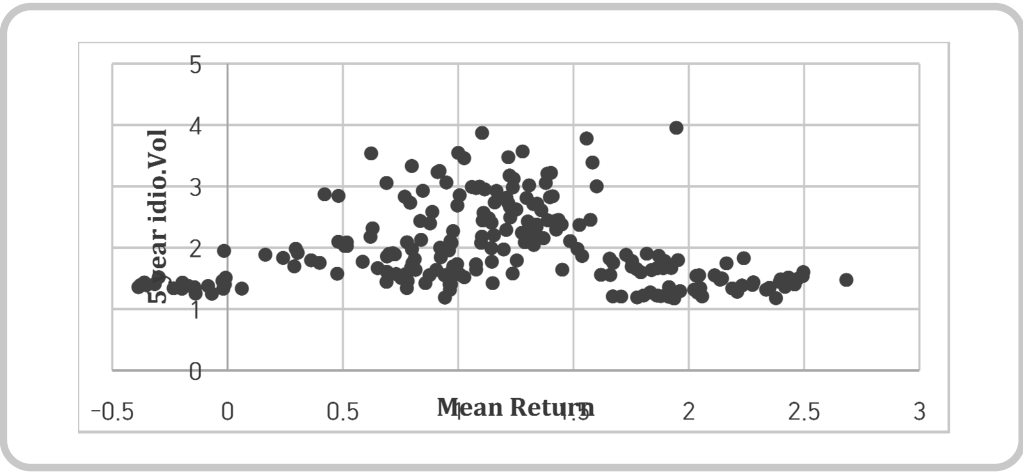
Hence, we consider all the random portfolios in different periods at the same time. 20 random portfolios are newly constructed every 5 years, and there are 220 portfolios across all 11 periods. Since every portfolio is randomly selected and fixed for 5 years, these 220 random portfolios from all periods can be thought as different cross-sectional data points. This is possible because $\beta_{p,2}^2 Var_{t-1}(F_{2,t})$ in equation (9) will get close to $\beta_{p,2}^2 Var(F_{2,t}) = \beta_{p,2}^2 \times constant$ as you estimate the idiosyncratic volatility by taking average of 5-year data. Hence, we gather all the portfolios from the entire 11 periods and examine the relationship between the mean returns and the idiosyncratic volatilities of the 220 portfolios.

Fig. 2. Mean Return and Idiosyncratic Volatility for 220 Portfolios

Panel A. Value-Weighted Random Portfolios



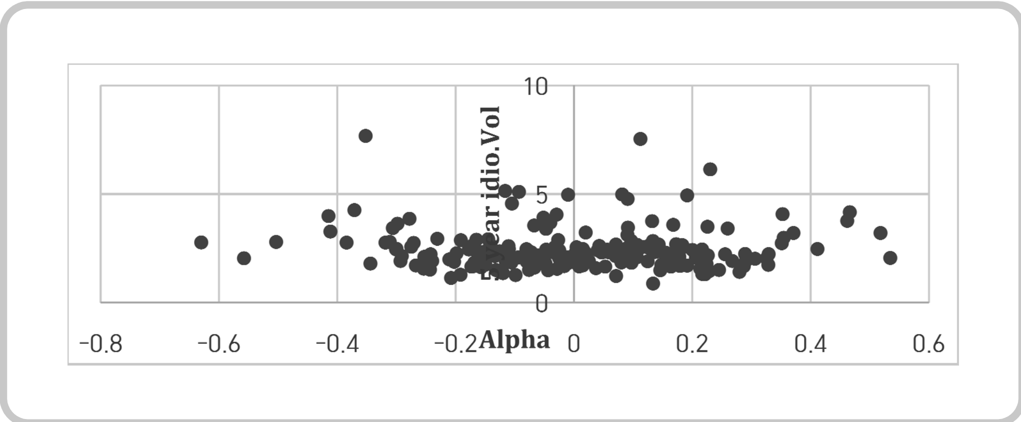
Panel B. Equally-Weighted Random Portfolios



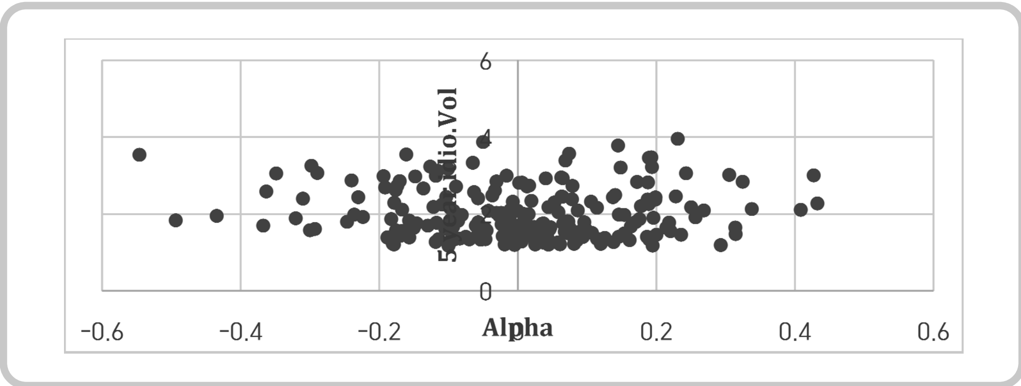
We estimate the idiosyncratic volatility relative to the FF-5 model with the entire 5-year data. This method gives more reliable data of idiosyncratic volatility than just a 5-year average of monthly idiosyncratic volatilities[6]. Figure 2 shows the scatter plot of mean returns and idiosyncratic volatilities. The horizontal axis represents mean return and the vertical axis represents idiosyncratic volatility. Plots in Figure 2 show that there is no distinct and significant pattern in both value-weighted (Panel A) random portfolios and equally weighted (Panel B) random portfolios. Although the patterns of value-weighted portfolios and equally weighted portfolios are quite different, both patterns indicate that mean returns and idiosyncratic volatilities are not related. These results also support that the effect of missing factors has no important role in explaining the idiosyncratic volatility puzzle.

Fig. 3. Alpha and Idiosyncratic Volatility for 220 Portfolios

Panel A. Value-Weighted Random Portfolios



Panel B. Equally-Weighted Random Portfolios



6) With 5-year average of monthly idiosyncratic volatilities, we obtain the similar results.

Now, we examine the relationship between the alphas of random portfolios from Fama-Macbeth regression and the idiosyncratic volatilities. For 20 portfolios in each period, we run monthly Fama-Macbeth regression with the factors of the FF-5 model. In the first stage, by using 5-year data, the betas of 5 factors are estimated by monthly time-series regression. The time-series regressions are conducted for each of the 20 portfolios. In the second stage, with the estimated betas as a regressor, we run the cross-sectional regression without intercept to calculate alphas for each of 20 portfolios and alphas for each portfolio are calculated by the sum of residuals in second stage regression[7)]. The cross-sectional regressions are conducted each month in the 5-year period. We repeat this procedure 11 times for the entire 11 periods, and we obtain 220 alphas for each of 220 random portfolios. Figure 3 shows the scatter plot of alphas and idiosyncratic volatilities. The horizontal axis represents the estimated alpha from the Fama-Macbeth regression and the vertical axis represents the idiosyncratic volatility. Figure 3 shows that there are no distinctive patterns in the cases of both value-weighted (Panel A) and equally weighted (Panel B) random portfolios. Similar to the results in Figure 2, these results in Figure 3 also support that the effect of a missing factor has no significant role in explaining the idiosyncratic volatility puzzle.

Ⅳ. Conclusion

In this paper, we study the impact of a missing factor in the context of the idiosyncratic volatility puzzle. Although asset pricing theory suggests that idiosyncratic risk shouldn't be priced or positively priced under certain conditions, Ang et al. (2006) find that idiosyncratic risk is negatively related to the cross-sectional stock returns. This phenomenon is called the idiosyncratic volatility puzzle and it poses a challenge to the traditional asset pricing theory. However, since idiosyncratic risk is unobservable, it should be estimated for empirical analysis by using a specific asset pricing model. Since all asset pricing models are not perfect, there must be some missing factors and these can distort the measurement of the idiosyncratic volatility. In such case, one might argue that the idiosyncratic volatility puzzle is driven by the effect of the missing factors, not by the pure idiosyncratic volatility. Hence, one needs to check whether the effect of the missing factors is really important in the context of the idiosyncratic volatility puzzle.

We carefully examine the effect of missing factors by using well-diversified portfolios. By constructing well-diversified portfolios, we eliminate the pure idiosyncratic risk. Then, the estimated idiosyncratic risk of the well-diversified portfolios is solely represented by the term related with the missing factors. By examining the relationship between the mean returns and the estimated idiosyncratic volatilities of the well-diversified portfolios, we can evaluate the significance of the missing factors in the idiosyncratic volatility puzzle. The results from the several tests confirm that missing factors do not have a significant relationship with the returns of the well-diversified portfolios. These results suggest that the effect of missing factors does not largely distort the measurement of the idiosyncratic volatility, and the idiosyncratic volatility

7) To calculate the alphas for each of 20 portfolios, it is necessary to run the cross-sectional regression without intercept. If we include the intercept in the regression, we can only obtain the alphas for each month.

puzzle is indeed driven by pure idiosyncratic risk. Hence, this paper supports the results of the previous literature about the idiosyncratic volatility puzzle.

References

Ang, A., R. J. Hodrick, Y. Xing and X. Zhang (2006), "The Cross-Section of Volatility and Expected Returns", *The Journal of Finance*, 61(1), 259-299.

Avramov, D., T. Chordia, G. Jostova and A. Philipov (2013), "Anomalies and Financial Distress", *Journal of Financial Economics*, 108(1), 139-159.

Bali, T. G. and N. Cakici (2008), "Idiosyncratic Volatility and the Cross Section of Expected Returns", *Journal of Financial and Quantitative Analysis*, 43(1), 29-58.

Campbell, J. Y., J. Hilscher and J. Szilagyi (2008), "In Search of Distress Risk", *The Journal of Finance*, 63(6), 2899-2939.

Campbell, J. Y., M. Lettau, B. G. Malkiel and Y. Xu (2001), "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk", *The Journal of Finance*, 56(1), 1-43.

Fama, E. F. and K. R. French (1993), "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33(1), 3-56.

Fama, E. F. and K. R. French (2015), "A Five-Factor Asset Pricing Model", *Journal of Financial Economics*, 116(1), 1-22.

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

Han, Y. and D. Lesmond (2011), "Liquidity Biases and the Pricing of Cross-Sectional Idiosyncratic Volatility", *The Review of Financial Studies*, 24(5), 1590-1629.

Huang, W., Q. Liu, S. G. Rhee and L. Zhang (2009), "Return Reversals, Idiosyncratic Risk, and Expected Returns", *The Review of Financial Studies*, 23(1), 147-168.

Hou, K. and R. K. Loh (2016), "Have We Solved the Idiosyncratic Volatility Puzzle?", *Journal of Financial Economics*, 121(1), 167-194.

Jin, L. L. (2013), "Idiosyncratic Volatility, Arbitrage Risk, and Anomaly Returns", *Working Paper*, University of Pennsylvania.

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

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

Shumway, T. (1997), "The Delisting Bias in CRSP Data", *The Journal of Finance*, 52(1), 327-340.

Stambaugh, R. F., J. Yu and Y. Yuan (2015), "Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle", *The Journal of Finance*, 70(5), 1903-1948.