

The Idiosyncratic Volatility Puzzle: A Time-Specific Anomaly

Xindong Zhang, Jianying Li*, Xiaoli Wang, Xiaoxin Hu

School of Economics and Management, Shanxi University, Taiyuan, China

Email: *lee_amaris@yeah.net

How to cite this paper: Zhang, X.D., Li, J.Y., Wang, X.L. and Hu, X.X. (2021) The Idiosyncratic Volatility Puzzle: A Time-Specific Anomaly. *Journal of Mathematical Finance*, 11, 294-312.

<https://doi.org/10.4236/jmf.2021.112017>

Received: March 31, 2021

Accepted: May 28, 2021

Published: May 31, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Given the availability of daily data over 1926-1962, it is surprising that there is no research examining the idiosyncratic volatility (IV) puzzle over this early period. This paper conducts an out-of-sample test on the IV phenomenon. We find that the negative relation between IV and expected returns only exists during the period 07/1963-12/1989, implying that the puzzle may be a result of data snooping bias. The result on time-special anomaly is robust for different sorting breakpoints and alternative measure of idiosyncratic volatility. Infrequent trading cannot account for the low average returns of stocks with high idiosyncratic volatility. With a striking contrast, the involving of short-term return reversals eliminates this dilemma.

Keywords

Idiosyncratic Volatility Puzzle, Out-of-Sample Test, Data Snooping Bias, Short-Term Return Reversals

1. Introduction

The systematic risk principle shows that firm-specific risk or idiosyncratic volatility (IV) should not carry a premium, whereas IV should positively predict return under Merton's [1] incomplete-information model¹ [2] [3] [4] [5]. However, Ang *et al.* [6] [7] find that IV predicts returns negatively. This contradictory finding to theory, which is referred to as the idiosyncratic volatility puzzle, has attracted a great deal of attention in the literature² [8]-[21]. In this paper, we in-

¹There are studies reporting consistent evidence with Merton [1], e.g., Malkiel and Xu [2], Spiegel and Wang [3], Fu [4], and Chua *et al.* [5].

²Studies have attempted to find the causes of the IV puzzle, e.g., the abnormal turnover (Connolly and Stivers, [8]), the expected idiosyncratic skewness (Boyer *et al.* [9]), coskewness (Chabi-Yo and Yang [10]), short-term price reversal (Huang *et al.* [11]), illiquidity (Bali and Cakici [12]; Han and Lesmond [13]), earnings surprises (Jiang *et al.* [14]; Wong [15]), maximum daily return (Bali *et al.* [16]), average variance (Chen and Petkova [17]), proportion of retail trading (Han and Kumar [18]), and lottery preferences of investors (Barberis and Huang [19]; Hou and Loh, [20]). An early study by Johnson [21] also shows the negative IV-return relation being due to unpriced information risk.

investigate whether the IV effect is a result of data snooping bias. In particular, it is surprising that there is no study examining the IV-return relation over the early 1926-1962 period when daily data have been available since 2006. Thus, we attempt to test whether the idiosyncratic volatility puzzle is a result of data snooping bias. Furthermore, we try to explore the question of what drives the puzzle in a specific period in which idiosyncratic volatility is robust.

Specifically, we tackle the following issues in this paper. First, we focus on the performance of IV puzzle in different periods. We use the standard deviation of residuals from the Fama-French three-factor model and CRSP breakpoints to allocate all stocks to ten groups (Low to High) during the periods 07/1926-06/1963 and 07/1963-12/2014. The results of portfolios analysis suggest that the significant negative relation between IV and subsequent stock returns exists during 07/1963-12/2014. To further examine the time-specific performance of the IV puzzle, we divide 07/1963-12/2014 into two subsamples in average. We find that the puzzle is distinct during the first period 07/1963-12/1989. To make the results more robust, we also use NYSE breakpoints and the standard deviation of residuals from the Fama-French five-factor model to repeat the process, respectively. The results do not make us disappoint, and increase our interest in continuing to explore the reasons for the prominent puzzle over this period.

Second, we examine the role of infrequent trading in dissecting the idiosyncratic volatility puzzle. Liu [22] shows that the addition of liquidity factor based on zero daily trading volumes can explain established market anomalies such as size, book-to-market, and long-termcontrarian premiums. Similarly, some studies also provide the evidence that infrequent trading has a significant effect on the efficiency of stock markets [23] [24] [25] [26]. As a consequence, we construct a rearranged CRSP daily return considering the influence of infrequent trading and use a five-factor model to estimate the IV. Similarly, we construct decile portfolios based on the new proxy for IV. The results of portfolio returns suggest that the negative relationship between IV and expected returns is not alleviated when infrequent trading is considered during the prominent puzzle period 07/1963-12/1989.

Third, following Fu [4] and Huang *et al.* [11], who find that the return reversals help explain the well-known IV puzzle, we examine whether the puzzle disappears after short-return reversals are controlled for. We proceed the cross-sectional regression and calculate the time-series averages of estimated coefficients. As we expected, the negative relation between IV and expected returns is particularly significant when the stock returns in the previous month are not involved. However, the negative significant relationship is gone after involving the individual stock return during the previous month. We don't stop footstep there. We move on with a portfolio analysis as before. Specifically, we leave a one-month gap between the portfolio formation date and holding period. We form ten portfolios every month by sorting all stocks with standard deviation of the residual from a five-factor model. Supporting the findings of Fu [4] and Huang *et al.*

[11], we find that the H-L portfolio return is negative but insignificant during 07/1963-12/1989.

We contribute to the literature in the following ways. First, we extend the time horizon of empirical studies on IV anomaly. Given the fact that the research on U.S. stock market anomalies begins in July 1963 in general, yet the studies using pre-1963 data are rare. Motivated by Davis *et al.* [27], we extend the data of Ang *et al.* [7] back to 1926. Second, and perhaps most importantly, this paper provides evidence that the idiosyncratic volatility puzzle is a result of data snooping bias. Specifically, we find the puzzle exists only during 07/1963-12/1989 and is not robust either the period 07/1926-06/1963 or the period 01/1990-12/2014. In this regard, our paper adds to the growing strand of literature that attempts to explain the IV puzzle in a limited sample. Third, we explore an alternative proxy for idiosyncratic volatility. We use the standard deviation of residuals from Fama-French five-factor model to measure idiosyncratic volatility. As is well-known, a five-factor model that adds profitability and investment factors to the three-factor model of Fama and French [28] performs better in the interpretation of anomalies [29]. This consideration doesn't alter our results to a certain degree. Fourth, although the causes of the IV puzzle still remain controversial, our results shed some light on the issue by supporting the role of short-time return reversals in explaining the puzzle.

The remainder of the paper is organized as follows. Section 2 describes the data and the measure of realized idiosyncratic volatility. Section 3 empirically analyzes the performance of IV puzzle in different sample periods. Section 4 shows how does the infrequent trading and short-term return reversals play a role in explaining the puzzle. Section 5 concludes the paper.

2. Data and Sample Construction

2.1. Data Sources and Sample

Our data include daily and monthly returns of NYSE, AMEX, and NASDAQ common stocks with share codes 10 or 11 from July 1926 to December 2014. We obtain daily and monthly returns, daily trading volume, and share code data from the CRSP and the book value of individual companies from Compustat. We follow the procedure adjusting for delisting firms used by Shumway [30] in daily and monthly returns. We use the one-month Treasury bill rate as the risk-free rate. Moreover, we take into account the Size, B/M, OP, AGR, and Zero Volume in the explaining of the IV puzzle. Size is the natural logarithm of market capitalization and B/M is the natural logarithm of book-to-market in month t . OP is the book-equity-deflated operating profitability and AGR is the total asset growth rate. To consider the impact of infrequent trading, we also define Zero_Volume as the stocks with zero trading volume.

2.2. Measure of Realized Idiosyncratic Volatility

We measure realized idiosyncratic volatility following the approach in Ang *et al.*

[6] [7] and Bali and Cakici [12]. Specifically, for one month, we perform the following Fama and French [28] three-factor regression for firms that have more than 15 daily return observations in that month:

$$r_{d,m}^i - r_f = \alpha_m^i + \beta_{\text{MKT}}^{i,m} (\text{MKT}_{d,m} - r_f) + \beta_{\text{SMB}}^{i,m} \text{SMB}_{d,m} + \beta_{\text{HML}}^{i,m} \text{HML}_{d,m} + \varepsilon_{d,m}^i \quad (1)$$

$$\text{IV}_{m\text{-FF3FM}}^i = \sqrt{\text{Var}(\varepsilon_{d,m}^i)} \quad (2)$$

where, for day d in month m , $r_{d,m}^i$ is stock i 's market return, $\text{MKT}_{d,m} - r_f$ is the market excess return, $\text{SMB}_{d,m}$ and $\text{HML}_{d,m}$ are the returns on portfolios formed to capture size and book-to-market effects, respectively³. Specifically, $\text{SMB}_{d,m}$ is the return on a diversified portfolio of small stocks minus the return on a diversified portfolio of big stocks, $\text{HML}_{d,m}$ is the difference between the returns on diversified portfolios of high and low stocks. $\varepsilon_{d,m}^i$ denotes the resulting residual. We use the standard deviation of daily residuals in month m to measure the individual stock's idiosyncratic volatility $\text{IV}_{m\text{-FF3FM}}^i$ for this month.

A five-factor model that adds profitability (RMW) and investment (CMA) factors performs better than the three-factor model [29] [31]. Therefore, we also consider a idiosyncratic volatility proxy based on the residuals of following Fama and French [31] five-factor model:

$$r_{d,m}^i - r_f = \alpha_m^i + \beta_{\text{MKT}}^{i,m} (\text{MKT}_{d,m} - r_f) + \beta_{\text{SMB}}^{i,m} \text{SMB}_{d,m} + \beta_{\text{HML}}^{i,m} \text{HML}_{d,m} + \beta_{\text{RMW}}^{i,m} \text{RMW}_{d,m} + \beta_{\text{CMA}}^{i,m} \text{CMA}_{d,m} + \varepsilon_{d,m}^i \quad (3)$$

$$\text{IV}_{m\text{-FF5FM}}^i = \sqrt{\text{Var}(\varepsilon_{d,m}^i)} \quad (4)$$

where $\text{RMW}_{d,m}$ is the difference between the returns on diversified portfolios of stocks with robust and weak profitability, and $\text{CMA}_{d,m}$ is the difference between the returns on diversified portfolios of stocks of low (conservative) and high (aggressive) investment firms. $\text{IV}_{m\text{-FF5FM}}^i$ is the individual stock's idiosyncratic volatility for m month derived from the Equation (4).

2.3. Summary Statistics

To see whether the IV anomaly exists only for a specific period of time. In other words, whether the IV puzzle can be explained by data snooping bias. We divide the whole sample period (07/1926-12/2014) into two subsample periods (07/1926-06/1963 and 07/1963-12/2014). **Table 1** presents the descriptive statistics and Spearman rank correlation for the main variables involved in this paper. Panel A reports the number and the proportion of Zero_Volumestocks during 07/1926-12/2014, 07/1926-06/1963 and 07/1963-12/2014. We notice that the number of Zero_Volume stocks is 17359 and account for 71.99% of the total common stocks. In the whole sample (07/1926-12/2014) and two subsamples (07/1926-06/1963 and 07/1963-12/2014), the proportion of Zero_Volume stocks is 12.95%, 25.77% and 11.07%, respectively. This implies that the number of

³We sincerely thank Kenneth French for making the data available on his website.

Table 1. Descriptive statistics for basic data.

Panel A: The statistics for the proportion of zero daily volumes						
Periods	R_original	R_rearrange	Zero_Volume	Proportion		
07/1926-12/2014	70,614,037	61,470,621	9,143,416	12.95%		
07/1926-06/1963	9,027,597	6,701,477	2,326,120	25.77%		
07/1963-12/2014	61,586,440	54,769,144	6,817,296	11.07%		
Number of stocks	24,114	24,103	17,359	71.99%		
Panel B: The descriptive statistics of main variables during 07/1926-12/2014						
Variables	N	Mean	Std.	Min	Max	
IV_FF3FM	3,305,450	0.0272	0.0085	0.0104	0.0625	
IV_FF3FM_Infre	2,777,844	0.0282	0.0092	0.0106	0.0839	
Size	2,254,957	4.7406	2.1995	-4.5504	13.4553	
B/M	2,254,957	0.9808	4.1963	0.0000	3040.55	
OP	2,254,957	14.7785	2334.16	-194,133.33	942,375.00	
AGR	2,254,957	19.2963	767.7825	-98.0832	789,850.00	
Panel C: The correlation of main variables during 07/1926-12/2014						
Variables	IV_FF3FM	IV_FF3FM_Infre	Size	B/M	OP	AGR
IV_FF3FM	1.0000	0.9897	-0.4348	0.1317	-0.0675	0.0136
IV_FF3FM_Infre	0.9897	1.0000	-0.5788	0.1337	-0.0751	0.0199
Size	-0.4348	-0.5788	1.0000	-0.2919	0.0758	0.0462
B/M	0.1317	0.1337	-0.2919	1.0000	-0.0781	-0.0899
OP	-0.0675	-0.0751	0.0758	-0.0781	1.0000	0.0412
AGR	0.0136	0.0199	0.0462	-0.0899	0.0412	1.0000

stocks with zero trading volume is quite substantial and Zero_Volume stocks shouldn't be ignored for the research on empirical asset pricing especially anomalies in stock market, which is in accordance with the argument of Liu [22] that understanding the role of stocks with zero daily trading volumes is an important key to explain anomalies.

Panel B and C report the number, means, standard deviations, minimum and maximum, and pairwise correlations for main variables during 07/1926-12/2014, respectively. IV_FF3FM is the idiosyncratic volatility defined as the standard deviation of residuals from the regression of daily return with respect to local Fama-French three-factor model in the previous month. The daily return for IV_FF3FM is based on CRSP original daily return, while is rearranged CRSP daily return after considering the influence of infrequent trading for IV_FF3FM_Infre. Size is the natural logarithm of market capitalization. B/M is the natural logarithm of book-to-market in month t . OP is the book-equity-deflated operating profitability and AGR is the total asset growth rate. We see that the value of IV_FF3FM_Infre is higher than IV_FF3FM in terms of mean, standard deviations, minimum and maximum. As expected, the IV_FF3FM has a downward bias compared with new idiosyncratic volatility measure considering the influ-

ence of infrequent trading, which provides strong evidence that the calibration for the daily return of Zero_Volume stocks is feasible. Panel C shows that IV_FF3FM_Infre is highly correlated with IV_FF3FM and the coefficient of association is 0.9897, which indicates that IV_FF3FM_Infre could be a reasonable proxy for idiosyncratic risk. Moreover, the correlation between IV_FF3FM and Size means small firms have high idiosyncratic volatility. The IV_FF3FM is also negatively correlated with OP, indicating that firms with low profitability have higher volatility. All the representation of Spearman rank correlation of IV_FF3FM is tenable for the IV_FF3FM_Infre.

3. A Reexamination of the Idiosyncratic Volatility Puzzle

In this section, we empirically test whether the IV puzzle is robust for different sample periods. In doing so, we first examine the existence of IV puzzle during two subsamples (07/1926-06/1963 and 07/1963-12/2014) employing three-factor model. Further, we conduct the same process during 07/1963-12/1989 and 01/1990-12/2014. We reexamine afterwards the IV puzzle with five-factor model.

3.1. Compared Performance for Different Sample Periods

We duplicate the research method of Ang *et al.* [6] [7] and examine the IV puzzle during 07/1926-06/1963 and 07/1963-12/2014. In this process, idiosyncratic volatility is measured by IV_FF3FM, which is the standard deviation of residuals from Fama-French three-factor model. At the end of each month, stocks are allocated to ten groups (Low to High) according to IV_FF3FM and using CRSP breakpoints. We compute the average returns and intercepts from the Fama-French three factor regressions for decile portfolios in **Table 2**. Panel A of **Table 2** shows that the difference of return (-0.28) between the highest decile IV portfolio and the lowest decile IV portfolio is inapparent for portfolios formed by value-weight during 07/1926-06/1963. Interestingly, the difference of return (0.18) between highest and lowest portfolio presents a positive trend when equal-weighted portfolios are formed, although it is significant at close to the 10% level. In stark contrast to Panel A, the H-L returns are -1.12 ($t = -3.20$) and -0.03 ($t = -0.07$) for value-weighted and equal-weighted decile portfolios during 07/1963-12/2014 in Panel B, implying that portfolios with high idiosyncratic volatility have lower return than portfolios with low idiosyncratic volatility. The Alpha has similar performance for the two samples. The results show that the idiosyncratic volatility puzzle is not robust during 07/1926-06/1963 and only exists during 07/1963-12/2014. Therefore, we infer that the idiosyncratic volatility puzzle is a result of data snooping bias.

To further examine the time-specific performance of the IV puzzle, we divide 07/1963-12/2014 into two subsamples in average, the first subsample period is 07/1963-12/1989 and the second subsample period is 01/1990-12/2014. **Table 3** presents that the average returns and intercepts from the Fama-French three

Table 2. Average return and Alpha of portfolios sorted by IV_FF3FM based on CRSP breakpoints during 07/1926-06/1963 and 07/1963-12/2014. This table reports the average returns and Alpha of ten portfolios sorted by idiosyncratic volatility relative to the Fama and French [28] model. Portfolios formed every month are based on idiosyncratic volatility computed using CRSP original daily return over the previous month. Portfolio Low (High) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. Returns are measured in monthly percentages. We consider the average monthly return of value-weighted and equal-weighted portfolios, where the weights are based upon market capitalization at the end of month. Panel A reports the results of decile portfolios using CRSP breakpoints during July 1926 to June 1963. Panel B reports the results of decile portfolios using CRSP breakpoints during July 1963 to December 2014. Alpha is the intercept of the Fama-French three-factor model. The row of H-L refers to the difference between highest and lowest idiosyncratic volatility portfolio. The Newey and West [32] t -statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Value-weight			Equal-weight		
	Return	Alpha	(t -value)	Return	Alpha	(t -value)
Panel A Portfolio return and Alpha of three-factor model during 07/1926-06/1963						
Low IV_FF3FM	0.99	0.18***	(4.04)	1.00	0.22***	(3.89)
2	1.07	0.10*	(1.85)	1.19	0.23***	(4.53)
3	1.04	-0.06	(-0.77)	1.26	0.14***	(2.03)
4	1.04	-0.13*	(-1.75)	1.32	0.11*	(1.74)
5	1.08	-0.19**	(-2.22)	1.43	0.12	(1.52)
6	1.02	-0.30***	(-2.90)	1.37	0.00	(0.00)
7	1.10	-0.22**	(-2.12)	1.43	0.01	(0.16)
8	0.86	-0.52***	(-3.94)	1.48	-0.12	(-1.18)
9	0.85	-0.50***	(-3.19)	1.44	-0.20*	(-1.83)
High IV_FF3FM	0.71	-0.63*	(-1.95)	1.81	0.08	(0.34)
H-L	-0.28	-0.81***		0.81	-0.14	
(t -value)	(-0.63)	(-2.45)		(1.36)	(-0.53)	
Panel B Portfolio return and Alpha of three-factor model during 07/1963-12/2014						
Low IV_FF3FM	0.92	0.10*	(1.73)	1.10	0.16***	(2.72)
2	0.95	0.05	(1.01)	1.21	0.14***	(2.82)
3	0.97	0.00	(0.07)	1.32	0.19***	(3.76)
4	1.01	-0.00	(-0.04)	1.38	0.20***	(3.98)
5	1.04	-0.00	(-0.03)	1.39	0.19***	(3.51)
6	1.10	0.02	(0.26)	1.36	0.12*	(1.92)
7	0.85	-0.29	(-0.78)	1.32	0.03	(0.43)
8	0.64	-0.59***	(-4.77)	1.19	-0.12	(-1.09)
9	0.27	-1.03***	(-6.05)	1.07	-0.30*	(-1.96)
High IV_FF3FM	-0.20	-1.50***	(-7.33)	1.07	-0.33	(-1.40)
H-L	-1.12***	-1.60***		-0.03	-0.49**	
(t -value)	(-3.20)	(-6.97)		(-0.07)	(-1.99)	

factor regressions for decile portfolios sorted by idiosyncratic volatility during 07/1963-12/1989 and 01/1990-12/2014. We see that H-L value-weighted portfolio return is significantly negative (-1.26) during 07/1963-12/1989, however, is indistinctive (-0.97) during 01/1990-12/2014. When the portfolios are formed

Table 3. Average return and Alpha of portfolios sorted by IV_FF3FM and based on CRSP breakpoints during 07/1963-12/1989 and 01/1990-12/2014. This table reports the average returns and Alpha of ten portfolios sorted by idiosyncratic volatility relative to the Fama and French [28] model. Portfolios formed every month are based on idiosyncratic volatility computed using CRSP original daily return over the previous month. Portfolio Low (High) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. Returns are measured in monthly percentages. We consider the average monthly return of value-weighted and equal-weighted portfolios, where the weights are based upon market capitalization at the end of month. Panel A reports the results of decile portfolios using CRSP breakpoints during July 1963 to December 1989. Panel B reports the results of decile portfolios using CRSP breakpoints during January 1990 to December 2014. Alpha is the intercept of the Fama-French three-factor model. The row of H-L refers to the difference between highest and lowest idiosyncratic volatility portfolio. The Newey and West [32] *t*-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Value-weight			Equal-weight		
	Return	Alpha	(<i>t</i> -value)	Return	Alpha	(<i>t</i> -value)
Panel A Portfolio return and Alpha of three-factor model during 07/1963-12/1989						
Low IV_FF3FM	0.89	0.01	(0.07)	1.06	0.02	(0.28)
2	0.92	-0.00	(-0.07)	1.23	0.11*	(1.93)
3	1.05	0.08	(1.53)	1.41	0.23***	(4.43)
4	1.14	0.16***	(2.68)	1.51	0.28***	(5.46)
5	1.13	0.06	(0.82)	1.50	0.23***	(4.51)
6	1.16	0.05	(0.59)	1.47	0.13**	(2.19)
7	0.99	-0.24***	(-2.43)	1.37	-0.05	(-0.63)
8	0.76	-0.51***	(-4.60)	1.29	-0.18*	(-1.82)
9	0.26	-1.14***	(-8.64)	1.01	-0.60***	(-4.63)
High IV_FF3FM	-0.37	-1.87***	(-10.54)	0.87	-0.95***	(-4.51)
H-L	-1.26***	-1.88***		-0.28	-0.97***	
(<i>t</i> -value)	(-3.28)	(-8.75)		(-0.71)	(-4.34)	
Panel B Portfolio return and Alpha of three-factor model during 01/1990-12/2014						
Low IV_FF3FM	0.98	0.22***	(2.91)	1.15	0.35***	(4.59)
2	1.01	0.12*	(1.71)	1.20	0.24***	(3.27)
3	0.92	-0.05	(-0.62)	1.26	0.22***	(2.86)
4	0.89	-0.16	(-1.51)	1.27	0.16**	(2.34)
5	0.97	-0.11	(-0.94)	1.30	0.19**	(2.08)
6	1.07	-0.07	(-0.50)	1.27	0.11	(0.99)
7	0.73	-0.48***	(-2.74)	1.28	0.08	(0.57)
8	0.55	-0.79***	(-3.75)	1.11	-0.11	(-0.59)
9	0.31	-1.10***	(-1.06)	1.15	-0.11	(-0.39)
High IV_FF3FM	0.01	-1.34***	(-3.75)	1.38	0.13	(0.30)
H-L	-0.97	-1.56***		0.23	-0.22	
(<i>t</i> -value)	(-1.62)	(-3.97)		(0.40)	(-0.51)	

by equal-weight, the difference of return between highest and lowest idiosyncratic volatility portfolio is negative (-0.28) for the first subsample but is positive (0.23) for the second subsample, although neither is significant. Similarity, the H-L Alpha is significant negative (-1.88 and -0.97) for value-weighted and equal-weighted portfolio during 07/1963-12/1989. The result suggests that the

idiosyncratic volatility puzzle only exists in the period of July 1963 to December 1989. That is to say, even for the second subsample, the negative relation between idiosyncratic volatility and expected returns is not robust, we can draw the conclusion that the anomaly really is a time-specific phenomenon.

In addition, to form decile portfolios with a relatively more balanced average market share, we also use the NYSE breakpoints which start with Fama and French [33]⁴ to examine the idiosyncratic volatility puzzle. **Table 4** reports the average return and Alpha of portfolios sorted by IV_FF3FM and based on NYSE breakpoints during 07/1963-12/1989 and 01/1990-12/2014. For value-weighted portfolios, the H-L return is -0.56% ($t = -1.63$) during 07/1963-12/1989 and is -0.54% ($t = -1.09$) during 01/1990-12/2014. For equal-weighted portfolios, the H-L return is -0.08% ($t = -0.25$) and 0.14% ($t = 0.18$), respectively. In total, for NYSE breakpoints, there is no negative and significant relation between idiosyncratic volatility and expected returns. However, in the first subsample period (07/1963-12/1989), the t-statistics is very close to the 10% significant level, so we still believe that the idiosyncratic volatility puzzle is significant during this period.

3.2. Alternative Test for the Idiosyncratic Volatility Puzzle

A five-factor model that adds profitability and investment factors to the three-factor model of Fama and French [28] largely absorbs the patterns in average returns [31] [34]. Especially in the anomalies of dissecting, the five-factor model shrinks the troublesome intercepts of the three-factor model [29]. Therefore, we will try again to examine the idiosyncratic volatility puzzle based on the five-factor model in this section. Different from the above measure of idiosyncratic volatility, we will use IV_FF5FM which is the standard deviation of residuals from Fama-French five-factor model to measure idiosyncratic volatility. To be more practical, we apply the NYSE breakpoints to allocate stocks. **Table 5** reports the average return and Alpha of portfolios sorted by IV_FF5FM and based on NYSE breakpoints during 07/1963-12/1989 and 01/1990-12/2014. The difference of return between highest and lowest idiosyncratic volatility portfolio is significantly negative (-0.58 , $t = -1.76$) during 07/1963-12/1989 for value-weighted portfolio, however, is inapparent (-0.47 , $t = -0.97$) during 01/1990-12/2014. Similarly, the H-L Alpha is also significantly negative (-0.92 , $t = -6.43$) during 07/1963-12/1989, however, is not significant (-0.20 , $t = -0.84$) during 01/1990-12/2014. The results show that the well-known idiosyncratic volatility puzzle (*i.e.*, a negative relation between the monthly realized idiosyncratic volatility in the previous month and the value-weighted portfolio return in the subsequent month) exists during 07/1963-12/1989. The results of equal-weighted portfolio are also support our conjecture that the idiosyncratic volatility puzzle is a time-specific anomaly.

⁴All NYSE, AMEX, and NASDAQ stocks are allocated to the 10 size portfolios using the NYSE breakpoints in their article.

Table 4. Average return and Alpha of portfolios sorted by IV_FF3FM and based on NYSE breakpoints during 07/1963-12/1989 and 01/1990-12/2014. This table reports the average returns and Alpha of ten portfolios sorted by idiosyncratic volatility relative to the Fama and French [28] model. Portfolios formed every month are based on idiosyncratic volatility computed using CRSP original daily return over the previous month. Portfolio Low (High) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. Returns are measured in monthly percentages. We consider the average monthly return of value-weighted and equal-weighted portfolios, where the weights are based upon market capitalization at the end of month. Panel A reports the results of decile portfolios using NYSE breakpoints during July 1963 to December 1989. Panel B reports the results of decile portfolios using NYSE breakpoints during January 1990 to December 2014. Alpha is the intercept of the Fama-French three-factor model. The row of H-L refers to the difference between highest and lowest idiosyncratic volatility portfolio. The Newey and West [32] t -statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Value-weight			Equal-weight		
	Return	Alpha	(t -value)	Return	Alpha	(t -value)
Panel A Portfolio return and Alpha of three-factor model during 07/1963-12/1989						
Low IV_FF3FM	0.86	-0.02	(-0.31)	1.02	-0.02	(-0.34)
2	1.01	0.12**	(2.09)	1.25	0.14**	(2.25)
3	1.05	0.11*	(1.75)	1.41	0.26***	(4.22)
4	1.13	0.18***	(2.95)	1.43	0.24***	(4.28)
5	1.09	0.08	(1.27)	1.49	0.27***	(4.78)
6	1.25	0.20***	(2.62)	1.56	0.31***	(5.88)
7	1.18	0.09	(1.11)	1.50	0.20***	(3.74)
8	1.14	-0.02	(-0.25)	1.51	0.12**	(1.97)
9	0.99	-0.28***	(-2.85)	1.38	-0.06	(-0.76)
High IV_FF3FM	0.30	-1.09***	(-9.40)	0.94	-0.70***	(-4.76)
H-L	-0.56	-1.07***		-0.08	-0.68***	
(t -value)	(-1.63)	(-7.15)		(-0.25)	(-4.03)	
Panel B Portfolio return and Alpha of three-factor model during 01/1990-12/2014						
Low IV_FF3FM	0.99	0.24**	(2.40)	1.11	0.37***	(4.55)
2	1.07	0.26***	(3.19)	1.12	0.23***	(3.02)
3	1.05	0.24**	(2.40)	1.22	0.28***	(3.65)
4	0.95	-0.00	(-0.01)	1.23	0.21***	(2.73)
5	0.95	-0.05	(-0.46)	1.24	0.20**	(2.40)
6	0.79	-0.27**	(-2.54)	1.21	0.13*	(1.69)
7	1.05	-0.01	(-0.09)	1.36	0.26***	(3.00)
8	1.02	-0.08	(-0.59)	1.31	0.17	(1.62)
9	1.12	-0.13	(-0.84)	1.36	0.15	(1.16)
High IV_FF3FM	0.45	-0.91***	(-4.31)	1.25	-0.01	(-0.03)
H-L	-0.54	-1.15***		0.14	-0.38	
(t -value)	(-1.09)	(-4.23)		(0.28)	(-1.34)	

In addition, compared with **Table 4**, the intercept term in **Table 5** has a massive improvement. Specifically, five of the ten value-weighted portfolios are significant (0.24, 0.26, 0.24, -0.27, -0.91) during 01/1990-12/2014 in **Table 4**, however, all are not significant in **Table 5**. The Alpha of highest value-weighted portfolio is -1.09 ($t = -9.40$) during 07/1963-12/1989 in **Table 4**, however,

Table 5. Average return and Alpha of portfolios sorted by IV_FF5FM and based on NYSE breakpoints during 07/1963-12/1989 and 01/1990-12/2014. This table reports the average returns and Alpha of ten portfolios sorted by idiosyncratic volatility relative to the Fama and French [31] model. Portfolios formed every month are based on idiosyncratic volatility computed using CRSP original daily return over the previous month. Portfolio Low (High) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. Returns are measured in monthly percentages. We consider the average monthly return of value-weighted and equal-weighted portfolios, where the weights are based upon market capitalization at the end of month. Panel A reports the results of decile portfolios using NYSE breakpoints during July 1963 to December 1989. Panel B reports the results of decile portfolios using NYSE breakpoints during January 1990 to December 2014. Alpha is the intercept of the Fama-French five-factor model. The row of H-L refers to the difference between highest and lowest idiosyncratic volatility portfolio. The Newey and West [32] *t*-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Value-weight			Equal-weight		
	Return	Alpha	(<i>t</i> -value)	Return	Alpha	(<i>t</i> -value)
Panel A Portfolio return and Alpha of five-factor model during 07/1963-12/1989						
Low IV_FF3FM	0.87	-0.05	(-0.78)	1.02	-0.07	(-0.84)
2	0.99	0.06	(0.92)	1.24	0.10	(1.43)
3	1.07	0.14**	(2.27)	1.38	0.24***	(3.86)
4	1.11	0.16**	(2.46)	1.47	0.25***	(4.31)
5	1.16	0.15**	(2.10)	1.49 0.4	0.27***	(4.68)
6	1.15	0.14*	(1.81)	1.52	0.28***	(5.42)
7	1.23	0.19**	(2.33)	1.51	0.24***	(4.43)
8	1.12	0.03	(0.31)	1.46	0.16**	(2.57)
9	1.07	-0.08	(-0.80)	1.41	0.09	(1.17)
High IV_FF3FM	0.29	-0.97***	(-8.82)	0.95	-0.49***	(-3.35)
H-L	-0.58*	-0.92***		-0.07	-0.42**	
(<i>t</i> -value)	(-1.76)	(-6.43)		(-0.20)	(-2.51)	
Panel B Portfolio return and Alpha of five-factor model during 01/1990-12/2014						
Low IV_FF3FM	0.96	-0.01	(-0.13)	1.10	0.21***	(2.59)
2	0.96	-0.03	(-0.34)	1.12	0.05	(0.69)
3	1.13	0.11	(1.26)	1.22	0.09	(1.36)
4	0.92	-0.17*	(-1.85)	1.19	0.02	(0.23)
5	0.96	-0.14	(-1.26)	1.27	0.09	(1.23)
6	0.83	-0.21*	(-1.82)	1.22	0.07	(1.01)
7	1.07	0.02	(0.12)	1.33	0.21**	(2.48)
8	1.00	-0.01	(-0.11)	1.33	0.22**	(2.07)
9	1.13	0.17	(1.13)	1.34	0.35***	(2.73)
High IV_FF3FM	0.49	-0.21	(-1.15)	1.26	0.61**	(2.32)
H-L	-0.47	-0.20		0.16	0.40	
(<i>t</i> -value)	(-0.97)	(-0.84)		(0.32)	(1.53)	

shrinks to -0.97 ($t = -8.82$) in **Table 5**. A similar pattern holds when portfolios are formed by equal-weight. All of these imply that the five-factor model has a good power in explaining anomalies. Therefore, we will emphatically focus on the five-factor model in the remainder.

4. Who Can Explain the Idiosyncratic Volatility Puzzle?

4.1. The Role of Infrequent Trading in Explaining IV Puzzle

In calculating a daily return, we need the closing price for previous day and current day. However, the information of closing prices can be missing due to no trading. In view of this case, CRSP generally uses the average of bid and ask prices to calculate the daily return. For zero trading volume stocks with high expected returns, if we use the daily returns of CRSP to estimate idiosyncratic volatility, the series of daily returns will tend to be smooth and the measure of idiosyncratic volatility will have a lower estimation which has been confirmed in the Section 2.3. Ultimately, it will have a real impact on the relationship between idiosyncratic volatility and returns. Therefore, we will try to use the rearranged CRSP daily return after considering the influence of infrequent trading and a five-factor model to estimate the idiosyncratic volatility (Here, we call it as IV_FF5FM_Infre) and test the idiosyncratic volatility puzzle during 07/1963-12/1989 again. **Table 6** reports the average return and Alpha of portfolios sorted by IV_FF5FM_Infre during 07/1963-12/1989. Whether using CRSP breakpoints or NYSE breakpoints, the return of H-L portfolio is significantly negative (-1.25, -0.58) for value-weight portfolios. Similarly, the difference of intercept between the highest and lowest portfolio is also significantly negative (-1.67, -0.92). The result suggests that the negative relationship between idiosyncratic volatility and returns is not alleviated when infrequent trading is considered.

4.2. The Role of Short-Time Return Reversals in Explaining IV Puzzle

The omission of the previous month's stock returns can lead to a negatively biased estimate of the relation [11]. As a consequence, we construct a model following the approach in Huang *et al.* [11], which is similar to Fama and Macbeth [35] and Fama and French [31], with the exceptions that we include realized idiosyncratic volatility and the prior month's individual stock returns. We test whether the data mining issue of idiosyncratic volatility puzzle can be explained by short-term return reversals. We run the following cross-section regression:

$$R_{i,t+1} - R_{f,t+1} = \alpha + \beta_1 IV_{FF5FM} + \beta_2 R_{i,t} + \beta_3 \text{Size} + \beta_4 \text{B/M} + \beta_5 \text{OP} + \beta_6 \text{AGR} + \varepsilon_{i,t+1} \quad (5)$$

where $R_{i,t+1}$ is stock i 's return in month $t+1$, and $R_{f,t+1}$ is the one-month Treasury bill rate in month $t+1$. The first element is a constant in the right-hand side of equation. IV_{FF5FM} is realized idiosyncratic volatility, given as the standard deviation of the residual from a five-factor model. We include other variables known to explain the cross-sectional variation in stock returns, such as the previous month's return $R_{i,t}$, Size, B/M, OP and AGR. Size is the natural logarithm of market capitalization and B/M is the natural logarithm of book-to-market in month t . OP is the book-equity-deflated operating profitability and AGR is the total asset growth rate. $\varepsilon_{i,t+1}$ is the model residual.

Table 6. Average return and Alpha of portfolios sorted by IV_FF5FM_Infre during 07/1963-12/1989. This table reports the average returns and Alpha of ten portfolios sorted by idiosyncratic volatility relative to the Fama and French [31] model. Portfolios formed every month are based on idiosyncratic volatility measured by IV_FF5FM_Infre . IV_FF5FM_Infre is rearranged CRSP daily return after considering the influence of infrequent trading and based on the five-factor model. Portfolio Low (High) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. Returns are measured in monthly percentages. We consider the average monthly return of value-weighted and equal-weighted portfolios, where the weights are based upon market capitalization at the end of month. Panel A reports the results of decile portfolios using CRSP breakpoints during July 1963 to December 1989. Panel B reports the results of decile portfolios using NYSE breakpoints during July 1963 to December 1989. Alpha is the intercept of the Fama-French five-factor model. The row of H-L refers to the difference between highest and lowest idiosyncratic volatility portfolio. The Newey and West [32] t -statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Value-weight			Equal-weight		
	Return	Alpha	(t -value)	Return	Alpha	(t -value)
Panel A Portfolio return and Alpha of five-factor model during 07/1963-12/1989 (CRSP breakpoints)						
Low IV_FF3FM	0.88	-0.06	(-0.74)	1.13	0.03	(0.35)
2	0.88	-0.06	(-0.99)	1.28	0.16**	(2.39)
3	1.07	0.15***	(2.68)	1.43	0.30***	(5.01)
4	1.14	0.17***	(2.78)	1.54	0.34***	(6.23)
5	1.12	0.12	(1.57)	1.49	0.28***	(5.29)
6	1.14	0.11	(1.22)	0.41.54	0.29***	(4.98)
7	0.99	-0.12	(-1.20)	1.54	0.15**	(2.00)
8	0.90	-0.19	(-1.61)	1.45	-0.03	(-0.33)
9	0.36	-0.84***	(-6.04)	1.28	-0.37***	(-2.87)
High IV_FF3FM	-0.37	-1.73***	(-9.73)	1.02	-0.83***	(-3.81)
H-L	-1.25***	-1.67***		0.60	-0.86***	
(t -value)	(-3.17)	(-7.96)		(-1.26)	(-3.55)	
Panel B Portfolio return and Alpha of five-factor model during 07/1963-12/1989(NYSE breakpoints)						
Low IV_FF3FM	0.87	-0.04	(-0.68)	1.11	0.00	(0.05)
2	1.00	0.07	(1.12)	1.29	0.14**	(2.13)
3	1.07	0.13**	(2.07)	1.45	0.31***	(5.02)
4	1.10	0.17***	(2.63)	1.54	0.31***	(5.45)
5	1.14	0.11	(1.56)	1.50	0.26***	(4.46)
6	1.16	0.18**	(2.25)	1.54	0.31***	(5.80)
7	1.22	0.18**	(2.15)	1.55	0.26***	(5.02)
8	1.16	0.06	(0.62)	1.55	0.27***	(4.50)
9	1.05	-0.09	(-0.92)	1.44	0.11	(1.56)
High IV_FF3FM	0.29	-0.96***	(-8.55)	0.81	-0.57***	(-3.94)
H-L	-0.58*	-0.92***		-0.30	-0.57***	
(t -value)	(-1.74)	(-6.29)		(-0.85)	(-3.28)	

We conduct cross-sectional regressions with Equation (5) each month and then report the time-series averages of estimated coefficients in Panel A of **Table 7**. To make a better comparison for different sample periods, we repeat the

Table 7. Relation between idiosyncratic risk and expected return: cross-section evidence. This table reports the average coefficients in the Fama and MacBeth cross-sectional regressions. Panel A reports the results from Equation (5) and corresponds to all NYSE/AMEX/NASDAQ individual stocks over the period from 07/1963 to 12/1989. Panel B reports the results from Equation (5) during 01/1990-12/2014. However, we only consider IV_FF3FM, Size and B/M during 07/1926-06/1963 in Panel C. IV_{FF5FM} is realized idiosyncratic volatility, given as the standard deviation of the residual from a five-factor model. IV_{FF3FM} is realized idiosyncratic volatility, given as the standard deviation of the residual from a three-factor model. $R_{i,t}$ is the individual stock return in month t . $Size$ is the natural logarithm of market capitalization. B/M is the natural logarithm of book-to-market in month t . OP is the book-equity-deflated operating profitability. AGR is the total asset growth rate. The Newey and West [32] t -statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

Panel A The outcome of cross-section regression during 07/1963-12/1989						
α	IV_{FF5FM}	$R_{i,t}$	Size	B/M	OP	AGR
0.0151*** (5.34)	-0.1729** (-2.29)					
0.0119*** (2.95)		-0.0576*** -10.10				
0.0140*** (4.83)	-0.0940 (-1.07)	-0.0642*** (-10.86)				
0.0257*** (5.69)	-0.2344*** (-4.22)		-0.0019*** (-3.38)	0.0025** (2.53)	0.0004** (2.27)	-0.0001*** (-6.19)
0.0220*** (4.83)	-0.1212 (-0.86)	-0.0772*** (-11.99)	-0.0014** (-2.45)	0.0014 (1.45)	0.0004** (2.22)	-0.0001*** (-6.92)
Panel B The outcome of cross-section regression during 01/1990-12/2014						
α	IV_{FF5FM}	$R_{i,t}$	Size	B/M	OP	AGR
0.0121*** (4.14)	-0.0293 (-0.52)					
0.0229*** (5.46)	-0.0525 (-0.92)		-0.0014*** (-3.24)	0.0035** (2.88)	0.0000 (1.18)	-0.0001*** (-8.47)
Panel C The outcome of cross-section regression during 07/1926-06/1963						
α	IV_{FF5FM}	$R_{i,t}$	Size	B/M		
0.0122*** (3.28)	-0.0279 (-0.36)					
0.0297*** (5.03)	-0.0504 (-0.68)		-0.0042*** (-4.33)	0.0019** (2.01)		

cross-sectional regression with Equation (5) in Panel B of **Table 7** for the period 01/1990-12/2014, which has been proved to be free of idiosyncratic volatility anomaly in Section 3, so we needn't to worry about the effect of prior monthly returns. Moreover, we also use IV_{FF3FM} , Size and B/M as regressor during 07/1926-06/1963 in Panel C. Consistent with the results of the foregoing portfolio analysis, idiosyncratic volatility anomaly only exists in the period 07/1963-12/1989. Specifically, the coefficient on idiosyncratic volatility is significant and negative at the 5% level (-0.1729 , $t = -2.29$) when only IV_{FF5FM} is included in the model during 07/1963-12/1989. However, the coefficient on idiosyncratic volatility is insignificant (-0.0293 , -0.0279) during the period

Table 8. Average return and Alpha of portfolios sorted by IV_FF5FM and based on NYSE breakpoints during 07/1963-12/1989 and 01/1990-12/2014. This table reports the average returns and Alpha of ten portfolios sorted by idiosyncratic volatility relative to the Fama and French [31] model. Portfolios formed every month are based on idiosyncratic volatility computed using CRSP original daily return over the previous month. Portfolio Low (High) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. Returns are measured in monthly percentages. We consider the average monthly return of value-weighted and equal-weighted portfolios, where the weights are based upon market capitalization at the end of month. Panel A reports the results of decile portfolios using NYSE breakpoints during July 1963 to December 1989. Panel B reports the results of decile portfolios using NYSE breakpoints during January 1990 to December 2014. Alpha is the intercept of the Fama-French five-factor model. The row of H-L refers to the difference between highest and lowest idiosyncratic volatility portfolio. The Newey and West [32] *t*-statistics are reported in parentheses. *, **, *** denote statistical significance at the 10%, 5% and 1% levels, respectively.

	Value-weight			Equal-weight		
	Return	Alpha	(<i>t</i> -value)	Return	Alpha	(<i>t</i> -value)
Panel A Portfolio return and Alpha of five-factor model during 07/1963-12/1989 (NYSE breakpoints)						
Low IV_FF3FM	0.87	-0.05	(-0.70)	1.27	0.14*	(1.87)
2	1.00	0.08	(1.21)	1.27	0.11*	(1.75)
3	1.00	0.04	(0.67)	1.33	0.16***	(2.62)
4	1.12	0.20***	(3.16)	1.41	0.20***	(3.33)
5	1.15	0.18**	(2.45)	1.41 0.4144	0.19***	(3.38)
6	1.13	0.09	(1.15)	1.44	0.20***	(3.66)
7	1.15	0.15*	(1.83)	1.41	0.15***	(2.82)
8	1.07	-0.02	(-0.20)	1.38	0.09	(1.39)
9	1.06	-0.01	(-0.09)	1.25	-0.09	(-1.09)
High IV_FF3FM	0.59	-0.56***	(-4.88)	1.05	-0.36**	(-2.45)
H-L	-0.28	-0.51***		-0.22	-0.50**	
(<i>t</i> -value)	(-0.83)	(-3.45)		(-0.61)	(-3.10)	
Panel B Portfolio return and Alpha of five-factor model during 01/1990-12/2014(NYSE breakpoints)						
Low IV_FF3FM	0.91	0.02	(0.22)	1.14	0.25***	(2.93)
2	1.00	-0.05	(-0.62)	1.15	0.07	(0.99)
3	0.90	-0.12	(-1.46)	1.22	0.10	(1.53)
4	1.03	-0.00	(-0.02)	1.21	0.06	(0.91)
5	1.04	-0.10	(-0.95)	1.31	0.13*	(1.82)
6	0.92	-0.10	(-0.93)	1.25	0.10	(1.33)
7	1.11	0.15	(1.32)	1.32	0.23***	(2.69)
8	0.88	-0.06	(-0.51)	1.28	0.21*	(2.11)
9	0.98	0.10	(0.68)	1.31	0.35**	(2.56)
High IV_FF3FM	0.70	-0.12	(-0.65)	1.24	0.56**	(2.12)
H-L	-0.21	-0.14		0.10	0.31	
(<i>t</i> -value)	(-0.47)	(-0.60)		(0.20)	(1.17)	

01/1990-12/2014 and 07/1926-06/1963. This again supports our conjecture that the idiosyncratic volatility puzzle is a result of data snooping bias. The third row in Panel A of **Table 7** indicates that the coefficient on idiosyncratic volatility is not significant (-0.0940 , $t = -1.07$) once we control for return reversals. We also

control for several explanatory variables that might be related to past returns or idiosyncratic volatility in our cross-sectional regression, such as Size, B/M, OP and AGR. Interestingly, the coefficient on idiosyncratic volatility is still significant negative (-0.2344 , $t = -4.22$) when ignore the prior monthly returns. The coefficient becomes insignificant (-0.1212 , $t = -0.86$) once we include stock returns in the previous month. The result confirms the critical role of short-term return reversals in explaining idiosyncratic volatility puzzle.

To mitigate the effect of short-term return reversals, we leave a one-month gap between the portfolio formation date and holding period. Then we form ten value-weighted and equal-weighted portfolios every month by sorting all stocks with IV_FF5FM, which is a proxy for idiosyncratic risk given as the standard deviation of the residual from a five-factor model. **Table 8** reports the average return and Alpha of portfolios sorted by IV_FF5FM and based on NYSE breakpoints during 07/1963-12/1989 and 01/1990-12/2014. We find that idiosyncratic risk is no longer negatively related to expected returns (*i.e.* the idiosyncratic volatility anomaly disappears) when the short-term return reversals are considered during 07/1963-12/1989 in Panel A. Specifically, the H-L portfolio return is negative but insignificant (-0.28 , $t = -0.83$) for value-weighted portfolios. The same is true for equal-weighted portfolios (-0.22 , $t = -0.61$). Although the difference of Alpha between the highest and lowest portfolio is still significantly negative (-0.51 , $t = -3.45$), the significance compared with **Table 5** (-0.92 , $t = -6.43$) is greatly diminished. As previously found, the negative correlation between idiosyncratic volatility and expected stock returns is not significant during 01/1990-12/2014 in Panel B. All the results suggest that the idiosyncratic volatility puzzle is driven by the force of short-term return reversals and is a data snooping bias.

5. Conclusion

In this paper, we use data with a broader time horizon to examine some possible explanations for the idiosyncratic volatility puzzle. Our results show that the well-known idiosyncratic volatility puzzle is driven by data snooping bias. Specifically, we find that the puzzle is significant only for the period 07/1963-12/1989 and is not true for any other period. Different sorting breakpoints and measure of idiosyncratic volatility do not still change our results. Motivated by the standpoint in Liu [22] and Huang *et al.* [11], we attempt to solve the time-specific anomaly from the perspective of infrequent trading and short-term return reversals. We conclude that idiosyncratic volatility is no longer negatively related to subsequent stock returns (*i.e.* the idiosyncratic volatility anomaly disappears) when the short-term return reversals are considered during the prominent puzzle period 07/1963-12/1989. As a byproduct, we find that a five-factor model proposed by Fama and French [31] indeed performs better in the interpretation of anomalies. This is reflected in the result that the intercept term in **Table 5** has a massive improvement compared with **Table 4**. Overall, the

data snooping bias is an important feature to consider in empirical research and our study provides direct evidence that the relation between IV and expected stock returns is not in contradiction to theory.

Funding

This work is supported by the Fund for Shanxi “1331 Project” Key Innovation Research Team [grant number 1331KIRT], the National Natural Science Foundation of China [grant number 71371113] and China Scholarship Council [grant number 202008140168].

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Merton, R.C. (1987) A Simple Model of Capital Market Equilibrium with Incomplete Information. *The Journal of Finance*, **42**, 483-510. <https://doi.org/10.1111/j.1540-6261.1987.tb04565.x>
- [2] Malkiel, B.G. and Xu, Y. (2002) Idiosyncratic Risk and Security Returns. Working paper, University of Texas at Dallas, Richardson.
- [3] Spiegel, M.I. and Wang, X. (2005) Cross-Sectional Variation in Stock Returns: Liquidity and Idiosyncratic Risk. Working Paper, Yale University, New Haven.
- [4] Fu, F. (2009) Idiosyncratic Risk and the Cross-Section of Expected Stock Returns. *Journal of Financial Economics*, **91**, 24-37. <https://doi.org/10.1016/j.jfineco.2008.02.003>
- [5] Chua, C.T., Goh, J. and Zhang, Z. (2010) Expected Volatility, Unexpected Volatility, and the Cross-Section of Stock Returns. *Journal of Financial Research*, **33**, 103-123. <https://doi.org/10.1111/j.1475-6803.2010.01264.x>
- [6] Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X. (2006) The Cross-Section of Volatility and Expected Returns. *The Journal of Finance*, **61**, 259-299. <https://doi.org/10.1111/j.1540-6261.2006.00836.x>
- [7] Ang, A., Hodrick, R.J., Xing, Y. and Zhang, X. (2009) High Idiosyncratic Volatility and Low Returns: International and Further US Evidence. *Journal of Financial Economics*, **91**, 1-23. <https://doi.org/10.1016/j.jfineco.2007.12.005>
- [8] Connolly, R. and Stivers, C. (2003) Momentum and Reversals in Equity-Index Returns during Periods of Abnormal Turnover and Return Dispersion. *The Journal of Finance*, **58**, 1521-1556. <https://doi.org/10.1111/1540-6261.00576>
- [9] Boyer, B., Mitton, T. and Vorkink, K. (2010) Expected Idiosyncratic Skewness. *The Review of Financial Studies*, **23**, 169-202. <https://doi.org/10.1093/rfs/hhp041>
- [10] Chabi-Yo, F. and Yang, J. (2010) Default Risk, Idiosyncratic Coskewness and Equity Returns. Working Paper, Ohio State University, Columbus. <https://doi.org/10.2139/ssrn.1572661>
- [11] Huang, W., Liu, Q., Rhee, S.G. and Zhang, L. (2010) Return Reversals, Idiosyncratic Risk, and Expected Returns. *The Review of Financial Studies*, **23**, 147-168. <https://doi.org/10.1093/rfs/hhp015>
- [12] Bali, T.G. and Cakici, N. (2008) Idiosyncratic Volatility and the Cross Section of

- Expected Returns. *Journal of Financial and Quantitative Analysis*, **43**, 29-58.
<https://doi.org/10.1017/S002210900000274X>
- [13] Han, Y. and Lesmond, D. (2011) Liquidity Biases and the Pricing of Cross-Sectional Idiosyncratic Volatility. *The Review of Financial Studies*, **24**, 1590-1629.
<https://doi.org/10.1093/rfs/hhq140>
- [14] Jiang, G.J., Xu, D. and Yao, T. (2009) The Information Content of Idiosyncratic Volatility. *Journal of Financial and Quantitative Analysis*, **44**, 1-28.
<https://doi.org/10.1017/S0022109009090073>
- [15] Wong, P. (2011) Earnings Shocks and the Idiosyncratic Volatility Anomaly in the Cross-Section of Stock Returns. Working Paper, Ohio State University, Columbus.
- [16] Bali, T.G., Cakici, N. and Whitelaw, R.F. (2011) Maxing out: Stocks as Lotteries and the Cross-Section of Expected Returns. *Journal of Financial Economics*, **99**, 427-446.
<https://doi.org/10.1016/j.jfineco.2010.08.014>
- [17] Chen, Z. and Petkova, R. (2012) Does Idiosyncratic Volatility Proxy for Risk Exposure? *The Review of Financial Studies*, **25**, 2745-2787.
<https://doi.org/10.1093/rfs/hhs084>
- [18] Han, B. and Kumar, A. (2013) Speculative Retail Trading and Asset Prices. *Journal of Financial and Quantitative Analysis*, **48**, 377-404.
<https://doi.org/10.1017/S0022109013000100>
- [19] Barberis, N. and Huang, M. (2008) Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *American Economic Review*, **98**, 2066-2100.
<https://doi.org/10.1257/aer.98.5.2066>
- [20] Hou, K. and Loh, R.K. (2016) Have We Solved the Idiosyncratic Volatility Puzzle? *Journal of Financial Economics*, **121**, 167-194.
<https://doi.org/10.1016/j.jfineco.2016.02.013>
- [21] Johnson, T.C. (2004) Forecast Dispersion and the Cross Section of Expected Returns. *The Journal of Finance*, **59**, 1957-1978.
<https://doi.org/10.1111/j.1540-6261.2004.00688.x>
- [22] Liu, W. (2006) A Liquidity-Augmented Capital Asset Pricing Model. *Journal of Financial Economics*, **82**, 631-671. <https://doi.org/10.1016/j.jfineco.2005.10.001>
- [23] Fisher, L. (1966) Some New Stock-Market Indexes. *The Journal of Business*, **39**, 191-225. <https://doi.org/10.1086/294848>
- [24] Dimson, E. (1979) Risk Measurement When Shares Are Subject to Infrequent Trading. *Journal of Financial Economics*, **7**, 197-226.
[https://doi.org/10.1016/0304-405X\(79\)90013-8](https://doi.org/10.1016/0304-405X(79)90013-8)
- [25] Lo, A. and MacKinlay, A.C. (1990) An Econometric Analysis of Infrequent Trading. *Journal of Econometrics*, **45**, 181-211.
[https://doi.org/10.1016/0304-4076\(90\)90098-E](https://doi.org/10.1016/0304-4076(90)90098-E)
- [26] Al Khazali, O. (2011) Does Infrequent Trading Make a Difference on Stock Market Efficiency? Evidence from the Gulf Cooperation Council (GCC) Countries. *Studies in Economics and Finance*, **28**, 96-110. <https://doi.org/10.1108/10867371111137102>
- [27] Davis, J.L., Fama, E.F. and French, K.R. (2000) Characteristics, Covariances, and Average Returns: 1929 to 1997. *The Journal of Finance*, **55**, 389-406.
<https://doi.org/10.1111/0022-1082.00209>
- [28] Fama, E.F. and French, K.R. (1993) Common Risk Factors in the Returns on Stocks and Bonds. *Journal of financial economics*, **33**, 3-56.
[https://doi.org/10.1016/0304-405X\(93\)90023-5](https://doi.org/10.1016/0304-405X(93)90023-5)
- [29] Fama, E.F. and French, K.R. (2016) Dissecting Anomalies with a Five-Factor Model.

- The Review of Financial Studies*, **29**, 69-103. <https://doi.org/10.1093/rfs/hhv043>
- [30] Shumway, T. (1997) The Delisting Bias in CRSP Data. *The Journal of Finance*, **52**, 327-340. <https://doi.org/10.1111/j.1540-6261.1997.tb03818.x>
- [31] Fama, E.F. and French, K.R. (2015) A Five-Factor Asset Pricing Model. *Journal of Financial Economics*, **116**, 1-22. <https://doi.org/10.1016/j.jfineco.2014.10.010>
- [32] Newey, W. and West, K. (1987) A Simple Positive Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica*, **55**, 703-705. <https://doi.org/10.2307/1913610>
- [33] Fama, E.F. and French, K.R. (1992) The Cross-Section of Expected Stock Returns. *Journal of Finance*, **47**, 427-465. <https://doi.org/10.1111/j.1540-6261.1992.tb04398.x>
- [34] Fama, E.F. and French, K.R. (2017) International Tests of a Five-Factor Asset Pricing Model. *Journal of Financial Economics*, **123**, 441-463. <https://doi.org/10.1016/j.jfineco.2016.11.004>
- [35] Fama, E.F. and MacBeth, J.D. (1973) Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, **81**, 607-636. <https://doi.org/10.1086/260061>