ERASMUS UNIVERSITY ROTTERDAM
Erasmus School of Economics

Master Thesis Financial Economics

Decomposing the idiosyncratic volatility puzzle

An examination of the existing explanations and an investigation into the difference through time and between a developed and an emerging stock market.

Abstract

This study investigates and decomposes the idiosyncratic volatility puzzle, making use of a portfolio strategy proposed by Ang et al. (2006) and a decomposition methodology initiated by Hou and Loh (2016). The study shows that the puzzle is still significantly occurring nowadays and tends to differ in occurrence through time and between developed and emerging stock markets. Furthermore, the study finds that approximately 35-45% of the puzzle can be explained by the investigated candidate explanations, where explanations based on market frictions, with bid-ask spread being the largest contributor, contribute the most in explaining the puzzle. Lastly, the study shows that the explanatory power of the explanations for the puzzle tends to differ throughout time and between a developed and an emerging stock market, suggesting that the IVOL puzzle is most likely to be explained by different explanations for each dimension and that an ultimate explanation for the whole IVOL puzzle most likely tends to be dimensionally bound.

Name student: Brian de Bruijn
Student ID number: 545277

Supervisor: dr. Sjoerd van den Hauwe
Second assessor: dr. Philippe Versijp

Date final version: 31-07-2023

Table of contents

1. Introduction

In today's complex and dynamic financial markets, the investigation of several factors that can predict asset prices has become a subject of significant interest among researchers and investors. One of these factors that has gained substantial attention in the recent decades is the volatility of a security, as a measure of risk. Traditional asset pricing models like the CAPM, introduced by Sharpe (1964) and Lintner (1965), assume that only the systematic risk of a security should be priced into the price of a security since idiosyncratic risk can be diversified away. However, additional literature has shown that the idiosyncratic risk associated with stocks has a positive correlation with their expected returns. For instance, studies conducted by Merton (1987) and Xu and Malkiel (2004) demonstrate that investors cannot always maintain an optimally diversified portfolio and therefore demand a compensation for securities' idiosyncratic risk. Therefore, it is predicted that idiosyncratic risk should either not be related with expected returns or should be positively related.

However, contrary to the aforementioned perspective, Ang et al. (2006) present in their highly influential study that idiosyncratic volatility, which serves as a proxy for idiosyncratic risk, is negatively related with subsequent stock returns, from which they conclude that their results on idiosyncratic volatility present a substantial puzzle, as it could not be explained why taking more risk would lead to lower subsequent returns. Several papers have been published trying to explain the idiosyncratic volatility (IVOL) puzzle, building on the research of Ang et al. (2006). Some of these potential explanations, however, are based on quite different economic mechanisms and economic factors. Therefore, Hou and Loh (2016) introduced a method to assess the explanatory power of the individual explanations for the IVOL puzzle and to investigate the extent to which different explanations can explain the puzzle in relation to the other explanations. Their study reveals that the majority of the explanations only accounts for a small fraction of the puzzle. Furthermore, the study indicates that combining multiple explanations cannot explain the entire puzzle either.

However, since the findings discussed by Hou and Loh (2016) might have become outdated, the main interest of this research is to evaluate to what extent the several candidate explanations can explain the puzzle with current data and explanations published after Hou and Loh's paper and what fraction of the puzzle is still unexplained. In addition, this study investigates the possible differences in the explanatory power of the different candidate variables and the occurrence of the IVOL puzzle through time and between a developed and an emerging stock market. Possible differences that can be found, may provide new insights into solving the IVOL puzzle. In order to investigate these perspectives, this study examines the following research question: *To what extent do the various candidate explanations for the idiosyncratic volatility puzzle explain the idiosyncratic volatility puzzle nowadays*?

To answer the research question, the study firstly uses a 1/0/1 portfolio strategy proposed by Ang et al. (2006), where stocks are sorted on their one-month historical IVOL estimate and held for one month, in order to investigate if the IVOL puzzle still occurs in the examined U.S. sample from 1982-2022. The IVOL estimate is computed as the monthly standard deviation of the residuals from a regression of daily stock returns, in excess of the risk-free rate, on the Fama and French (1993) three factors. A significant return on a portfolio, long in stock with a high historical IVOL estimate and short in stock with a low historical IVOL estimate, points to a relationship between IVOL and subsequent stock returns. To examine a possible difference in occurrence of the puzzle through time and between a developed and an emerging stock market, the 1/0/1 portfolio strategy is also applied to several sub-samples. The U.S. will be considered as a developed stock market and China as an emerging stock market, where the study uses a Chinese sample from 2000-2021.

Next, the decomposition methodology proposed by Hou and Loh (2016) is used to quantify the fraction of the puzzle that each candidate explanation from the literature can explain on their own and the fraction that can be explained after controlling for the other possible explanations. To investigate a possible difference in the explanatory power of the different candidate variables through time and between a developed and an emerging stock market, the decomposition methodology is also applied to several sub-samples. The difference in explanatory power of the variables between the different sub-samples is determined using a Z-test initiated by Clogg et al. (1995).

The investigated explanations in this research are sorted into three groups. The first group contains explanations that tend to explain the IVOL puzzle using proxies for the lottery preferences of investors, such as the skewness, co-skewness, expected idiosyncratic skewness and the maximum daily return of a stock. The second group of explanations relate the puzzle to different types of market frictions, such as one-month return reversals, Amhihud’s illiquidity measure, bid-ask spread and zero return proportion. Lastly, the explanations that cannot be classified under lottery preferences of investors and market frictions, as analyst forecast uncertainty, unexpected earning shocks and growth option, are included in the third group.

The results of the used 1/0/1 strategy show that the IVOL puzzle still tends to occur among U.S. stocks. The study finds a significant average negative return of 1.28% - 1.43% per month on the long-short portfolio, sorted by historical IVOL. However, the results show that a subset of small firms with high IVOL estimates are the driving force behind the IVOL puzzle. In addition, the subsample analysis shows that the occurrence of the IVOL puzzle appears to be time-varying and to differ between developed and emerging stock markets.

The decomposition analysis of Hou and Loh (2016) shows that most of the existing explanations only explain less than 10% of the IVOL puzzle on their own in the U.S. sample. The explanation based on the maximum daily return of a stock proposed by Bali et al. (2011) shows the most promising result and explains 81.91% of the IVOL puzzle. However, due to the high correlation with the IVOL estimate (0.93), there is a high probability that the maximum daily return of a stock is just a proxy for IVOL. When all explanations for the puzzle are examined simultaneously, the results show that explanations related to market frictions contribute the most in explaining the puzzle with 18-21%, followed by the other explanations with 10-14% and the explanations based on the lottery preferences of investors with 7-10%. However, 55-65% of the puzzle remains unexplained.

Moreover, when the puzzle is decomposed for sub-samples of stocks, where the IVOL puzzle tends to occur more strongly, the study finds that the examined explanation can explain approximately 53% of the puzzle among these sub-samples. Lastly, the study finds evidence that the explanatory power of the candidate variables tends to differ through time and between a developed and an emerging stock market, indicating that the IVOL puzzle is most likely to be explained by different explanations for each dimension and an ultimate explanation that can explain the whole IVOL puzzle most likely tends to be dimensionally bound.

In a robustness test, the study repeats the 1/0/1 portfolio strategy indicated by Ang et al. (2006) using IVOL estimates relative to the Carhart (1997) 4-factor model, the Fama and French (2015) 5-factor model and the Fama and French (2018) 6-factor model, to determine whether the puzzle still occurs when IVOL is estimated relative to asset pricing models other than the Fama and French (1993) three factors. The results show the IVOL puzzle still tends to occur among U.S. and Chinese stocks, regardless of which of the investigated asset pricing models is used to determine the IVOL estimate.

This study contributes in several ways to the existing literature among the investigation of the IVOL puzzle. The study is the first to decompose the IVOL puzzle among Chinese stocks and is one of the first studies to investigate the explanatory power of the several candidate explanations between different time periods and between developed and emerging stock markets, which has not been done in the existing literature. Furthermore, the study examines the IVOL puzzle for a more recent time frame and uses an explanation proposed after Hou and Loh (2016) published their paper. Traders and researchers could take the findings of this study into account for the future, where other researchers can build forward on the new insights to possibly find the ultimate explanations for the puzzle.

The remainder of the study continues as follows. Section two discusses the literature concerning the IVOL puzzle and the several candidate explanations, from which the different hypotheses for this research have been formulated. Section three describes the methodology and the construction of the several candidate explanations. Section four discusses how the data is obtained and what modifications have been made to the data. Section five evaluates the results of the study and elaborates on the formulated hypotheses. Section six performs a robustness check and Section 7 concludes and discusses the limitations of the study and the avenues for future research.

2. Literature Review

2.1 Definition of the idiosyncratic volatility puzzle

The occurrence of the so-called ‘anomalies’ have been thoroughly investigated in the financial literature. The term anomaly refers to studies that find contradictory evidence against the findings and correctness of some of the traditional economic theories, which cannot be explained by these traditional economic theories (Frankfurter and McGoun, 2001). One of these traditional economic theories is the ‘Capital Asset Pricing Model’ (CAPM), introduced by Sharpe (1964) and Lintner (1965). The model states that expected stock returns are a linear function of the riskiness of a stock (beta) and the market risk premium. Beta is a measure of a stock's degree of risk relative to its market portfolio, computed as the sensitivity of stock return variance to market return variance. A beta above 1 refers to a more volatile/risky stock than the market portfolio. A beta below 1 refers to a less volatile/risky stock than the market portfolio. Therefore, the CAPM predicts that an increase in beta should lead to an increase in the expected returns. Furthermore, the CAPM states that investors should only receive a reward for the so-called ‘systematic risk’ and not for a stock’s so-called ‘idiosyncratic risk’. Systematic risk refers to the risk associated with market shocks in economic activity that cannot be diversified away, as this risk applies to all stocks. Conversely, idiosyncratic risk refers to risk that is correlated to a specific stock, which can be diversified away by investing in multiple stocks. In other words, the CAPM states that only market risk should be priced into stock prices and demand compensation for the risk taken, with the assumption that the equity market is frictionless, and investors hold well-diversified portfolios.

However, traditional economic theories, including the CAPM, received a lot of criticism since several studies found results that contradicted these theories. One of the most important of these studies has been done by Fama and French (1993). They find that the beta of a stock, as a measure of the systematic risk of that stock, is not able to explain all of the variation in stock returns in their research sample. Based on their contradictory findings, they developed a 3-factor model in addition to the CAPM, which also included a factor for book-to-market ratio and company size. Fama and French (1993) argued that these two factors should also be considered as risk factors in addition to a stock's beta in determining a stock's return.

In addition, the fact that idiosyncratic risk would not be compensated also received criticism. Various economic studies suggested that idiosyncratic risk should be positively correlated with expected stock returns since the equity market is not frictionless and complete. For instance, Merton (1987) predicted that investors do not have complete information at their disposal and therefore only invest in stocks of which they know the risk and return characteristics. As a result, investors will hold sub-optimally diversified portfolios and will demand compensation for securities' idiosyncratic risk.

According to Xu and Malkiel (2004), when investors are unable to maintain a diversified portfolio due to exogenous factors, investors will demand compensation and a premium for investing in companies with high idiosyncratic volatility (IVOL), since high IVOL should be seen as a risk factor. This positive relationship between idiosyncratic risk and expected stock returns was obtained in multiple studies such as Barberis and Huang (2001) and Goyal and Santa-Clara (2003). Nevertheless, despite the criticism of the incompleteness of the CAPM and some of the traditional economic theories, the relationship between a stock's risk and expected returns remained positive.

In contrast to the studies mentioned above, Ang et al. (2006) observed that there appears to be a negative link between IVOL and future stock returns for U.S. stocks. They examined IVOL at the firm level and relative to the Fama and French (1993) 3 factor model and found that stocks with high IVOL tend to have low expected returns on average. In their more recent study, Ang et al. (2009) also found this negative relationship between IVOL and subsequent stock returns for stocks in other G7 countries. Moreover, Blitz and van Vliet (2007) noted the same negative relationship between stock volatility and expected returns. The findings of Ang et al. (2006) and Blitz and van Vliet (2007) are puzzling because the theories previously discussed either predict no relationship between IVOL and subsequent stock returns, considering the market is frictionless and investors are well diversified, or predict a positive relationship, assuming that investors aren't able to keep a well-diversified portfolio and demand compensation for this. Ang et al. (2006) provided several checks in their study to identify the robustness of their results and observed that their results are robust after controlling for several variables and different specifications. Additionally, they are unable to find a complete explanation for the observed relationship in practice and concluded that their results on idiosyncratic volatility represent a substantial puzzle.

The idiosyncratic volatility puzzle can therefore be defined as the fact that many traditional economic theories predict either no relationship or a positive relationship between IVOL and subsequent stock returns, but in practice, this relationship appears to be negative and to a substantial extent unexplainable.

2.2 Candidate explanations for the idiosyncratic volatility puzzle and the stronger occurrence of the puzzle

Building on the research of Ang et al. (2006), several papers have been published trying to explain the IVOL puzzle by proposing and testing possible explanations for the puzzle based on their specific theory that links IVOL to the subsequent stock returns. However, some of these candidate explanations are based on quite different economic mechanisms and economic aspects. This section, therefore, provides an overview of the several types of explanations for the puzzle. Furthermore, this section also provides an overview of several studies, which find subsamples of firms where the IVOL puzzle tends to be more prominent and find a difference in the occurrence of the puzzle between different time periods.

2.2.1 Investors' lottery preferences

The first type of explanation for the IVOL puzzle relates to the lottery preferences of investors when buying certain stocks. Barberis and Huang (2008) state that investors will overestimate the small odds of exceptionally high returns and will favor lottery-type stocks. These investors will prefer and buy up positively skewed stocks because of this preference, which will cause these positively skewed stocks to become overpriced and have low future returns.

In addition, Kumar (2009) related this thought process to IVOL. He argues that stocks with higher IVOL are more likely to be seen as lotteries because the level of IVOL could affect idiosyncratic skewness estimates. Investors may think that extreme return observations from the past are more likely to occur again when the volatility of a stock is high. On the other hand, if a stock with a low price and high skewness has a low IVOL, the occurrence of extreme return events may be considered as outliers and there is a smaller chance of the incidents occurring again. Boyer et al. (2010) observed a negative correlation between expected idiosyncratic skewness and stock returns and found that a measure for expected idiosyncratic skewness can help to explain the IVOL puzzle. Chabi-Yo and Yang (2010) showed that the effect of idiosyncratic volatility on stock returns is related to a stock's co-skewness with the market portfolio. Furthermore, Bali et al. (2011) mentioned that stocks with lagged positive extreme returns tend to have low future returns. Their positive extreme returns can therefore be seen as a rough measure of skewness.

The studies mentioned in this section, therefore, share the same line of thought. They conclude that the IVOL puzzle can be explained by the lottery preferences of investors and lottery-type stock characteristics such as skewness. Investors that seek skewness/lottery-like stocks will pay a higher price for stocks with high idiosyncratic volatility mainly because these stocks offer high skewness, which causes these stocks to become overpriced with low future stock returns. (Boyer et al., 2007).

2.2.2 Market frictions

The second type of candidate explanations for the findings of Ang et al. (2006) are associated with frictions in the equity market. Fu (2009) argues that the IVOL puzzle can be largely explained by the return reversal of stocks with high IVOL. He shows that idiosyncratic volatilities are time-varying and states that the results of Ang et al. (2006) are influenced by a subset of small firms with high IVOL. These small stocks tend to realise an extremely high IVOL estimate and abnormally high returns in the same month. However, these returns tend to reverse in the following month. As a result, the returns of high IVOL firms will be abnormally low in the following month. Fu (2009) therefore mentions that the relationship between IVOL and subsequent stock returns is negative, but the relationship between IVOL and realized returns tends to be positive. In addition, he finds that, after controlling for one-month return reversals, the relationship between IVOL and subsequent stock returns is not significantly negative anymore. Fu (2009) states that the lagged estimate of IVOL is not a good proxy for the realized value of IVOL and mentions that the findings of Ang et al. (2006) should not be used to suggest a link between expected returns and idiosyncratic risk.

Further explanations for the IVOL puzzle associated with equity market frictions are based on certain stocks' liquidity. Han and Lesmond (2011) argue that microstructure influences are essential for estimating idiosyncratic volatility and predicting future returns. They mention that the pricing ability of IVOL depends on the underlying liquidity costs, since these costs influence the estimation of the IVOL and these costs are negatively related to future returns. They refer to the bid-ask bounce and trading days with zero returns which influences the estimate for the IVOL. After controlling for liquidity costs as mentioned above and the liquidity measure of Amihud (2002), they show that their measure of IVOL no longer can predict stock prices anymore, suggesting that the liquidity cost component of the IVOL measure tends to drive the results.

The studies mentioned in this section conclude that the negative returns associated with IVOL can be explained by several types of market frictions, which influences the estimate for the IVOL. These market frictions tend to capture a substantial component of the IVOL measure, and after controlling for these frictions the IVOL measure seems to have no return predicting ability anymore.

2.2.3 Other candidate explanations

The third group of candidate explanations for the IVOL puzzle includes explanations that cannot be categorized neatly into the categories of investor lottery preferences or market frictions. Johnson (2004) finds that there is a relation between fundamental uncertainty around a stock, as indicated by the dispersion of analyst forecasts, and future stock returns. He argues that dispersion may proxy for idiosyncratic parameter risk when fundamentals are unobservable. Furthermore, Johnson (2004) states that we observe a negative relationship between uncertainty and subsequent stock returns because a stock is a call option on a levered firm's underlying assets. Therefore, IVOL could proxy for the fundamental uncertainty surrounding a stock, which then should be negatively correlated with subsequent stock returns (Hou and Loh, 2016). George and Huang (2013) supported this conclusion in their paper. They found that among low analyst coverage firms, which are more affected by uncertainty and disagreement, the IVOL puzzle tends to be stronger. They conclude that high IVOL predicts low stock returns because low-coverage firms' optimistic mispricing is caused by uncertainty and disagreement, which is later corrected.

Another type of explanation for the IVOL puzzle was found by Wong (2013). He demonstrated that large and negative pre- and post-formation earnings shocks are associated with stocks with high IVOL, which clarifies why these stocks at first showed high idiosyncratic volatility. He finds that high IVOL stocks consistently experience negative post-formation earnings shocks that cause their realized returns to appear lower than their expected returns. Furthermore, Wong (2013) shows that after controlling for those negative earnings shocks on future stock returns, the IVOL estimate has barely any return prediction ability, implying that earnings shocks are a possible explanation for the puzzle.

Liu et al. (2021) found that the negative association between IVOL and expected returns can be attributed to the growth options of firms. According to them, a corporation can be viewed as a combination of existing assets and growth potential, and IVOL must be decomposed into these components. They argue that a company with more growth options can scale down much more easily than one with more assets in place, indicating that growth options are less risky than assets currently in place. Therefore, they expected that a firm with higher IVOL that can be associated with growth options tends to be less risky and will generate lower future stock returns. In line with their expectations, they observe a negative relation between IVOL related to growth options and future returns.

2.2.4 Stronger occurrence of the puzzle

Apart from several studies that have been published to explain the IVOL puzzle, several studies have also investigated the stronger occurrence of the IVOL puzzle among subsamples of firms. For example, Avramov et al. (2013) argue in their paper that the IVOL puzzle only tends to exist among firms with low credit ratings, which corresponds to companies in financial trouble. They show that the profitability of an investment strategy, where an investor buys low IVOL stocks and short sells high IVOL stocks, is concentrated in the stocks with low credit ratings. The strategy tends to be unprofitable when companies with a rating of BB+ or lower are excluded from the sample. These findings are consistent with the findings of Wong (2013), who argued that negative earnings shocks associated with high IVOL firms are likely to drive these companies into financial distress.

George and Huang (2013) showed in their study several subsamples of firms where the IVOL puzzle seems to occur more strongly. They mention that high-volatility stocks are frequently used for tax-loss selling. This selling pressure has an especially strong impact on penny stocks and their positive January returns hide a strong and persistent negative relationship between future returns and idiosyncratic volatility. Similar to this approach, they find that the negative returns associated with IVOL are stronger for firms with a share price higher than $5 and in non-January months.

Furthermore, as mentioned in section 2.2.3, they find that the puzzle occurs more strongly among stocks with low analyst coverage.

According to Stambaugh, Yu and Yuan (2015), high short-interest stocks, which are more difficult to short-sell, tend to show a strong negative relation between IVOL and future stock returns. They argue that higher IVOL, which results in greater arbitrage risk, allows for more mispricing and that the relationship between IVOL and stock returns is negative for overpriced stocks and positive for underpriced stocks. They show that, among the overvalued stocks, the negative relationship between IVOL and future stock returns is stronger for stocks that are less easy to short.

Johnson (2004) mentioned in his study that firms with more leverage are associated with higher dispersion among analysts and more disagreement among investors. He shows that these firms with more leverage show more fundamental uncertainty surrounding the stock, where IVOL could be a proxy for, and have lower subsequent stock returns.

Brandt et al. (2010) approached the puzzle from another point of view. They tried to investigate the difference in the occurrence of the puzzle between different time periods and find differences through time. They observe a rise in the strength of the puzzle through the 1990s, followed by a significant drop around 2003. They argue that the IVOL puzzle is more of a temporary situation than an actual trend, which can be explained in part by the number of retail investors.

2.3 The explanatory power of the candidate explanations for the idiosyncratic volatility puzzle

As mentioned in section 2.2, many studies have come up with quite different possible explanations for the IVOL puzzle. However, based on the numerous studies, it is difficult to determine which explanation can best explain the puzzle. Since each study uses its specific method to evaluate the possible explanation for the puzzle, Hou and Loh (2016) designed a methodological framework that allowed for the comparison of the different explanations. Their method allows to quantify the fraction of the puzzle that can be explained by each candidate explanations and allows them to determine what fraction of the puzzle can be explained and is still left unexplained. Their study finds that many of the existing explanations, on their own, explain less than 10% of the IVOL puzzle in the US. Together, all existing explanations, for which they control, represent 29-54% of the puzzle in individual US stocks for the period between 1963 and 2012. Although approximately half of the puzzle can be explained by existing theories, they find that most of the explainable part of the puzzle can be explained by explanations based on investors' lottery preferences and market frictions. However, they also conclude that a significant fraction of 46-71% remains unexplained.

Zhong (2018) used the decomposition method of Hou and Loh (2016) to investigate to what extent the IVOL puzzle could be explained for the Australian equity market between 1993 and 2013. He suggests that the potential causes of the IVOL puzzle are due to mispricing in the Australian equity market, as the puzzle is concentrated among the most overpriced stocks, which is consistent with the findings of Stambaugh, Yu and Yuan (2015). However, in contrast to Hou and Loh (2016), Zhong (2018) finds that almost all the explanations he checks for can explain more than 10% of the IVOL puzzle on their own. With all the explanations taken together, Zhong can explain approximately 70% of the puzzle and, like Hou and Loh (2016), finds that the puzzle can largely be explained by explanations based on investors' preference for lottery-like stocks.

Annaert et al. (2022) investigated the IVOL puzzle in Euro area stocks and used data from thirteen different European countries from 1999 to 2019. Based on the explanations they check for, they can explain approximately 30% of the IVOL puzzle with an almost equal contribution coming from explanations based on lottery features and market frictions and conclude that many of the explanations individually can explain less than 10% of the puzzle, which is in line with Hou and Loh (2016). However, also for Euro area stocks, about 70% of the puzzle remains unexplained.

According to the different studies, which have decomposed the IVOL puzzle in various equity markets and time periods, a significant portion of the puzzle remains unexplained, concluding that the leading candidate explanations from the literature do not have the explanatory power to explain the entire puzzle.

2.4 Developed markets versus emerging markets: implications for the idiosyncratic volatility puzzle

As section 2.3 shows, several studies examined the decomposition of the IVOL puzzle for countries with developed stock markets but the decomposition analysis for the puzzle among emerging stock markets is hardly discussed. However, the literature reports significant differences in characteristics between developed and emerging markets, which might influence the occurrence of the IVOL puzzle and the explanatory power of the candidate explanations on the relevant markets.

For example, Kohers et al. (2006) show in their study that emerging markets in general are more risky and more volatile than developed markets. Moreover, they note that emerging markets differ from developed markets in terms of their liquidity risk and the lack of high-quality, large-capitalization stocks. Lesmond (2005) and Bekaert et al. (2007) also mention that emerging markets tend to be less liquid and contain relatively more microcap stocks. Additionally, they demonstrate that emerging markets have considerably more short-selling constraints, making it more difficult to short-sell stocks. Furthermore, according to Li and Wang (2010), emerging stock markets tend to be more volatile and risky compared to developed markets, due to the fact that emerging markets contain relatively more uninformed retail investors than developed markets. As a result, noise traders may start to dominate the market, leading to more volatile stock returns as noise traders trade on unreliable information.

According to the studies mentioned above, there are several differences in characteristics between developed and emerging markets, which may imply differences for the IVOL puzzle on the relevant markets. It may be that the IVOL puzzle for emerging markets is different than for developed markets due to the difference in stock illiquidity, as Han et al. (2023) state that illiquidity causes the level of IVOL. Furthermore, as described in section 2.2.4, the magnitude of the IVOL puzzle tends to increase among stocks with more short-sell constraints (Stambaugh, Yu and Yuan, 2015). Since emerging markets tend to have more short-selling constraints compared to developed markets, this might imply that the IVOL puzzle also tends to be different between the two types of markets. Moreover, it could also be the case that the difference in fractions of retail traders between developed and emerging markets (Li and Wang, 2010) causes differences in the appearance of the IVOL puzzle, as Brandt et al.(2010) argue that rises in the magnitude of the puzzle can be explained by rises in the fraction of retail investors active on the relevant stock market.

Several studies have shown that developed and emerging stock markets tend to vary considerably in specific characteristics, while particularly these characteristics, according to the literature, also tend to influence the occurrence of the IVOL puzzle. Considering these distinct characteristics, it is plausible that the decomposition results that apply to developed markets may not apply to emerging markets and that the occurrence of the IVOL puzzle and the explanatory power of the candidate explanations might differ between the two market types.

2.5 Hypotheses

The occurrence of the IVOL puzzle has been observed in various countries, time periods and subsamples of firms. Whereas traditional economic theories expect no or a positive relationship between IVOL and subsequent returns, Ang et al. (2006), in their influential paper, indicated a negative relationship. Many studies have subsequently investigated why we observe this relationship in practice and showed different explanations for the negative relationship.

Hou and Loh (2016) designed a methodological framework that allowed for the comparison of the different explanations but concluded that most of the puzzle was still unexplained for the U.S. stock market. This conclusion was later also found for the Australian stock market and Euro area stocks (Zhong, 2018; Annaert et al., 2022). Combing these findings, it is not expected that the various candidate explanations for the puzzle to date, are able to explain the whole IVOL puzzle for the U.S. stock market with more recent data. However, the main goal of this study is to investigate to what extent the IVOL puzzle can be explained nowadays based on the various candidate explanations and to examine to what extent the puzzle and explanations differ across different data samples, to get more insight into the puzzle and hopefully get an indication of what the ultimate explanation might be related to.

To answer the main research question, this study tests four hypotheses. First, several papers find that explanations based on investors' lottery preferences and market frictions can largely explain the IVOL puzzle (Hou and Loh, 2016; Zhong, 2018; Annaert et al., 2022). However, these studies look at different time periods and/or countries compared to this study, but there is no clear evidence that there should be a difference in the explanatory power of these types of explanations. Since there is no clear evidence that the explanatory power of the explanations should be different for a different time period and/or country, the first hypothesis has been formulated as follows:

Hypothesis 1: Explanations based on investors' lottery preferences and market frictions can best explain the puzzle.

Secondly, the decomposition studies of the IVOL puzzle (Hou and Loh, 2016; Zhong, 2018; Annaert et al., 2022) show that still a large fraction of the IVOL puzzle is left unexplained. Therefore, it is expected that this study will not be able to explain the whole IVOL puzzle with more recent data, but it is still interesting to investigate what fraction of the IVOL puzzle can be explained by all the candidate explanations to date taken together. However, as it is expected that this study will be unable to explain the whole IVOL puzzle, the second hypothesis has been constructed as follows:

Hypothesis 2: Candidate explanations from the literature are not able to explain the whole IVOL puzzle.

Thirdly, a possible time-varying occurrence of the IVOL puzzle and the explanatory power of the candidate explanations will be examined. For example, Fu (2009) shows that IVOL varies over time and Brandt et al. (2010) find evidence that the occurrence of the IVOL puzzle differs over time. It could be the case that with the occurrence of the IVOL puzzle, which tends to differ over time, also the explanatory power of the candidate explanations possibly will differ over time. Therefore, the third hypothesis states:

Hypothesis 3: The candidate explanations and the occurrence of the IVOL puzzle differ over time.

Lastly, the studies that decompose the IVOL puzzle (Hou and Loh, 2016; Zhong, 2018; Annaert et al., 2022) solely focus on developed stock markets, where, to my knowledge, no research has been done yet that focuses on emerging stock markets. However, as section 2.4 discusses, it is plausible that the occurrence of the IVOL puzzle and the fraction that can be explained by the candidate explanations might differ between developed and emerging markets, since both markets tend to be quite different based on distinct characteristics (Lesmond, 2005; Bekaert et al., 2007; Li and Wang, 2010) and emerging markets tend to be more volatile than developed markets (Kohers et al., 2006). To investigate whether there is a difference between developed and emerging markets, a comparison will be made between the U.S. stock market and the Chinese stock market, where the U.S. stock market will be considered as a developed stock market and the Chinese stock market as an emerging stock market. This was chosen since the U.S. stock market is the largest contributor to the MSCI World Index and the Chinese stock market is the largest contributor to the MSCI Emerging Markets Index. Since it is expected that due to the considerably different characteristics between developed and emerging markets, the IVOL puzzle also might be different between the two types of markets, the fourth hypothesis concerns the following:

Hypothesis 4: The candidate explanations and the occurrence of the IVOL puzzle differ between developed and emerging stock markets.

3. Methodology

3.1 Computation of the idiosyncratic volatility (IVOL) measure

To check whether the IVOL puzzle still exists and to what extent we can explain it, it is important to first establish the idiosyncratic volatility estimate. The computation method of Ang et al. (2006) has been followed to estimate IVOL. They compute IVOL as the standard deviation of the residuals from a regression of daily stock returns in a specific month relative to the Fama and French (1993) 3 factors. To be able to calculate IVOL, the following time-series regression will first be performed:

$$r_t^i - rf_t = a_t^i + \beta_{Mkt}^i Mkt_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \varepsilon_t^i, \tag{1}$$

where $r_t^i - rf_t$ is the daily return on the stock of firm i in excess of the risk-free rate on day t and a_t^i is the constant of the regression. Mkt_t is the market return in excess of the risk free rate on day t, SMB_t is the return difference of small cap company stocks relative to large cap company stocks on day t, HML_t is the difference in return between high book-to-market value company stocks and low book-to-market value company stocks on day t, and ε_t^i is the residual estimate of the regression for firm i on day t. β_{Mkt}^i, β_{SMB}^i and β_{HML}^i are factor loadings on Mkt_t, SMB_t and HML_t, respectively. These factor loadings, used to compute the residual estimate for firm i on day t, are estimated using an expanding monthly window. That is, at the end of the relevant month, the factor loadings are estimated based on daily data from the relevant month plus data from previous months if such data is available to an investor at that time. The factor loadings are thus re-estimated every month based on data available up to time t.

Then, the IVOL estimate for a specific month is defined as the standard deviation of the residuals from equation (1) at the end of the relevant month using the following equation:

$$IVOL_t^i = \sqrt{Var(\varepsilon_t^i)}, \tag{2}$$

where $IVOL_t^i$ is the idiosyncratic volatility estimate for firm i in month t and $Var(\varepsilon_t^i)$ is the variance of the daily residuals available for firm i over month t. A minimum of ten daily observations is required to estimate IVOL, as done by Hou and Loh (2016), to overcome biased estimates of IVOL due to a lack of observations.

3.2 Examining the occurrence of the IVOL puzzle and the occurrence over time

In order to test whether the IVOL puzzle occurs in the sample used for this research, stocks will be sorted into quintile portfolios based on their monthly IVOL estimate of equation (2). The first portfolio will contain stocks with the lowest IVOL estimate and the fifth portfolio will contain stocks with the highest IVOL estimate. These portfolios will be used to create a trading strategy as done by Ang et al.

(2006). They describe their strategy as an L/M/N strategy, where at month t, a historical period of L months is used to estimate their IVOL measure, a waiting period of M months and a holding period of N months. This study will make use of the 1/0/1 strategy, as this strategy is also the main strategy in Ang et al. (2006). This means that at the end of every month t, stocks are divided into quintile portfolios based on their IVOL estimate over the past month and these portfolios are held for 1 month. The portfolios will be rebalanced for each month. However, as an addition to the strategy of Ang et al. (2006), equal-weighted portfolios and value-weighted portfolios will be used in this study.

To check whether there is a negative relationship between IVOL and subsequent stock excess returns, an additional portfolio will be set up that goes long in the stocks with the highest IVOL estimate (portfolio five) and short in the stocks with the lowest IVOL estimate (portfolio one). A significant negative return on this portfolio will suggests a negative relation between IVOL and subsequent returns and will indicate that the IVOL puzzle tends to occur in the sample. The significance of the return of the additional portfolio will be determined by making use of a t-test. In addition, for each quintile and the long-short portfolio, the alpha estimate will be calculated relative to the Fama and French (1993) factors. Alpha shows the average return on a portfolio adjusted for risk factors. Alpha will be estimated as the constant term in a regression of monthly portfolio returns on the associated Fama and French (1993) factors.

The occurrence of the IVOL puzzle will be tested based on a portfolio strategy, as a regression analysis assumes a linear relationship between the IVOL estimate and subsequent stock returns, which of course does not have to be the case as the relationship could also be nonlinear. A portfolio analysis does not make such an assumption about the relationship between the IVOL estimate and subsequent returns and will therefore account for a possible non-linear relationship (Bali, Engle & Murray, 2016).

After the occurrence of the IVOL puzzle over the entire U.S. sample has been tested, the U.S. sample will be divided into several smaller subsamples, to check if the puzzle still occurs through different time periods, in which the puzzle's occurrence may vary due to possible changes in a variety of market conditions. The sample will be divided into ten sub-samples, each containing four years. The subsamples that will be investigated are 1982-1985, 1986-1989, 1990-1993, 1994-1997, 1998-2001, 2002-2005, 2006-2009, 2010-2013, 2014-2017 and 2018-2022 to examine if the puzzle tends to be time-varying. Furthermore, the relevant sub-samples will be investigated because four sub-samples capture a specific crisis, namely the Early 1990s recession (1990-1993), the DotCom bubble around the year 2000 (1998-2001), the financial crisis around the year 2008 (2006-2009) and the Covid-19 crisis around the year 2020 (2018-2022). Therefore, the results of these sub-samples will automatically show how the IVOL puzzle behaves during periods that capture a substantial crisis compared to periods that do not capture a crisis. The occurrence will be determined based on a significant negative t-test of the returns of the long-short portfolio for the relevant subsample.

Lastly, to check for a possible difference in the presence of the IVOL puzzle between developed and emerging markets, the occurrence of the IVOL puzzle in the U.S. stock market and the Chinese stock market will be compared through different time periods. The occurrence of the puzzle on the US stock market and Chinese stock market will be compared for the sub-samples 2000-2003, 2004-2007, 2008-2011, 2012-2015, 2016-2019 and 2020-2021, as little reliable data is available for Chinese stocks before the year 2000 and the CSMAR database does not provide data for the year 2022 for Chinese companies. These sub-samples will be investigated in order to also compare the occurrence of the IVOL puzzle when considering smaller time periods, in which the occurrence of the puzzle may differ due to changes in several market characteristics. The occurrence of the IVOL puzzle will also be determined based on a significant negative t-test of the returns of the long-short portfolio for the relevant subsample.

3.3 Computation of the candidate variables related to investors' lottery preferences

3.3.1 Skewness (Skew)

For this study, several candidate variables have been formulated, which according to previous literature, as discussed in section 2.2, should be a possible explanation for the IVOL puzzle. The first group of candidate variables to be discussed are the candidate explanation variables related to investors' lottery preferences. Barberis and Huang (2008) showed that investors tend to overestimate the small chances of remarkably high returns (lottery-type stocks) and will bid up positively skewed stocks, causing them to become overpriced. Kumar (2009) argued that stocks with higher IVOL are more likely to be seen as lotteries, as the level of IVOL could affect the skewness estimate and assigns the IVOL puzzle to IVOL being related to a stocks' skewness. Therefore, skewness (Skew) will be measured as the skewness of the daily excess returns in month t for a specific firm.

3.3.2 Expected idiosyncratic Skewness (E(IS))

Furthermore, as mentioned in the literature review, Boyer et al. (2010) found that expected idiosyncratic skewness helps explain the IVOL puzzle. They make use of a predictive regression model to obtain a proxy for the expected idiosyncratic skewness and show that expected idiosyncratic skewness is negatively related to subsequent stock returns. The method by Boyer et al. (2010) to estimate expected idiosyncratic skewness is slightly changed for this research. They make use of two rolling windows of sixty months to obtain a proxy for expected idiosyncratic skewness. This means that a minimum of ten years of data is required to generate an estimate for expected idiosyncratic skewness. However, they state that their results are robust for horizons ranging from six months to their base case of sixty months. Therefore, to keep as many observations as possible, expected idiosyncratic skewness will be measured using rolling windows of twelve months.

First, the following regression will be applied:

$$is_t^i = a_t + \beta_t^1 is_{t-12}^i + \beta_t^2 iv_{t-12}^i + \beta_t^3 MOM_{t-12}^i + \beta_t^4 TURN_{t-12}^i + \beta_t^5 SMALL_{t-12}^i + \beta_t^6 MEDIUM_{t-12}^i + \beta_t^7 NASDAQ_{t-12}^i + \beta_t^8 SIC_{t-12}^i + \varepsilon_t^i, \tag{3}$$

where is_t^i is the estimate of historical idiosyncratic skewness of firm i in month t, measured as the skewness of the daily residuals from equation (1) of the past twelve months. is_{t-12}^i is the lagged estimate of is_t^i for twelve months. iv_{t-12}^i is the lagged estimate of historical idiosyncratic volatility for twelve months of firm i in month t, measured as the standard deviation of the daily residuals from equation (1) of the past twelve months. MOM_{t-12}^i and $TURN_{t-12}^i$ are the twelve month lagged measures of momentum and share turnover, defined as discussed in Boyer et al. (2010), for firm i in month t. $SMALL_{t-12}^i$ and $MEDIUM_{t-12}^i$ are twelve month lagged dummy variables for small size and medium size firms, respectively, where companies are divided into three groups of the same size each month for small, medium and large based on their market capitalization. Lastly, $NASDAQ_{t-12}^i$ and SIC_{t-12}^i are twelve month lagged dummy variables for firms that trade on the NASDAQ stock exchange and an industry classification based on a firm's two digit SIC code, respectively.

After equation (3) has been applied, the factor loadings of equation (3) are used to estimate the expected idiosyncratic skewness for each company:

$$E_t\left[is_{t+12}^i\right] = a_t + \beta_t^1 is_t^i + \beta_t^2 iv_t^i + \beta_t^3 MOM_t^i + \beta_t^4 TURN_t^i + \beta_t^5 SMALL_t^i + \beta_t^6 MEDIUM_t^i + \beta_t^7 NASDAQ_t^i + \beta_t^8 SIC_t^i, \tag{4}$$

where $E_t\left[is_{t+12}^i\right]$ is the proxy for expected idiosyncratic skewness 12 months later. The other variables are defined as mentioned above. The candidate variable expected idiosyncratic skewness will be named E(IS) for the remainder of this study. Equations (3) and (4) will be used to examine the E(IS) for common stocks in the U.S.. For Chinese common stocks, equations (3) and (4) will be applied but without a Nasdaq dummy and a company's two digit SIC code will be replaced with their Industry Code B.

3.3.3 Co – Skewness (CoSkew)

In addition to using the raw measure of skewness as a candidate variable, as mentioned in section 3.2.1, a variable based on co-skewness, as indicated by Chabi-Yo and Yang (2010), will also be used. Chabi-Yo and Yang mentioned that the low returns of high IVOL stock can be explained by a stock's co-skewness with the market portfolio. To measure the co-skewness (CoSkew) of a firm, in an equivalent way as Hou and Loh (2016) have done, the following regression will be applied for each month:

$$(r_t^i - rf_t)^2 = a_t^i + \beta_1^i Mkt_t + \varepsilon_t^i, \tag{5}$$

where $(r_t^i - rf_t)^2$ is the estimate of squared daily excess returns of firm i on day t, Mkt_t is the market return in excess of the risk free rate on day t and β_1^i is the factor loading on Mkt_t of firm i for a specific month. Co-skewness will be measured as regression coefficient β_1^i.

3.3.4 Maximum return (MaxRet)

As another possible measure of the skewness of a stock, a variable based on maximum return, mentioned by Bali et al. (2011), will also be used. Bali et al. (2011) argued that stock with lagged positive extreme returns is preferred by investors who want to buy lottery-type stocks, resulting in overpricing and low subsequent stock returns. The variable maximum return (MaxRet) will be measured as the highest daily excess stock return in month t for a specific firm.

3.4 Computation of the candidate variables related to market frictions

3.4.1 Return reversal (LagRet)

The second group of candidate variables are the candidate explanations related to market frictions. As mentioned in section 2.2.2, Fu (2009) states that the IVOL puzzle can be explained by the return reversal of stocks with high IVOL in the previous month. As a result, returns reversals (LagRet), like those measured by Hou and Loh (2016), will be measured as the previous month's monthly excess return for a specific firm.

3.4.2 Measures of illiquidity

Several studies have shown that the IVOL puzzle can be explained by the relationship with the illiquidity of a stock. For example, Han and Lesmond (2011) show that after controlling for liquidity costs and Amihud's (2002) liquidity measure, IVOL no longer exhibits pricing power. Therefore, as candidate explanations for the IVOL puzzle, this study will use the three variables that Han and Lesmond (2011) use in their study as proxies for illiquidity.

The first proxy for a stock’s illiquidity will be the Amihud measure (Amihud, 2002). Amihud’s measure will be computed using the following equation:

$$Amihud_t^i = \frac{1}{D_t^i} \sum_{t=1}^{D_t^i} \left| R_{td}^i \right| / \ VOLD_{td}^i, \tag{6}$$

where $Amihud_t^i$ is the estimate of Amihud’s measure for firm i in month t. D_t^i is the number of trading days with sufficient data for firm i in month t. R_{td}^i is the excess stock return of firm i on day d of month t and $VOLD_{td}^i$ is the daily dollar trading volume in millions of firm i on day d of month t. In other words, Amihud's measure of illiquidity (AMIHUD) is defined as the monthly average of the daily absolute excess return divided by the dollar trading volume in millions on that particular day.

The second proxy for a stock's illiquidity will be defined by the bid-ask spread, as Han and Lesmond (2011) state that bid-ask bounces are one of the significant explanations for the IVOL puzzle. The bid-ask spread (SPREAD) will be defined as:

$$Spread_t^i = \frac{1}{D_t^i} \sum_{t=1}^{D_t^i} \frac{(ASK_{td}^i - BID_{td}^i)}{(\frac{1}{2}(ASK_{td}^i + BID_{td}^i))}, \quad (7)$$

where $Spread_t^i$ is the estimate of the bid-ask spread form firm i in month t. D_t^i is the number of trading days with sufficient data for firm i in month t, ASK_{td}^i is the closing ask price for the stock of firm i on day d in month t and BID_{td}^i is the closing bid price, respectively. Spread will therefore be defined as the monthly average of the daily percentage bid-ask spread of the specific stock. The variable Spread will not be established for Chinese firms due to limited data availability.

The third proxy for illiquidity will be the fraction of trading days with a return of zero (ZeroRet) for a specific firm in a specific month. ZeroRet will be computed as:

$$ZeroRet_t^i = \frac{Zero_t^i}{D_t^i}, \quad (8)$$

where $ZeroRet_t^i$ is the fraction of trading days with a return of zero for firm i in month t, $Zero_t^i$ is the amount of trading days with a return of zero in month t for firm i and D_t^i is the number of trading days with sufficient data for firm i in month t.

3.5 Computation of the candidate variables related to other explanations

3.5.1 Analyst forecast uncertainty (Dispersion)

The third group of candidate variables are the candidate explanations related to other factors than investors' lottery preferences and market frictions. As discussed in section 2.2.3, Johnson (2004) states that fundamental uncertainty surrounding a particular stock, measured as analyst forecast dispersion, is negatively associated with subsequent stock returns. George and Huang (2013) demonstrated that high IVOL tends to indicate lower stock returns due to the presence of uncertainty and disagreement, which leads to optimistic mispricing. As IVOL could proxy for the fundamental uncertainty around a stock, uncertainty (DISPERSION) will be measured based on the following equation, as done by Hou and Loh (2016):

$$Dispersion_t^i = \frac{SDFY1_t^i}{\mathrm{MEANFY1}_t^i}. \quad (9)$$

$SDFY1_t^i$ is the standard deviation of the one fiscal year analysts' forecast of earnings per share for firm i announced in month t, $\mathrm{MEANFY1}_t^i$ is the mean estimate of the forecasted earnings per share by the analysts for firm i announced in month t. $Dispersion_t^i$ is therefore the standard deviation of the one fiscal year analysts' forecast of earnings per share scaled by the mean forecast for firm in announced in month t. Analysts forecast data is obtained and based on data from I/B/E/S unadjusted summary statistics file for U.S. common stocks. In the absence of sufficiently reliable observations for the Chinese stock market, this variable will not be constructed for the Chinese sub-sample.

3.5.2 Unexpected earning shocks (SUE)

Wong (2013) discovered an alternative explanation for the IVOL puzzle. He finds, as discussed in section 2.2.3, that once negative earnings shocks in relation to subsequent stock returns are considered, the IVOL estimate loses its predictive power. This finding implies that a firm's negative earning shocks can serve as a possible explanation for the puzzle. Hence, this study will also make use of the standardized unexpected earning shocks (SUE), as computed by Hou and Loh (2016). SUE will be estimated with the following equation:

$$SUE_t^i = \frac{QEAR_q^i - QEAR_{q-4}^i}{\sigma_q^i}, \tag{10}$$

where $QEAR_q^i$ is the quarterly earnings before extraordinary items for firm i announced in quarter q and $QEAR_{q-4}^i$ quarterly earnings before extraordinary items four quarters ago, respectively. σ_q^i is the standard deviation of the unexpected quarterly earnings ($QEAR_q^i - QEAR_{q-4}^i$) over the previous eight quarters. SUE_t^i is the most recently announced standardized unexpected earning shocks for firm i in month t. This candidate variable will not be taken into consideration for the Chinese subsample, due to the lack of sufficient observations across the entire Chinese sample.

3.5.3 Growth Options (MEBE)

The last candidate variable that will be used as a possible explanation for the IVOL puzzle is a variable based on a firm's growth options, as mentioned by Liu et al. (2021). They find a negative relation between IVOL related to growth options and subsequent stock returns, as previously discussed in the literature review. Therefore, we will use MEBE (Liu et al., 2018) as a proxy for the growth options of a firm. MEBE will be computed as:

$$MEBE_t^i = \frac{SHROUT_t^i \times PRC_t^i}{\mathrm{CEQ}_t^i}, \tag{11}$$

where $SHROUT_t^i$ is the amount of shares outstanding in thousands for firm i in month t, PRC_t^i is the closing stock price for firm i in month t and CEQ_t^i is the total book value of common/ordinary equity for firm i in month t. in other words, $MEBE_t^i$ is the market equity value divided by the book equity value of firm i in month t.

3.6 Decomposing the IVOL puzzle from different candidate explanations

To determine to what extent the different candidate explanations can possibly explain the IVOL puzzle, a decomposition method like that of Hou and Loh (2016) will be used. Multiple individual firm-level cross-sectional Fama-MacBeth (1973) regressions will be performed. Fama-MacBeth regressions will be used, as they weight each time period equally instead of a panel regression, which gives greater weight to time periods with more observations. Furthermore, Fama-MacBeth regressions are frequently used to investigate the relationship between IVOL and subsequent stock returns.

The first step of the decomposition method is to regress the cross-section of excess stock returns of firm i in month t on their IVOL estimate for month t-1. This will be done as follows:

$$R_t^i = a_t + \Upsilon_t IVOL_{t-1}^i + \varepsilon_t^i, \tag{12}$$

where R_t^i is firm i's monthly stock return in excess of the monthly risk-free rate and $IVOL_{t-1}^i$ is the month t-1 estimate as obtained from equation (2). A negative significant coefficient for factor loading Υ_t, will indicate a negative relationship between IVOL and subsequent stock returns and hence the IVOL puzzle.

For the second step, the IVOL estimate will be cross-sectionally regressed on the candidate variables:

$$IVOL_{t-1}^i = a_{t-1} + \delta_{t-1} Cand_{t-1}^i + \mu_{t-1}^i, \tag{13}$$

where $Cand_{t-1}^i$ is one of the candidate explanatory variables of firm i in month t-1. The regression of equation (13) allows to check whether there is a relationship between IVOL and the candidate variable. A relationship between IVOL and the candidate variable is a requirement, as Hou and Loh (2016) state that any possible explanation for the puzzle must be correlated with IVOL. With the regression coefficient estimates of equation (13), Hou and Loh (2016) show that it is possible to decompose $IVOL_{t-1}^i$ into two components: $\delta_{t-1} Cand_{t-1}^i$ is the part of $IVOL_{t-1}^i$ associated with the candidate variable. The residual part (a_{t-1} + μ_{t-1}^i) is the component of $IVOL_{t-1}^i$ that is dissociated with the candidate variable.

The final step is to decompose the estimated Υ_t coefficient of equation (12) into the part that can be explained by the candidate variable and the part that is left unexplained by the candidate variable. This will be done using the linearity of covariances as shown by Hou and Loh (2016):

$$
\begin{aligned}
\Upsilon_t &= \frac{Cov\left[R_t^i, IVOL_{t-1}^i\right]}{Var\left[\,IVOL_{t-1}^i\right]} \\
&= \frac{Cov[R_t^i, \delta_{t-1}Cand_{t-1}^i + a_{t-1} + \mu_{t-1}^i]}{Var[\,IVOL_{t-1}^i]} \\
&= \frac{Cov[R_t^i, \delta_{t-1}Cand_{t-1}^i]}{Var[\,IVOL_{t-1}^i]} + \frac{Cov[R_t^i, a_{t-1} + \mu_{t-1}^i]}{Var[\,IVOL_{t-1}^i]} \\
&= \Upsilon_t^C + \Upsilon_t^U. \qquad (14)
\end{aligned}
$$

Υ_t^C is the part of Υ_t that is related to the candidate variable and Υ_t^U is the part of Υ_t that is related to the residual part. The fraction of the IVOL puzzle that can be explained by the candidate variable and the fraction of the IVOL puzzle that is left unexplained are given as:

$$
E(\frac{\Upsilon_t^C}{\Upsilon_t}) \approx \frac{E(\Upsilon_t^C)}{E(\Upsilon_t)},\ E(\frac{\Upsilon_t^U}{\Upsilon_t}) \approx \frac{E(\Upsilon_t^U)}{E(\Upsilon_t)}. \qquad (15)
$$

A key advantage of the decomposition method of Hou and Loh (2016) is that due to the fact that Υ_t^C and Υ_t^U are required to be estimated through the variation of IVOL, both coefficients exactly sum up to the original Υ_t coefficient. This makes it possible to directly check what percentage of the IVOL puzzle can be explained by the candidate explanatory variable and what percentage remains unexplained. In addition, another important advantage is that the applied decomposition method allows to decompose Υ_t using multiple candidate variables simultaneously, to determine the marginal contribution of each individual candidate explanation in explaining the IVOL puzzle.

This study will also investigate the difference in the IVOL puzzle through different dimensions by making use of different subsample analyses. Additional control variables will be constructed based on these different dimensions such as time periods, country, and subsamples of stocks where the puzzle seems to be more prevalent, as discussed in section 2.2.4. These variables will be used to decompose the Υ_t coefficient of equation (12) to check to what extent variables based on these dimensions can explain the IVOL puzzle. This will also be done based on equations (12) – (15).

First, additional control variables based on situations where the puzzle seems to occur more strongly will be investigated for U.S. stocks. For this purpose, a subsample analysis will be performed for stocks with a price above $5, companies with a low credit rating and companies with high leverage, to investigate to what extent the different candidate variables can explain the puzzle among these situations. Next, additional control variables will be created based on different time periods, to investigate the IVOL puzzle through time for U.S. stocks. Four sub-samples based on time will be created. The subsamples that will be investigated will be 1982-1990, 1991-1999, 2000-2010 and 2011-2022 to examine if the explanatory power of the puzzle differs through time. This four time periods will be investigated, as it is likely that various market conditions differ considerably between the time periods examined. Lastly, additional control variables will be generated based on a specific country, where this study will look at the difference between the U.S. stock market and the Chinese stock market, as discussed earlier. These additional variables will investigate whether there is a difference in the explanatory power of the control variables for these countries.

To investigate whether the IVOL puzzle can be better explained over different time periods or between developed and emerging stock markets, the various regression coefficients of the candidate variables will be compared. A Z-test will be conducted to check whether the regression coefficients differ significantly from each other, and to conclude whether the explanatory power also differs per dimension. The equation provided by Clogg et al. (1995) will be used in the following way:

$$Z = \frac{\beta_1 - \beta_2}{\sqrt{(SE\beta_1)^2 + (SE\beta_2)^2}}, \tag{16}$$

where β_1 and β_2 are the regression coefficients of the candidate variables that are compared and $SE\beta_1$ and $SE\beta_2$ are their standard errors, respectively. The Z statistic follows a standard normal distribution under the null hypothesis of an equality of the two the regression coefficients of the candidate variables. This study will reject the hypothesis that the coefficients are the same if the Z-value is greater than the critical value of 1.645. Therefore, a confidence level of 95% will be used in this study.

4. Data

4.1 Dataset selection and cleaning

To decompose the IVOL puzzle, several databases are used to create the dataset of interest for this research. The main country of interest concerns the United States. To investigate the occurrence of the IVOL puzzle for U.S. stocks, the Wharton Research Data Services (WRDS) database was used. The WRDS database consists of several databases containing multiple several types of firm data. The WRDS platform is used to retrieve daily US common stock data from CRSP (share codes 10 or 11) listed on the NYSE, AMEX or NASDAQ (CRSP exchange code 1, 2, and 3) for a period from January 1982 to December 2022. This relatively large time frame will be examined to gain a better understanding of the behaviour of the IVOL puzzle. In addition, penny stocks (stocks with a price below 1 dollar) are removed from the sample as done by Hou and Loh (2016).

Additional databases were used to generate several candidate variables for stocks in the U.S.. Fundamental firm data for U.S. stocks is retrieved from Compustat's North America quarterly fundamentals database. Firm credit ratings are retrieved from Compustat's North America Ratings database. Analysts forecast and analyst coverage data is obtained from the Institutional Brokers Estimates System (IBES) using the unadjusted summary file. In order to compute the idiosyncratic volatility of a stock, as described in equation (1), the Fama and French (1993) three factors (Mkt, SMB, HML) are obtained from the Kenneth R. French Data Library.

To investigate the difference of the IVOL puzzle between developed and emerging stock markets, the study also makes use of Chinese stock data obtained from the China Stock Market and Accounting Research (CSMAR) database using the WRDS database. Daily Chinese common stock data is obtained for all A-share listed firms on the Shanghai Stock Exchange (SSE) or the Shenzhen Stock Exchange (SZSE) for a period from January 2000 to December 2021. Data before the year 2000 will not be taken into consideration in this study due to the lack of enough reliable data for the Chinese stock market for this period. All the observations are obtained in USD dollars. This study will only include Chinese A-shares, as these represent mainland China companies listed on the relevant stock exchange market. Penny stocks are also removed from the Chinese common stock sample. Firm fundamental data for Chinese stocks is also retrieved from the CSMAR database, accessed through WRDS. The Fama and French (1993) three factors are also obtained from the Kenneth R. French Data Library for the Chinese stocks.

Since the dataset contains some unrealistically extreme values that could potentially lead to biased results, all estimated variables are winsorized at the 1% level to limit the impact of these extreme values. This implies that the extreme positive (negative) outliers are replaced with the value corresponding to the ninety-ninth (first) percentile of the distribution for the estimated variable in question. In this study, winsorizing is used instead of removing the extreme observations, because some extreme values might be a genuine observation.

4.2 Descriptive statistics

Table 1 shows the descriptive statistics of the most important variables of the definitive dataset used in this study. Panel A shows the descriptive statistics for the U.S. sample. The U.S. sample contains more than two million three hundred thousand monthly firm observations. The average excess return is 0.5% per month with a standard deviation of 14.4%. The average IVOL estimate, computed using 1 month of daily returns, is 3.0% with a standard deviation of 2.2%. The average firm size, measured in millions and computed as the common shares outstanding multiplied by the monthly share price, is 1889.1 with a standard deviation of 6273.3. For the candidate variables related to investors' lottery preferences, the variable Skew shows an average value of 0.212, indicating that on average stock returns are positively skewed. The average MaxRet is 7.2%. For the candidate variables related to market frictions, the average LagRet is 0.7% and ZeroRet is 0.166 on average, which indicates that on average seventeen percent of the trading days in a specific month have a daily return of zero. The average monthly Spread is 2.6%. For the candidate variables related to other explanations Dispersion shows an average monthly value of 15.4% and SUE shows an average monthly value of 20.6%. The average monthly estimate for MEBE, which is a measure for a firm’s growth options, is 4.189, indicating that on average the monthly market value of a firm is four times larger than the book value of book value of common/ordinary equity.

Panel B of Table 1 shows the descriptive statistics for the Chinese sample. The Chinese sample consists of more than four hundred sixty thousand monthly firm observations, which is a considerably less than the US sample due to a smaller sample period for Chinese stocks. The Chinese sample displays an average monthly excess return of 0.8% per month with a standard deviation of 12.8%. The average IVOL estimate is 2.2% with a standard deviation of 1.1%. The average firm size is 1613.2 million, measured in U.S. dollars. Among the candidate variables related to investors' lottery preferences, the variable Skew shows an average value of 0.036 but CoSkew shows an average value of -0.011. This implies that stocks are on average positively skewed but the average skewness associated with the market returns are negative. For the candidate variables related to market frictions, the average monthly estimate of Amihud is 0.013, which is considerably lower than the average Amihud estimate for the U.S. sample. In addition, ZeroRet is 0.027 on average, which is also noticeably lower than the estimate for the U.S. sample. Lastly MEBE, the candidate variable related to other explanations, shows an average value of 3.764.

Table 1: Sample descriptive statistics

This table shows the descriptive statistics of the most important variables used in this study. Panel A shows the descriptive statistics for the U.S. sample from 1982 to 2022 and panel B shows the descriptive statistics for the Chinese sample from 2000 to 2021. N is the total number of monthly firm observations is the relevant sample. Excess Return is the average monthly stock return in excess of the risk-free rate. IVOL is the firm-monthly idiosyncratic volatility estimate, measured by equation (2). Size is the monthly market capitalization in million U.S. dollar. Skew is the monthly skewness of daily excess returns. Coskew is the co-skewness measure indicated by Chabi-Yo and Yang (2010). E(IS) is the expected idiosyncratic skewness indicated by Boyer et al. (2010). MaxRet is the maximum daily excess return of the relevant month. Lagret is the lagged monthly excess return. Amihud is the illiquidity measure by Amihud (2002). ZeroRet is the portion of trading days with a raw return of zero in a relevant month and Spread is the average monthly daily bid-ask spread. Dispersion is the dispersion in the analyst FY1 forecast indicated by Johnson (2004). SUE is the most recent standardized unexpected earnings and MEBE is a measure of a firms' growth options indicated by Liu et al. (2018).

Panel A: Descriptive statistics U.S. firms 1982-2022

Variable	N	Mean	St.dev	p1	p25	Median	p75	p99
Excess Return	2,353,062	0.005	0.144	-0.393	-0.069	-0.002	0.068	0.529
IVOL	2,376,434	0.030	0.022	0.004	0.015	0.024	0.038	0.122
Size ($mil)	2,376,434	1889.1	6273.3	2.9	37.4	156.2	804.2	47407.0
Candidate variables related to investors' lottery preferences								
Skew	2,366,278	0.212	0.918	-2.667	-0.260	0.175	0.661	3.100
CoSkew	2,376,434	0.004	0.115	-0.494	-0.018	0.000	0.020	0.568
E(IS)	1,601,151	0.491	0.331	-0.191	0.276	0.459	0.667	1.634
MaxRet	2,376,434	0.072	0.063	0.000	0.032	0.053	0.090	0.363
Candidate variables related to market frictions								
LagRet	2,366,573	0.007	0.146	-0.386	-0.068	-0.002	0.069	0.552
Amihud	2,261,506	1.838	6.097	0.000	0.004	0.054	0.637	44.307
ZeroRet	2,376,434	0.166	0.199	0.000	0.000	0.095	0.250	0.905
Spread	1,942,586	0.026	0.036	0.000	0.002	0.013	0.035	0.194
Candidate variables related to other explanations								
Dispersion	1,088,018	0.154	0.376	0.000	0.019	0.043	0.116	2.800
SUE	1,824,727	0.206	1.337	-2.842	-0.534	0.129	0.929	4.238
MEBE	2,276,914	4.189	10.036	-11.963	0.863	1.664	3.509	74.622

Table 1

Continued.

Panel B: Descriptive statistics Chinese firms 2000-2021								
Variable	N	Mean	St.dev	p1	p25	Median	p75	p99
Excess Return	462,157	0.008	0.128	-0.305	-0.070	-0.002	0.074	0.431
IVOL	468,056	0.022	0.011	0.006	0.014	0.020	0.028	0.060
Size ($mil)	468,056	1631.2	3271.3	76.7	338.6	647.8	1400.4	23730.7
Candidate variables related to investors' lottery preferences								
Skew	468,056	0.036	0.719	-1.849	-0.402	0.023	0.469	1.992
CoSkew	468,056	-0.011	0.036	-0.131	-0.027	-0.007	0.006	0.102
E(IS)	241,097	0.680	0.264	0.056	0.492	0.677	0.856	1.333
MaxRet	468,056	0.054	0.028	0.012	0.032	0.048	0.075	0.101
Candidate variables related to market frictions								
LagRet	466,854	0.010	0.131	-0.308	-0.070	-0.001	0.076	0.456
Amihud	468,056	0.013	0.025	0.000	0.002	0.004	0.011	0.158
ZeroRet	468,056	0.027	0.041	0.000	0.000	0.000	0.048	0.182
Candidate variables related to other explanations								
MEBE	467,935	3.764	3.651	0.319	1.711	2.716	4.424	25.156

5 Results

5.1 The occurrence of the IVOL puzzle and the difference in occurrence of the puzzle over time and between a developed and emerging stock market

This section discusses the results with regard to the occurrence of the IVOL puzzle and the occurrence of the IVOL puzzle over time. The presence of the IVOL puzzle is examined by making use of the established method discussed in section 3.2. Stocks are sorted into quintile portfolios based on their monthly IVOL estimate of equation (2) and are used to implement the 1/0/1 strategy as proposed by Ang et al. (2006). The first portfolio (P1) contains the stocks with the lowest IVOL estimate and the fifth portfolio (P5) contains the stocks with the highest IVOL estimate. The monthly excess returns of the 1/0/1 strategy are determined by using equal-weighted and value-weighted portfolios.

Table 2 reports the results of the 1/0/1 strategy for the U.S. sample for portfolios sorted on IVOL. In panel A, the results for the equally weighted portfolios are shown and the results for the value-weighted portfolios are shown in panel B. Panel A shows that the average monthly excess return equals 0.88% for quintile one (P1) and stays positive for quintile two (P2), three (P3) and four (P4). However, quintile portfolio five (P5), containing stocks with the highest IVOL estimate, shows an average excess return of -0.55% per month. The long-short portfolio (P5-P1), which reports the average monthly return difference between quintile portfolio five and quintile portfolio one, shows an average excess return of -1.43% per month, significant at the 1% level, indicating that high IVOL stocks tend to significantly underperform low IVOL stocks on average. Furthermore, the difference in the FF-3 alpha between quintile portfolio five and quintile portfolio one is -1.60% and also significant at the 1% level. This significant difference in the FF-3 alpha shows the average return on the long-short portfolio that cannot be explained by the Fama and French (1993) three factor model.

Panel B reports the results for the value-weighted portfolios and shows comparable results as panel A. The average monthly excess returns are positive for quintile portfolio one to four, while quintile portfolio five is once again showing negative monthly excess returns on average. The long-short portfolio has an average monthly return of -1.28%, also significant at the 1% level. The long-short portfolio FF-3 alpha is -1.47% and significant at the 1% level, suggesting that the FF-3 model is again unable to explain the average monthly excess returns on portfolio P5 - P1. The obtained results are in line with the results of Ang et al. (2006), who also find a significant negative monthly return and FF-3 alpha on their long-short portfolio sorted on IVOL.

It is interesting to note that the standard deviation (St.dev) of the quintile portfolios monotonically increases from quintile one to five and that the average size tends to monotonically decrease from quintile one to five for both equal-weighted and value-weighted portfolios.

Table 2: Portfolios sorted by IVOL for the U.S. sample

*The table reports the results from both equal- and value-weighted quintile portfolios sorted on the IVOL estimate of the past month. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. Portfolio P1 (P5) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. The Mean Monthly Excess Returns and St.dev are measured in monthly percentages. Size reports the monthly market capitalization in million U.S. dollar for firms in the portfolio. The FF-3 Alpha column reports the alpha obtained from a regression of portfolio returns on the Fama and French (1993) 3 factors. The t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The sample period is January 1982 to December 2022.*

Portfolio Rank	Mean Monthly Excess Returns	St.dev	Size	% Mkt Share	FF-3 Alpha
Panel A: Equal-weighted portfolios sorted by IVOL relative to FF-3 factors					
P1	0.88%	3.65%	5087.56	47.6%	0.27%***
					(4.88)
P2	0.96%	4.72%	2925.41	27.4%	0.21%***
					(3.96)
P3	0.87%	5.48%	1582.66	14.5%	0.08%
					(1.38)
P4	0.49%	6.41%	799.25	7.3%	-0.32%***
					(-3.88)
P5	-0.55%	7.00%	339.19	3.2%	-1.33%***
					(-9.23)
P5-P1	-1.43%***				-1.60%***
	(-6.30)				(-9.65)
Panel B: Value-weighted portfolios sorted by IVOL relative to FF-3 factors					
P1	0.84%	3.96%	5087.56	47.6%	0.16%***
					(3.30)
P2	0.82%	4.78%	2925.41	27.4%	0.04%
					(0.81)
P3	0.77%	5.79%	1582.66	14.5%	-0.07%
					(1.01)
P4	0.38%	7.09%	799.25	7.3%	-0.48%***
					(-4.48)
P5	-0.44%	7.92%	339.19	3.2%	-1.31%***
					(-8.00)
P5-P1	-1.28%***				-1.47%***
	(-4.53)				(-7.60)

This indicates that smaller firms tend to have higher IVOL estimates and more volatile returns than larger firms on average, which is generally known in the finance literature. Furthermore, Table 2 shows that the average market share (% Mkt Share) of the quintile portfolios strongly differs from the initial 20% of stocks in each portfolio. For example, quintile portfolio 5 contains 20% of the shares, but has a market share of only 3.2% on average. However, all these interesting results are in line with the results of Ang et al. (2006), even though this study focuses on a different and more recent period. Based on the significantly negative returns on the long-short portfolios in Table 2, it can therefore be concluded that the IVOL puzzle does indeed occur in the entire U.S. sample used for this study.

Although the results of Table 2 indicate that the IVOL puzzle appears to be present in the U.S. sample, it is also interesting to investigate if the puzzle also occurs when the analysis of Table 2 is repeated, but the sample is split into smaller time periods. This allows checking whether IVOL and the occurrence of the IVOL puzzle indeed varies over time, as indicated by Fu (2009) and Brandt et al. (2010) and allows to check how the IVOL puzzle behaves during periods that capture a substantial crisis. The total is sample is divided into ten sub-samples, where a difference in the significance of the returns of these sub-sample long-short portfolios may indicate a time-varying occurrence of the IVOL puzzle.

Table 3 reports the results of all the quintile and long-short portfolios for the ten different subsamples. Panel A reports the average monthly excess returns on the equal-weighted portfolios and the value-weighted portfolios are reported in panel B. The excess returns on the equal-weighted long-short portfolios are significantly negative in seven of the ten subsamples, with the most negative excess returns of -2.11% during the Covid-19 crisis (2018-2022), significant at the 1% level. Interestingly, panel A only shows a significant negative return for the financial crisis around the year 2008 (2006-2009) and the Covid-19 crisis around the year 2020 (2018-2022) but not for the other two examined crises, where the relationship between IVOL and subsequent returns appears to be insignificant.

However, panel B reports only significantly negative returns on the long-short portfolios for four out of the ten subsamples, with the most negative excess returns of -2.07% before the early 1990s recession in the U.S. (1986-1989), significant at the 1% level. Compared to panel A, in panel B only the Covid-19 crisis shows a significant negative return with regard to the different crisis periods examined. The difference in significance between the equal and value-weighted portfolios for the subsample between 2006 and 2017, indicates that the IVOL puzzle tends to be heavily dependent on the returns of small cap stocks for these periods, due to the insignificant returns on the value-weighted portfolios. This result was also found by Fu (2009), who showed that the results of Ang et al. (2006) are influenced by a subset of small firms with high IVOL estimates. Furthermore, the difference in the significance of the excess returns on the long-short portfolios for the different sub-samples seems to indicate that the IVOL puzzle is indeed time-varying, which is in line with the results of Brandt et al. (2010). Moreover, the crisis sub-samples show mixed results regarding the occurrence of the IVOL puzzle, making it difficult to draw conclusions about whether or not the IVOL puzzle significantly occurs during crisis periods.

In addition to table 3, which already gives an indication of the time-varying aspect of the IVOL puzzle, Figure 1 shows a time-series plot of the monthly mean of IVOL for the U.S. sample. The figure shows both the equal-weighted monthly mean IVOL estimate (black line) and the value-weighted monthly mean IVOL estimate (gray dashed line). The results of Figure 1 demonstrate that the monthly mean IVOL estimate varies widely over time. The figure shows some extreme fluctuations, most of which are centred around a crisis period (Early 1990s recession, DotCom bubble around the year 2000, The financial crisis around the year 2008 and the Covid-19 crisis around the year 2020). However, the figure also shows that after a peak, the IVOL estimate tends to decrease extremely. Furthermore, the figure shows evidence that the equal-weighted IVOL estimate is strongly influenced by small-cap stocks, as the figure shows that the average value-weighted IVOL estimate is always smaller than the equal-weighted estimate.

Out of the results of Table 3 and Figure 1, it can be concluded that IVOL and the occurrence of the IVOL puzzle indeed differs through time. These results are in line with Fu (2009), who showed that IVOL is time-varying and Brandt et al. (2010), who indicated that the IVOL puzzle is more an episodic phenomenon than a trend over time.

While Table 2 shows that the IVOL puzzle occurs in the U.S. sample, but Table 3 and Figure 1 show that the occurrence of the puzzle differs over time, the occurrence of the puzzle may also differ between developed and emerging markets, due to the considerable differences in market characteristics discussed in section 2.4. The U.S. stock market and the Chinese stock market are compared to determine whether there is a difference between developed and emerging markets, making use of the same analysis as for Table 2. In addition, the total Chinese sample is divided into six subsamples to check whether the IVOL puzzle also tends to differ over time for Chinese stocks and these subsamples are matched with the relevant subsamples for U.S. stocks. A difference in the significance of the returns of these sub-sample long-short portfolios may indicate a difference in the occurrence of the IVOL puzzle between developed and emerging markets. Only the results of the value-weighted portfolios have been compared, as Table 3 and Figure 1 showed that the IVOL puzzle often relies heavily on the performance of small-cap stocks.

Table 4 shows the results of the comparison between a developed and an emerging market for the different subsamples. Panel A reports the average monthly excess returns for U.S. stocks and panel B reports these for the Chinese stocks. The first column reports the results for the subsample 2000-2021 (whole Chinese sample). The excess returns on the long-short portfolios for this subsample are -1.04% for the U.S. and -1.05% for Chinese stocks, both significant at the 1% level. Therefore, these negative significant returns on the P5-P1 portfolios point to the presence of the IVOL puzzle in the Chinese sample and in the matched U.S. subsample. However, the occurrence of the puzzle differs between the U.S. and Chinese stock market when looking at the different sub-samples. The U.S. shows only one subsample where the return on the long-short portfolio is significantly negative (2012-2015), where China shows three subsample where returns on long-short portfolio is significantly negative

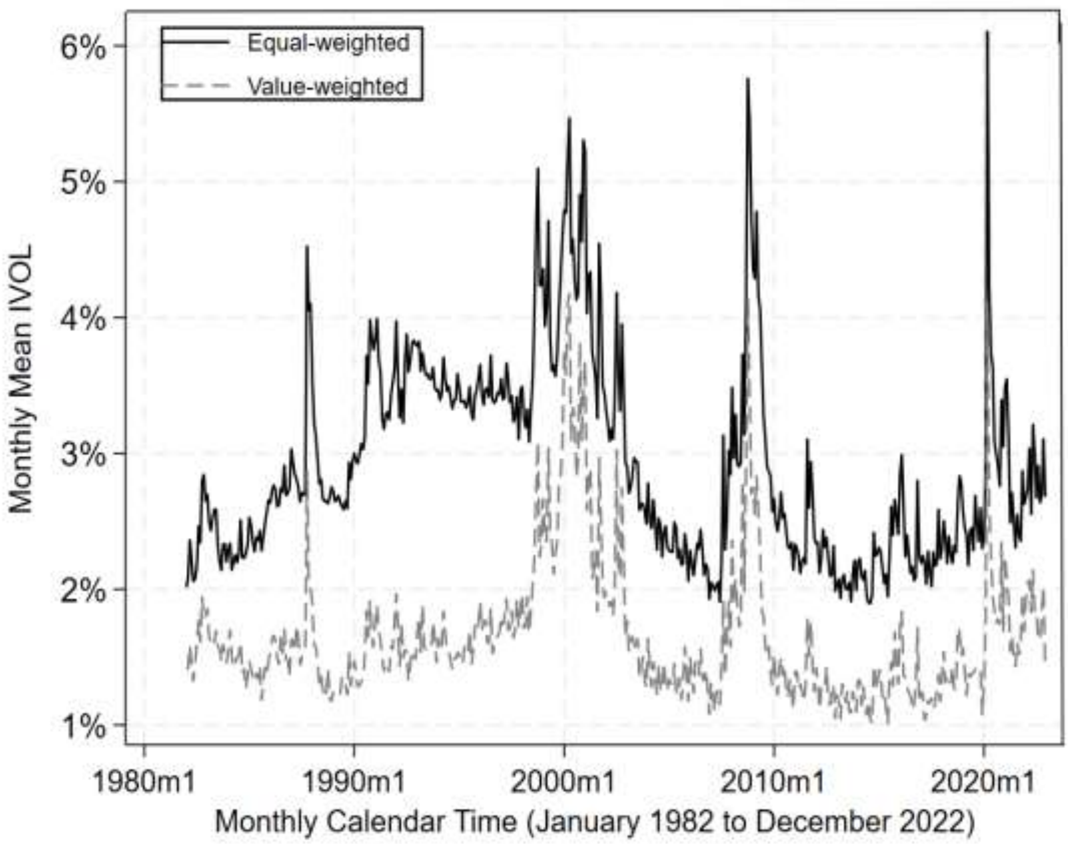
Table 3: Portfolios sorted by IVOL for different time periods for U.S. stocks

*The table reports the results from both equal- and value-weighted quintile portfolios sorted on the IVOL estimate of the past month. Portfolios are estimated separately for each sub-sample referring to a specific time period. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. Portfolio P1 (P5) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. The Mean Monthly Excess Returns is measured in monthly percentages. The t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The sub-samples are estimated over the period January 1982 to December 2022.*

Time period	1982 - 1985	1986 - 1989	1990 - 1993	1994 - 1997	1998 - 2001	2002 - 2005	2006 - 2009	2010 - 2013	2014 - 2017	2018 - 2022
Portfolio Rank	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns
Panel A: Equal-weighted portfolios sorted by IVOL relative to FF-3 factors										
P1	1.32%	0.37%	0.70%	1.33%	0.36%	1.17%	-0.02%	1.60%	1.26%	0.75%
P2	1.35%	0.44%	1.20%	1.26%	0.53%	1.37%	-0.21%	1.75%	1.13%	0.84%
P3	1.24%	0.30%	1.26%	0.96%	0.55%	1.40%	-0.22%	1.79%	1.00%	0.46%
P4	0.95%	-0.29%	0.94%	0.45%	0.02%	1.45%	-0.33%	1.65%	0.38%	-0.16%
P5	-0.63%	-1.38%	0.18%	-0.72%	-1.55%	1.04%	-1.33%	0.90%	-0.50%	-1.36%
P5-P1	-1.95%***	-1.75%***	-0.52%	-2.05%***	-1.91%	-0.13%	-1.31%**	-0.70%*	-1.76%***	-2.11%***
	(-3.46)	(-5.32)	(0.93)	(-3.31)	(-1.33)	(0.17)	(-2.55)	(-1.68)	(-3.40)	(-2.69)
Panel B: Value-weighted portfolios sorted by IVOL relative to FF-3 factors										
P1	1.14%	0.89%	0.75%	1.33%	0.44%	0.48%	-0.03%	1.38%	1.16%	0.84%
P2	0.91%	0.71%	0.94%	1.18%	0.48%	0.74%	-0.19%	1.57%	1.15%	0.75%
P3	0.88%	0.49%	1.33%	0.86%	0.39%	0.85%	-0.35%	1.70%	1.04%	0.46%
P4	0.69%	-0.20%	0.61%	0.14%	-0.58%	0.97%	-0.20%	1.46%	0.40%	0.51%
P5	-0.83%	-1.18%	-0.26%	-0.77%	-1.98%	0.81%	-0.91%	1.22%	0.40%	-0.78%
P5-P1	-1.97%***	-2.07%***	-1.00%	-2.10%**	-2.42%	0.32%	-0.89%	-0.16%	-0.76%	-1.62%*
	(-3.70)	(-4.29)	(1.52)	(-2.57)	(-1.31)	(0.35)	(-1.02)	(-0.34)	(-1.39)	(1.89)

Figure 1: Monthly Mean of IVOL for the U.S. sample

This figure shows the Equal-weighted Monthly Mean IVOL estimate (black line) and the Value-weighted Monthly Mean IVOL estimate (gray dashed line) for each month between January 1982 and December 2022. IVOL is measured using daily returns following the methodology described in section 3.1.



(2000-2003, 2004-2007 and 2016-2019), indicating that there may is a difference in the occurrence of the puzzle between developed and emerging markets..

This difference in occurrence of the puzzle may be explained by the fractions of retail investors who invest on the market in question. For example, the Chinese stock market is dominated by retail investors, unlike developed markets, which are dominated by institutional investors (Leippold et al., 2022). This difference in the fraction of retail investors, coupled with the findings of Brandt et al. (2010), who found that the IVOL puzzle is more prominent in times of increased retail investor trading activity, could possibly explain this difference in occurrence. Furthermore, contrary to developed markets, not all stocks on the Chinese stock market are allowed to be shorted (Chen et al., 2023). The difference in short-sell constraints, which tends to increase the occurrence of the puzzle (Stambaugh, Yu and Yuan, 2015), may also account for the difference in occurrence. In addition, it is also interesting that among the six subsamples, there is not a single subsample where both the U.S. and the Chinese market show a significant negative return on the long-short portfolio.

Table 4: Value-weighted portfolios sorted by IVOL for a developed and emerging stock market

*The table reports the results from value-weighted quintile portfolios sorted on the IVOL estimate of the past month. Portfolios are estimated separately for U.S. and Chinese stocks and for each sub-sample referring to a specific time period. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. Portfolio P1 (P5) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. The Excess Returns (Mean Monthly Excess Returns in Table 2 and Table 3) is measured in monthly percentages. The t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The sub-samples are estimated over the period January 2000 to December 2021.*

Time period	2000-2021	2000-2003	2004-2007	2008-2011	2012-2015	2016-2019	2020-2021
Portfolio Rank	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns	Excess Returns
Panel A: Value-weighted portfolios sorted by IVOL relative to FF-3 factors for U.S. stocks							
P1	0.78%	0.40%	0.42%	0.46%	1.06%	1.16%	1.55%
P2	0.77%	0.12%	0.49%	0.45%	1.10%	1.23%	1.64%
P3	0.63%	-0.42%	0.40%	0.52%	0.99%	1.08%	1.76%
P4	0.34%	-1.25%	0.43%	0.53%	0.35%	1.13%	1.39%
P5	-0.26%	-2.19%	0.12%	-0.17%	0.17%	0.61%	-0.01%
P5-P1	-1.04%**	-2.59%	-0.30%	-0.64%	-0.89%*	-0.55%	-1.56%
	(-2.34)	(-1.37)	(-0.61)	(-0.70)	(-1.72)	(-1.17)	(-0.80)
Panel B: Value-weighted portfolios sorted by IVOL relative to FF-3 factors for Chinese stocks							
P1	0.89%	0.34%	2.92%	-0.41%	1.24%	0.44%	0.71%
P2	0.86%	0.11%	3.03%	-0.57%	1.05%	0.52%	1.22%
P3	0.91%	-0.14%	3.42%	-0.84%	1.22%	0.46%	1.81%
P4	0.61%	-0.46%	2.49%	-0.47%	0.89%	0.10%	1.68%
P5	-0.16%	-1.07%	1.77%	-1.15%	0.32%	-1.08%	0.63%
P5-P1	-1.05%***	-1.41%***	-1.15%*	-0.74%	-0.92%	-1.52%**	-0.08%
	(3.86)	(-3.13)	(-4.29)	(1.09)	(-1.30)	(-2.38)	(-0.06)

In conclusion, the results of Table 2 indicate that the IVOL puzzle tends to occur in the U.S. sample and show that IVOL is significantly negative related to future stock returns, which is in accordance with Ang et al. (2006). However, the results of Table 3 and Figure 1 indicate that IVOL and the IVOL puzzle vary over time and the puzzle does not always occur when smaller time intervals are considered. These time-varying results are therefore consistent with Fu (2009) and Brandt et al. (2010). Apart from the time-varying aspect of the IVOL puzzle, Table 4 shows that there is also a difference in the occurrence of the puzzle between developed and emerging stock markets when looking at smaller time frames, which could possibly be explained by the difference in the fraction of retail investors or the difference in short-sell constraints, or both. To fully answer the proposed hypotheses, the IVOL puzzle will first be decomposed for the different explanations, of which the results will be discussed in the following sections.

5.2 Investigating the IVOL puzzle using one candidate explanation at a time

5.2.1 The idiosyncratic volatility puzzle controlling for the candidate explanation

In the previous section it was concluded that the IVOL puzzle is still present in the sample used for this study. In this section it will be investigated which candidate explanation can explain the puzzle for the U.S. sample. The estimated explanatory variables discussed in section 3.3, 3.4 and 3.5 are examined. Before the IVOL puzzle is decomposed using the method of Hou and Loh (2016), as described in section 3.6, Table 5 reports the Fama-MacBeth cross sectional regression results of monthly individual stock excess returns of month t on their month t−1 IVOL estimate and on the several candidate variables for the U.S. sample. All the coefficients are multiplied by 100. Based on the results of Table 5, the relationship between IVOL and excess returns, after controlling for the different candidate explanations, can be examined.

Model 1 regresses the monthly excess returns on the IVOL estimate of the previous month alone over the whole sample period. The regression uses a total of 2,353,062 firm-month observations with an average of 4,792 firms per month (not reported in Table 5). The average coefficient of IVOL is -32.783% and highly significant at the 1% level, indicating that IVOL is also negatively correlated with subsequent stock returns at the individual stock level. This negative correlation is in line with the results from Table 2 and the several different studies that found a negative relationship between IVOL and future returns. Models 2 through 5 add the various candidate variables related to investors' lottery preferences, Models 6 through 9 add the different candidate variables related to market frictions and models 10 through 12 add the different candidate variables related to other explanations as control variables one by one to Model 1. Using the several candidate variables as control variables allows to assess the relation between IVOL and subsequent returns when the candidate explanations are taken into account. There may be differences in the number of observations between the different models, because not every candidate explanation has a value for every observation due to data limitations.

However, Table 5 shows an interesting result. The average coefficient of IVOL is negative and highly significant at the 1% level in all of the 12 models after controlling for the different candidate explanations. Out of this interesting result, it can be concluded that none of the candidate explanations is able to absorb the entire negative relationship between IVOL and subsequent stock returns. Therefore, this implies that none of the candidate explanations can explain away the complete IVOL puzzle. However, using the decomposition approach described in section 3.6, it is possible to determine whether the various candidate explanations can explain at least part of the puzzle and what fraction of the puzzle can be explained by the relevant candidate explanatory variable. The following sections discus the decomposition of the IVOL puzzle using one candidate statement at a time.

5.2.2 Decomposing the IVOL puzzle using one candidate explanation at a time

This section discusses the decomposition of the IVOL puzzle using one candidate explanation at a time. Table 6 reports the results of the decomposition analysis. First, the decomposition analysis is discussed in detail using the candidate variable Skew. In stage 1, equation (12) is used to regress the monthly excess stock returns in month t on their IVOL estimate for month t-1. Therefore, this regression is approximately the same as Model 1 in Table 5, with the exception that stage 1 excludes monthly observations with a missing estimate for Skew, to ensure that the sample is kept the same when Skew is added to the analysis later on. The average coefficient of IVOL is -33.222% and is highly significant at the 1% level.

In stage 2, equation (13) is used to regress the monthly IVOL estimate on the Skew estimate for each month to determine whether the IVOL estimate is related to a stock's skewness. The average coefficient of Skew is 0.402% and significant at the 1% level. This implies that the Skew candidate variable is related to IVOL, where a change of one unit in Skew would result in a change of 0.402% in IVOL. However, the average adjusted R-squared shows that only 3.7% of the variation in IVOL can be explained by Skew. Based on the estimated coefficients of stage 2, IVOL can be divided into a component of IVOL ($\delta_{t-1}Skew_{t-1}^{i}$), which is related to Skew, and a residual component of IVOL (a_{t-1} + μ_{t-1}^{i}), which is not related to Skew for each month.

In stage 3, equation (14) is used to decompose the coefficient of IVOL from stage 1 (Υ_t) into a component that is related to Skew (Υ_t^{Skew}) and a residual component of the coefficient (Υ_t^{U}) that is not related to Skew and left unexplained. The average coefficient of Υ_t^{Skew} is -1.136% and the average coefficient of Υ_t^{U} is -32.087%, which together add up to the IVOL coefficient of stage 1 (-33.222%). This allows the fraction of the IVOL puzzle explained by Skew to be calculated as $\frac{-1.136\%}{-33.222\%}$ = 3.42% which is significant at the 1% level. The fraction of the IVOL puzzle left unexplained is calculated as $\frac{-32.087\%}{-33.222\%}$ = 96.58%, which is also significant at the 1% level. Therefore, it can be concluded that Skew can only explain 3.42% of the IVOL puzzle.

Table 5: The relation between IVOL and excess returns for the U.S. sample

*The table reports the results of firm-level Fama-MacBeth cross-sectional regressions estimated each month from January 1982 to December 2022. The dependent variable is a stocks monthly excess return. Time-series averages of the coefficients are multiplied by 100 and the associated time-series t-statistics are reported in parentheses. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. The several candidate explanatory variables are estimated as described is sections 3.3, 3.4 and 3.5. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.*

Variable	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8	Model 9	Model 10	Model 11	Model 12
Intercept	1.439***	1.454***	1.436***	1.175***	1.417***	1.396***	1.525***	1.569***	1.479***	1.358***	1.209***	1.512***
	(7.39)	(7.43)	(7.41)	(4.93)	(7.33)	(7.41)	(7.85)	(6.94)	(7.10)	(6.65)	(6.07)	(7.59)
IVOL	-32.783***	-33.169***	-32.758***	-23.677***	-22.605***	-32.167***	-36.076***	-33.141***	-38.039***	-22.149***	-19.876***	-31.098***
	(-9.41)	(-9.40)	(-9.41)	(-7.08)	(-4.72)	(-9.35)	(-9.49)	(-9.57)	(-8.81)	(-4.80)	(-5.57)	(-8.99)
Skew		-0.020										
		(-0.91)										
CoSkew			-0.303									
			(-1.34)									
E(IS)				0.296*								
				(1.70)								
MaxRet					-4.000***							
					(-4.25)							
LagRet						-2.144***						
						(-5.69)						
Amihud							0.022***					
							(2.90)					
ZeroRet								-0.391				
								(-0.78)				
Spread									2.988			
									(0.87)			
Dispersion										-0.300***		
										(-4.21)		
SUE											0.305***	
											(17.42)	
MEBE												-0.032***
												(-10.46)
Avg adj R^2	0.015	0.016	0.016	0.019	0.017	0.021	0.019	0.020	0.025	0.020	0.016	0.017
Num obs	2,353,062	2,343,230	2,353,062	1,586,845	2,353,062	2,343,808	2,240,614	2,353,062	1,923,687	1,079,080	1,811,952	2,256,667

Table 6: Decomposing the IVOL puzzle using one candidate explanation at a time for U.S. stocks

*The table reports the results of the decomposition analysis proposed by Hou and Loh (2016) using one candidate explanation at a time. The relationship between IVOL and monthly excess returns is decomposed into the fraction related to a candidate variable and a residual component using firm-level Fama-MacBeth cross-sectional regressions. Stage 1 regresses monthly excess returns of month t on the IVOL estimate of the previous month ($R_t^i = a_t + \Upsilon_t IVOL_{t-1}^i + \varepsilon_t^i$). Stage 2 regresses IVOL on the candidate variable ($IVOL_{t-1}^i = a_{t-1} + \delta_{t-1} Cand_{t-1}^i + \mu_{t-1}^i$) to decompose $IVOL_{t-1}^i$ into two components: $\delta_{t-1} Cand_{t-1}^i$ and ($a_{t-1} + \mu_{t-1}^i$). In Stage 4, the Υ_t coefficient from Stage 1 is decomposed as: $\Upsilon_t = \frac{Cov[R_t^i, IVOL_{t-1}^i]}{Var[IVOL_{t-1}^i]} = \frac{Cov[R_t^i, \delta_{t-1} Cand_{t-1}^i]}{Var[IVOL_{t-1}^i]} + \frac{Cov[R_t^i, a_{t-1} + \mu_{t-1}^i]}{Var[IVOL_{t-1}^i]} = \Upsilon_t^C + \Upsilon_t^U$. Dividing Υ_t^C by Υ_t measures the fraction of the IVOL-return relation captured by the candidate variable, and Υ_t^U divided by Υ_t measures the fraction of the relation left unexplained by the candidate explanation. Time-series averages of the coefficients are multiplied by 100 and the associated time-series t-statistics are reported in parentheses. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. The several candidate explanatory variables are estimated as described is sections 3.3, 3.4 and 3.5. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The sample period is January 1982 to December 2022*

Stage	Description	Variable	Skew		CoSkew		E(IS)		MaxRet		LagRet	
1	Excess returns on IVOL	Intercept	1.459***	(7.43)	1.439***	(7.39)	1.258***	(6.16)	1.439***	(7.39)	1.436***	(7.38)
		IVOL	-33.222***	(-9.55)	-32.783***	(-9.41)	-21.447***	(-5.75)	-32.783***	(-9.41)	-32.716***	(-9.36)
2	IVOL on candidate variable	Intercept	2.851***	(89.95)	2.913***	(87.94)	1.594***	(41.22)	0.786***	(85.34)	2.823***	(89.19)
		Candidate	0.402***	(36.50)	1.329***	(17.71)	2.360***	(36.86)	30.303***	(297.18)	1.180***	(11.90)
		Avg adj R^2	3.7%		1.0%		15.8%		77.2%		4.5%	
3	Decompose stage 1 IVOL coefficient	Candidate	-1.136		-0.359		-1.102		-26.854		-2.092	
			3.42%***	(4.27)	1.09%**	(2.51)	5.14%	(0.71)	81.91%***	(10.22)	6.39%***	(3.09)
		Residual	-32.087		-32.424		-20.345		-5.929		-30.624	
			96.58%***	(9.57)	98.91***	(9.43)	94.86%***	(7.33)	18.19%***	(5.49)	93.61%	(9.28)
		Total	-33.222***	(-9.55)	-32.783***	(-9.41)	-21.447	(-5.75)	-32.783***	(-9.41)	-32.716***	(-9.36)
			100%		100%		100%		100%		100%	
Avg # firms/mth			4786.4		4792.4		3487.6		4792.4		4775.4	
Num obs			2,343,230		2,353,062		1,586,845		2,353,062		2,343,808	

(Continued on next page)

Table 6:
Continued.

Stage	Description	Variable	Amihud		ZeroRet		Spread		Dispersion		SUE		MEBE	
1	Excess returns on IVOL	Intercept	1.497***	(7.51)	1.439***	(7.39)	1.492***	(7.18)	1.354***	(6.61)	1.332***	(6.67)	1.424***	(7.18)
		IVOL	-34.016***	(-9.60)	-32.783***	(-9.41)	-35.579***	(-9.57)	-24.030***	(-5.05)	-22.180***	(-6.13)	-31.635***	(-9.06)
2	IVOL on candidate variable	Intercept	2.714***	(90.82)	2.929***	(85.33)	2.101***	(66.05)	2.232***	(79.96)	2.830***	(87.92)	2.922***	(88.04)
		Candidate	0.102***	(38.42)	0.781***	(7.14)	41.636***	(84.19)	0.822***	(75.68)	-0.138***	(-59.85)	0.004***	(5.78)
		Avg adj R^2	12.8%		1.1%		27.1%		4.3%		1.0%		0.4%	
3	Decompose stage 1 IVOL coefficient	Candidate	-2.255		-0.199		-8.032		-2.808		-2.522		-0.699	
			6.63%***	(3.21)	0.61%	(0.61)	22.58%***	(4.93)	11.69%***	(5.52)	11.37%***	(14.10)	2.21%***	(5.27)
		Residual	-31.761		-32.584		-27.547		-21.222		-19.658		-30.936	
			93.37%***	(9.19)	99.39%***	(9.55)	77.42%***	(7.99)	88.31%***	(4.79)	88.63%***	(5.57)	97.79%***	(8.99)
		Total	-34.016***	(-9.60)	-32.783***	(-9.41)	-35.579***	(-9.57)	-24.030***	(-5.05)	-22.180***	(-6.13)	-31.635***	(-9.06)
			100%		100%		100%		100%		100%		100%	
	Avg # firms/mth		4563.4		4792.4		4007.7		2197.7		3690.3		4596.1	
	Num obs		2,240,614		2,353,062		1,923,687		1,079,080		1,811,952		2,256,667	

The other measures of skewness, CoSkew and E(IS), do not show very promising results either. Coskew explains only 1.09%, significant at the 5% level, and E(IS) explains 5.14% of the IVOL puzzle but shows no significant coefficient. Moreover, it is noteworthy that none of the measures of skewness in Table 2 is able to predict returns negatively after controlling for IVOL, where E(IS) even shows a significant positive coefficient. Table 6 therefore initiates that only the part of the skewness measures related to IVOL is in accordance with the lottery preference-based explanations. This finding is therefore inconsistent with the studies of Boyer et al. (2010), Barberis and Huang (2008) and Chabi-Yo and Yang (2010). As the last lottery preference-based explanation, MaxRet does show a promising result. The fraction of the puzzle explained by MaxRet is 81.91% and significant at the 1% level. However, Hou and Loh (2016) initiate in their study that this large fraction could be mechanical, due to the almost perfect collinearity between MaxRet and IVOL. Also in this study, the correlation between MaxRet and IVOL is 93%, which therefore initiates that the finding is likely to be biased.

Furthermore, Table 6 also shows the decomposition results for the candidate variables related to market frictions. LagRet can explain 6.39% of the IVOL puzzle, which is significant at the 1% level, suggesting that the return reversal of stocks can indeed explain a small fraction of the IVOL puzzle. As a measure of illiquidity, Amihud is able to explain 6.63% of the puzzle, significant at the 1% level. However, Table 2 also shows that the return predictability of Amihud after controlling for IVOL is positive. This also suggests for Amihud that only the part of Amihud related to IVOL can predict returns negatively. ZeroRet is unable to explain a significant portion of the puzzle and only explains 0.61%. As a final measure of illiquidity, Spread also shows a promising significant result and is able to explain 22.58% of the IVOL puzzle, but again, if Table 2 is also considered, only the portion of the Spread associated with IVOL tends to negatively predict returns, as Spread has no return predicting power after controlling for IVOL.

Lastly, Table 6 reports the decomposition results for the candidate variables related to other explanations. The table shows that both Dispersion and SUE are able to explain approximately one tenth of the IVOL puzzle (11.69% for Dispersion and 11.37% for SUE, both significant at the 1% level). This suggests that the fundamental uncertainty around a stock and the negative earnings shocks of high IVOL stocks are able to explain a part of the IVOL puzzle. MEBE, as a proxy for growth options of a firm, is only able to explain 2.21% of the IVOL puzzle, significant at the 1% level.

To conclude, except from E(IS) and Zero Ret, all of the candidate variables are able to explain a fraction of the IVOL puzzle. However, most of them can just capture a small fraction of the puzzle, where only four candidate statements can explain more than 10% of the puzzle. MaxRet is able to explain most of the IVOL puzzle, but there is a high probability that this result is biased due to the high correlation between MaxRet and IVOL. Following, Spread, Dispersion and SUE show the most promising results for explaining the IVOL puzzle independently. Although, this section indicates that explanations based on market frictions and other explanations can best explain the puzzle on their own, the following section investigates each variable's marginal contribution after controlling for the other candidate variables.

5.3 Investigating the IVOL puzzle using multiple candidate explanation at a time

This section discusses the decomposition of the IVOL puzzle using multiple candidate explanation at the same time, in order to examine the marginal contribution of each variable when the other explanations are also taken into consideration for U.S. stocks. As also done in Hou and Loh (2016), MaxRet is not included in the multivariate analysis due to the near-perfect relationship with IVOL. Table 7 reports the results of the multivariate analysis.

Model 1 in Table 7 shows the results of the multivariate analysis with all the variables included except from MaxRet. Model 1 has an average of only 1480 firm observations per month, due to the availability of some of the candidate explanatory variables. However, Model 1 shows a remarkable result. Stage 1 reports that for the firm-month observations included in Model 1, for which all ten remaining candidate explanations are available, IVOL is not significantly negatively related to future excess returns, which is in contradiction with the findings of Table 5 and Table 6 when considering the entire U.S. sample. This suggests that in the remaining sample for Model 1 the IVOL puzzle seems to not occur, and therefore no reliable conclusions can be drawn from the explanatory power of the candidate variables, as the Model 1 sample is most likely biased, due to the relatively small sample size. The stocks selected in Model 1 most likely do not accurately represent the full U.S. sample, because Model 1 only includes stocks that have no missing values for all candidate variables, leaving only a small group of stocks to analyze.

To avoid the exclusion sample bias of model 1, the candidate variable Dispersion is omitted in Model 2 to almost double the sample size to an average of 2587 firm observations per month. Stage 1 in Model 2 shows a strongly significant negative relationship between IVOL and subsequent excess returns, which is in line with the results of Table 5 and Table 6. The total fraction of the IVOL puzzle that can be explained by the remaining nine candidate variables is 40.77%, leaving 59.23% of the IVOL puzzle unexplained, which is significant at the 1% level. Spread shows the biggest contribution of 11.55%, although the fraction is not significantly different form zero. Following Spread, is SUE able to explain 10.52% of the puzzle, which is significant at the 1% level. In addition, Skew has a contribution of 5.79%, significant at the 1% level, and LagRet can explain 5.01% of the puzzle, which, however, does not differ significantly from zero. The other five explanatory variables (Coskew, E(IS), Amihud, ZeroRet and MEBE) are together able to explain 7.90% of the IVOL puzzle.

To increase the sample size even further, the candidate variable E(IS) is also omitted in Model 3 to increase the sample size to an average of 3124 firm observations per month. Stage 1 shows that the coefficient for IVOL (-25.163) decreases substantially compared to model 2. However, the total fraction of the IVOL puzzle being explained in Model 3 decreases slightly to 36.91%, leaving 63.09% of the IVOL puzzle unexplained, which is significant at the 1% level. Spread can again explain the largest fraction of the puzzle with a contribution of 19.31%, followed by SUE (6.88%), Skew (4.70%) and MEBE (2.67%), which are all significant at the 1% level. Interestingly, the contribution for LagRet decreases to 0.87% in Model 3 and does not significantly differ from zero. The other candidate variables (Coskew, Amihud and ZeroRet) can together explain 2.48% of the puzzle.

Based on the results of Model 2 and Model 3 from Table 7, the variables related to the lottery preferences of investors contribute to approximately 7-10% in explaining the IVOL puzzle, which is considerably less than found in the study by Hou and Loh (2016). This may be explained by the fact that this study extends the data set of Hou and Loh (2016) with an additional ten years of recent data. The candidate variables based on market frictions can approximately explain 18-21%, with Spread as the main contributor, what was also found by Anneart et al. (2022), and the candidate explanations related to other explanations can explain approximately 10-14% of the IVOL puzzle. For the other explanations, it must be taken into account that Dispersion is not included, while Dispersion did explain a significant part of the puzzle in Table 6. However, the unexplained fraction is about 60-63% and therefore still very large.

To gain a better understanding of the explanatory power of the candidate variables, the multivariate analysis of Model 2 and 3 from Table 7 is repeated for subsamples of stocks, where the IVOL puzzle seems to occur more strongly. This makes it possible to check whether there is a difference between the explanatory power of the candidate variables in the sub-sample and in the whole sample. The subsample analysis will be performed for stocks with a price above $5, companies with a low credit rating (the lowest three credit rating deciles) and companies with high leverage (the highest three leverage deciles). Table 8 reports the decomposition results for the sub-sample analysis. In order to not make Table 8 too extensive and cluttered, only the results of stage 3 are shown. The table shows that the considered candidate variables can explain the IVOL puzzle for 42-53% among stocks with a price higher than $5, 70% for stocks with a low credit rating and 36-48% for firms with high leverage, compared to the 37-40% for the whole U.S. sample. The results of Table 8 indicate that the average explained part related to the lottery preferences of investors across the three sub-samples is approximately 6% and 10% for Model 2 and Model 3, respectively. The explanatory candidate variables related to market frictions can explain on average 32% for Model 2 and 28% for Model 3 and the candidate explanations related to the other explanations can explain on average 15% of the IVOL puzzle for Model 2 and Model 3. The residual component accounts on average for 47% of the puzzle in Model 2 and Model 3.

In conclusion, the results of both Table 7 and Table 8 show that the candidate variables related to market frictions contribute the most in explaining the puzzle among U.S. stocks, followed by the variables related to other explanations and the candidate variables related to investors' lottery preferences. These results are not completely in accordance with the results of Hou and Loh (2016), Anneart et al. (2022) and Zhong (2018), who found that investors' lottery preferences and market frictions can largely explain the puzzle. However, the results of sections 5.2 and 5.3 show that the candidate variables related to investor lottery preferences seem to contribute the least to explaining the puzzle, compared to the explanations related to market frictions and the other explanations. Therefore, hypothesis 1 is rejected since hypothesis 1 stated that explanations based on investors' lottery preferences and market frictions can best explain the puzzle. Although the candidate variables related to market frictions and the other explanations seem to explain an important part of the IVOL puzzle, the unexplained (residual) component remains significant and relatively large in all analyses in sections 5.2 and 5.3, which is in line with the results of Hou and Loh (2016), Anneart et al. (2022) and Zhong (2018). Therefore, hypothesis 2 is accepted, as hypothesis 2 stated that the candidate explanations from the literature are not able to explain the whole IVOL puzzle.

Table 7: Decomposing the IVOL puzzle using multiple candidate explanation at a time for U.S. stocks

*The table reports the results of the decomposition analysis proposed by Hou and Loh (2016) using multiple candidate explanation at the same time. The relationship between IVOL and monthly excess returns is decomposed into a number of fractions each related to a candidate variable and a residual component using firm-level Fama-MacBeth cross-sectional regressions. Time-series averages of the coefficients are multiplied by 100 and the associated time-series t-statistics are reported in parentheses. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. The several candidate explanatory variables are estimated as described is sections 3.3, 3.4 and 3.5. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.*

Stage	Description	Variable	Model 1			Model 2			Model 3		
			Coeff.	Fraction	t-stat	Coeff.	Fraction	t-stat	Coeff.	Fraction	t-stat
1	Excess returns on	Intercept	1.072***		(5.03)	1.215***		(5.71)	1.347***		(6.29)
	IVOL	IVOL	-8.015		(-1.59)	-15.364***		(-4.06)	-25.163***		(-6.47)
2	IVOL on candidate	Intercept	1.567***		(54.92	1.565***		(56.75)	2.106***		(79.56)
	variables	Skew	0.131***		(15.88)	0.231***		(28.09)	0.284***		(31.99)
		Coskew	0.713***		(5.91)	0.572***		(7.14)	0.580***		(7.37)
		E(IS)	0.699***		(12.17)	1.151***		(22.61)			
		LagRet	0.349***		(3.98)	0.713***		(7.99)	0.556***		(5.65)
		Amihud	-0.150***		(-7.53)	0.001		(0.71)	-0.045		(-1.47)
		ZeroRet	-1.748***		(-20.27)	-2.524***		(-31.74)	-2.209***		(-22.55)
		Spread	58.434***		(36.66)	35.506***		(55.17)	51.461***		(70.77)
		Dispersion	0.452***		(44.42)						
		SUE	-0.050***		(-33.45)	-0.063***		(-41.42)	-0.088***		(-41.85)
		MEBE	0.006***		(10.89)	0.007***		(13.39)	0.008***		(15.51)
		Avg adj R^2	35.1%			43.2%			38.6%		
3	Decompose stage 1	Skew	-0.508	6.34%**	(2.41)	-0.890	5.79%***	(4.94)	-1.183	4.70%***	(5.45)
	IVOL coefficient	Coskew	-0.106	1.32%	(0.29)	-0.153	1.00%	(0.88)	-0.543	2.16%	(1.51)
		E(IS)	0.380	-4.74%	(-0.28)	-0.429	2.79%	(0.37)			
		LagRet	-0.220	2.74%	(0.29)	-0.770	5.01%	(1.40)	-0.219	0.87%	(0.34)
		Amihud	-0.020	0.25%	(0.12)	-0.080	0.52%	(0.47)	-0.052	0.21%	(0.23)
		ZeroRet	0.346	-4.32%	(-0.73)	-0.079	0.51%	(0.22)	-0.028	0.11%	(0.07)
		Spread	-2.982	37.21%*	(1.78)	-1.775	11.55%	(1.23)	-4.860	19.31%***	(2.67)
		Dispersion	-0.815	10.17%**	(2.32)						
		SUE	-2.038	25.43%***	(11.06)	-1.615	10.52%***	(12.80)	-1.730	6.88%***	(12.03)
		MEBE	-0.598	7.45%***	(4.28)	-0.473	3.08%***	(5.52)	-0.672	2.67%***	(4.44)
		Residual	-1.455	18.15%	(0.48)	-9.100	59.23%***	(3.79)	-15.876	63.09%***	(5.49)
		Total	-8.015	100%	(-1.59)	-15.364***	100%	(-4.06)	-25.163***	100%	(-6.47)
		Avg # firms/mth	1479.7			2586.7			3123.5		
		Num obs	671,803			1,174,355			1,499,268		

Table 8: Decomposing the IVOL puzzle using multiple candidate explanation at a time for U.S. stocks: subsample analysis

*The table reports the results of the decomposition analysis proposed by Hou and Loh (2016) using multiple candidate explanation at the same time for several subsamples: stocks with a price above $5 (Panel A), companies with a low credit rating (Panel B) and companies with high leverage (Panel C). Time-series averages of the coefficients are multiplied by 100 and the associated time-series t-statistics are reported in parentheses. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. The several candidate explanatory variables are estimated as described is sections 3.3, 3.4 and 3.5. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.*

Variable	Model 2			Model 3		
	Coeff.	Fraction	t-stat	Coeff.	Fraction	t-stat
Panel A: Price > $5						
Skew	-0.730	4.11%***	(3.60)	-0.893	3.18%***	(4.20)
Coskew	-0.388	2.19%	(1.14)	-0.680	2.42%	(1.49)
E(IS)	0.217	-1.22%	(-0.20)			
LagRet	-1.390	7.83%*	(1.89)	-0.704	2.51%	(0.86)
Amihud	-0.175	0.99%	(0.91)	-0.043	0.15%	(0.19)
ZeroRet	0.227	-1.28%	(-0.40)	0.372	-1.33%	(-0.57)
Spread	-4.175	23.52%***	(2.94)	-6.783	24.18%***	(4.07)
Dispersion						
SUE	-2.213	12.47%***	(13.53)	-2.288	8.16%***	(12.85)
MEBE	-0.733	4.13%***	(4.88)	-0.902	3.22%***	(4.45)
Residual	-8.388	47.26%***	(2.89)	-16.131	57.51%***	(4.65)
Total	-17.748***	100%	(-4.37)	-28.053***	100%	(-6.52)
Avg # firms/mth	2190.4			2551.7		
Num obs	994,439			1,224,809		
Panel B: Low credit rating						
Skew	-1.318	9.95%**	(2.00)	-1.042	8.16%*	(1.71)
Coskew	-1.384	10.45%	(1.03)	-1.591	12.46%	(1.23)
E(IS)	2.060	-15.55%	(-1.25)			
LagRet	-0.926	6.99%*	(0.74)	-0.863	6.76%	(0.68)
Amihud	0.337	-2.54%	(-0.21)	0.480	-3.76%	(-0.33)
ZeroRet	-0.589	4.45%	(0.58)	-0.120	0.94%	(0.12)
Spread	-5.011	37.84%	(1.17)	-3.140	24.59%	(0.74)
Dispersion						
SUE	-1.454	10.98%***	(4.04)	-1.664	13.03%***	(4.73)
MEBE	-1.008	7.61%**	(2.35)	-0.976	7.64%***	(3.00)
Residual	-3.951	29.82%	(0.86)	-3.851	30.16%	(0.82)
Total	-13.244*	100%	(-1.67)	-12.767*	100%	(-1.68)
Avg # firms/mth	331.6			359.5		
Num obs	124,013			134,451		
Panel C: High leverage						
Skew	-0.600	3.50%**	(2.31)	-0.497	1.93%	(0.83)
Coskew	-0.573	3.34%**	(1.99)	-0.763	2.96%	(1.53)
E(IS)	-0.029	0.17%	(0.02)			
LagRet	0.155	-0.90%	(-0.29)	0.258	-1.00%	(-0.34)
Amihud	-0.531	3.10%	(1.57)	0.100	-0.39%	(-0.06)
ZeroRet	0.131	-0.76%	(-0.36)	0.546	-2.12%	(-0.81)
Spread	-2.987	17.45%	(1.54)	-8.937	34.68%***	(2.80)
Dispersion						
SUE	-1.431	8.36%***	(9.18)	-1.574	6.11%***	(8.20)
MEBE	-0.362	2.11%**	(2.51)	-1.392	5.40%**	(2.29)
Residual	-10.894	63.63%***	(4.10)	-13.509	52.43%***	(4.68)
Total	-17.121***	100%	(-4.12)	-25.767***	100%	(-5.56)
Avg # firms/mth	776.3			890.1		
Num obs	352,427			427,257		

5.4 Examining the difference of the explanatory power of the candidate explanations through time

This section examines the possible differences in the explanatory power of the candidate variables through different periods of time for the U.S. sample. Since Fu (2009), Brandt et al. (2010) and the results of section 5.1 already indicated that the level of IVOL and the IVOL puzzle tend to differ over time, it is also interesting to examine to what extent the explanatory power of the different candidate explanations to predict negative returns associated with IVOL also differs over time. The entire U.S. sample is divided into four sub-samples for the periods 1982-1990, 1991-1999, 2000-2010 and 2011-2022. The coefficients of the various candidate variables of stage 3 from the decomposition analysis of Table 7 are compared between the different sub-samples. In order to detect a possible significant difference in the coefficients of the candidate variables between the different subsamples, a Z-test initiated by Clogg et al. (1995) is used, which is discussed in section 3.6. For each sub-sample the candidate variables included in Model 3 are compared with each other, since the results in section 5.3 were most likely biased when Dispersion was included and E(IS) is not available for a relatively large number of observations and seems to explain little of the puzzle in the analyses discussed earlier.

Table 9 reports the results of analysis described above. The table only shows the results of stage 3 to ensure space. Panel A reports the results of 1982-1990 compared to 1991-1999. It can be seen that the Total IVOL estimate for the period 1982-1990 shows a rather high coefficient (-43,043), which is also highly significant, compared to the coefficient for the period 1991-1999 (-6.502), which shows no significant coefficient. The Z statistic (-3.25) shows that there is a significant difference between these two coefficients, indicating that IVOL in the period 1982-1990 had a much more negative impact on subsequent returns than in the period 1991-1999. This therefore initiates that the IVOL puzzle appears to be time-varying and that the magnitude of the puzzle also differs over time, which is in line with Brandt et al. (2010) and the results of section 5.1. In addition, the Z statistics of panel A indicate a significant difference in the regression coefficient of the candidate variables Skew, Spread, and the residual component. This suggests that the part of these candidate variables that is related to IVOL is significantly more negatively related to subsequent returns in one time period than in the other. However, these significant differences should be interpreted with caution, as the coefficients for the candidate variables at Stage 3 are determined relative to the Total IVOL coefficient. The significant difference between the Total IVOL coefficient for the two subsamples examined could potentially lead to biased results for the Z statistics, since there is a high probability that there may already be a significant difference between the coefficients by construction.

Panel B of Table 9 reports the results of 1991-1999 compared to 2000-2010. The sub-sample 2000-2010 shows a Total IVOL coefficient of -22.151, significant at the 1% level. However, the Z statistic shows that there is no significant difference in Total IVOL between the two time periods

(1.45). In addition, the Z statistic for CoSkew (1.91) and SUE (1.87) indicate a significant difference in the coefficient and the relation with subsequent stock returns of the candidate variables between the two sub-samples. The significant Z statistic and observable difference in fractions for the candidate variables therefore initiates a difference in the explanatory power of the candidate variables over time.

Lastly, panel C of Table 9 shows the results of the sub-sample 2000-2010 compared to 2011-2022. There does not seem to be a significant difference in the Total IVOL coefficient between the two subsamples, as the Z statistic shows a value of 0.72. However, Panel C does indicate a significant difference in the coefficients between both sub-samples for the candidate variables CoSkew and MEBE (Z statistic of -1.88 and -2.25, respectively). These significant negative Z statistics therefore indicate that the variables Coskew and MEBE were more negatively related to future returns and explained more of the IVOL puzzle in the period 2000-2010 than 2011-2022.

In conclusion, the results from Table 9 show several interesting results. Panel A shows that there is a significant difference between the compared Total IVOL coefficients, indicating that the magnitude of the IVOL puzzle appears to vary over time. This finding is therefore in accordance with the paper by Brandt et al. (2010) and the results previously found in this study. In addition, each panel shows a significant Z-statistic for one of the examined candidate variables, indicating that the explanatory power of the candidate explanations tends to differ over time. In other words, some explanations seem to be significantly stronger/weaker related to the negative returns associated with IVOL in one period than in another, which may indicate that the IVOL puzzle in each period must also be explained by a different explanation, since the explanatory power is not constant over time and an ultimate explanation for the puzzle is most likely dimensionally bound. Based on the results from Table 3 and Table 9, the study concludes that the explanatory power of the candidate explanations and the occurrence of the IVOL puzzle tends to differ over time and hence per dimension. Therefore, hypothesis 3 is accepted.

Table 9: Decomposing the IVOL puzzle using multiple candidate explanation at a time for U.S. stocks through time.

*The table reports the results of the decomposition analysis proposed by Hou and Loh (2016) using multiple candidate explanation at the same time for four time periods, where the coefficients for the candidate variables are compared between the different periods using a Z-test initiated by Clogg et al. (1995). Panel A compares 1982-1990 to 1991-1999, panel B compares 1991-1999 to 2000-2010 and panel C compares 2000-2010 to 2011-2022. The Z-statistic for each comparison is reported in the last column. Time-series averages of the coefficients are multiplied by 100 and the associated time-series t-statistics are reported in parentheses. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. The several candidate explanatory variables are estimated as described is sections 3.3, 3.4 and 3.5. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.*

Variable	1982-1990			1991-1999			
	Coeff.	Std. Error	Fraction	Coeff.	Std. Error	Fraction	Z-statistic
Panel A: 1982-1990 compared to 1991-1999							
Skew	-1.926	0.538	4.48%***	-0.568	0.334	8.74%*	-2.15
Coskew	-2.284	1.646	5.31%	0.199	0.245	-3.06%	-1.49
LagRet	1.048	1.011	-2.44%	-0.010	0.967	0.15%	0.76
Amihud	-0.198	0.713	0.46%	-0.346	0.447	5.32%	0.18
ZeroRet	1.448	1.451	-3.36%	0.388	0.978	-5.96%	0.61
Spread	-15.399	6.043	35.78%**	-0.476	4.143	7.32%	-2.04
SUE	-1.143	0.216	2.65%***	-1.224	0.181	18.81%***	0.29
MEBE	-1.502	0.613	3.49%**	-0.723	0.195	11.11%***	-1.21
Residual	-23.087	5.410	53.64%***	-3.742	4.607	57.55%	-2.72
Total	-43.043***	8.952	100%	-6.502	6.790	100%	-3.25
Variable	1991-1999			2000-2010			
	Coeff.	Std. Error	Fraction	Coeff.	Std. Error	Fraction	Z-statistic
Panel B: 1991-1999 compared to 2000-2010							
Skew	-0.568	0.334	8.74%*	-1.074	0.475	4.85%**	0.87
Coskew	0.199	0.245	-3.06%	-0.628	0.356	2.84%*	1.91
LagRet	-0.010	0.967	0.15%	-1.004	1.801	4.53%	0.49
Amihud	-0.346	0.447	5.32%	-0.121	0.382	0.55%	-0.38
ZeroRet	0.388	0.978	-5.96%	-1.482	0.900	6.69%*	1.41
Spread	-0.476	4.143	7.32%	-1.721	2.740	7.77	0.25
SUE	-1.224	0.181	18.81%***	-1.770	0.230	7.99%***	1.87
MEBE	-0.723	0.195	11.11%***	-0.664	0.239	3.00%***	-0.19
Residual	-3.742	4.607	57.55%	-13.687	6.453	61.78%**	1.25
Total	-6.502	6.790	100%	-22.151***	8.386	100%	1.45

Table 9

Continued.

Variable	2000-2010			2011-2022			
	Coeff.	Std. Error	Fraction	Coeff.	Std. Error	Fraction	Z-statistic
Panel C: 2000-2010 compared to 2011-2022							
Skew	-1.074	0.475	4.85%**	-1.242	0.374	4.15%***	0.28
Coskew	-0.628	0.356	2.84%*	0.158	0.219	-0.53%	-1.88
LagRet	-1.004	1.801	4.53%	-0.509	0.930	1.70%	-0.24
Amihud	-0.121	0.382	0.55%	0.333	0.337	-1.11%	-0.89
ZeroRet	-1.482	0.900	6.69%*	-0.005	0.146	0.02%	-1.62
Spread	-1.721	2.740	7.77	-3.919	1.943	13.10%**	0.65
SUE	-1.770	0.230	7.99%***	-2.476	0.377	8.28%***	1.59
MEBE	-0.664	0.239	3.00%***	-0.078	0.102	0.26%	-2.25
Residual	-13.687	6.453	61.78%**	-22.170	5.700	74.13%***	0.98
Total	-22.151***	8.386	100%	-29.908***	6.705	100%	0.72

5.5 Examining the difference of the explanatory power of the candidate explanations between developed and emerging stock markets

This section investigates the possible differences in the explanatory power of the candidate variables between developed and emerging markets. Kohers et al. (2006) have shown that emerging markets are more volatile and the results in section 5.1 have already showed that there seems to be a difference in the occurrence of the IVOL puzzle between a developed and an emerging stock market, which may be explained by the difference in the fraction of retail investors on the relevant stock market. It is therefore interesting to investigate whether the explanatory power of the candidate variables for the negative returns related to IVOL differs between developed and emerging markets. The coefficients of stage 3 from the decomposition analysis are compared. The available candidate variables for both the U.S. sample (developed) and the Chinese sample (emerging) are compared for the period 2000-2021, using a Z-test as done in section 5.4 and discussed in section 3.6.

Table 10 reports the results of the comparison of the two samples. Repeatedly, Table 10 only shows the results of stage 3 of the decomposition analyses to ensure space. Panel A shows the results of the comparison, taken all the available candidate variables for both samples into account. The Total IVOL coefficient for the U.S. sample shows a value of -20.731% and the Chinese sample shows a value of -67.267%, both significant at the 1% level, indicating that IVOL is significantly negative related with subsequent stock returns for both countries. However, when both coefficients are compared, the table shows a significant Z statistic of 4.28, demonstrating that the magnitude of the IVOL puzzle seems to be higher in China than in the matching U.S. period. This result is consistent with the results and conclusions of section 5.1, which also found that the occurrence and the strength of the IVOL puzzle tend to differ between China and the US. Furthermore, panel A show several significant Z statistics for the candidate variables coefficients, indicating that the part of the relevant candidate variables that is related to IVOL is significantly different in relation to subsequent returns between the developed and the emerging stock market. However, as also observed in panel A of table 9, these significant differences need to be interpreted with caution, due to the significant difference in the Total IVOL coefficient, which possibly already causes a difference in the explanatory power of the candidate variables. There also seems to be a notable difference in the fractions that the candidate variables can explain of the IVOL puzzle between the two samples, which also indicates a difference in explanatory power of the candidate variables. For example, the variable LagRet explains about a third of the puzzle in the Chinese sample, while LagRet in the U.S. sample explains only 2.99%. These differences could be explained by the difference in the fraction of retail traders active on the relevant stock market, as retail traders seem to act differently than institutional traders.

Panel B drops the variable E(IS) to extent the Chinese subsample. However, the results are quite similar to the results of panel A. For instance, the Total IVOL coefficient for the Chinese sample (-63,733) is also significantly different from the coefficient for the U.S. sample (-32.759) with a Z statistic of 3.28. Additionally, panel B also shows several significant Z statistics for the relevant candidate variables coefficients, suggesting a significant difference in the explanatory power of these variables between the developed and the emerging stock market. The fractions that can be explained by the candidate explanations also seem to differ between the two samples, which is about the same as in panel A of Table 10.

To conclude, Table 10 shows that the magnitude of the IVOL puzzle seems to be stronger on the Chinese stock market than on the U.S. stock market, which is in line with the results of section 5.1 since a difference in occurrence was already found between both samples. Furthermore, Table 10 shows several significant Z statistics, which indicate a significant difference in the explanatory power of the candidate variables to predict future returns between an emerging and a developed market. This suggests that in a developed stock market the IVOL puzzle is most likely to be explained by different explanations than in an emerging stock market, since the explanatory power is not constant between developed and emerging markets and an ultimate explanation for the puzzle is most likely dimensionally bound, which was also concluded in section 5.4. Moreover, Table 10 shows that the fractions of the IVOL puzzle that can be explained by the candidate variables also seem to differ considerably between developed and emerging markets, which may be explained by the fractions of retail traders active on the relevant stock market. However, based on the results of Table 4 and Table 10, the study concludes that the occurrence of the IVOL puzzle and the explanatory power of the candidate variables to predict future returns, tends to differ between a developed and an emerging market and therefore differs per dimension. Therefore, hypothesis 4 is also accepted.

Table 10: Decomposing the IVOL puzzle using multiple candidate explanation at a time: difference between a developed and an emerging stock market

*The table reports the results of the decomposition analysis proposed by Hou and Loh (2016) using multiple candidate explanation at the same to test for the difference between the U.S. stock market (developed) and the Chinese stock market (emerging), where the coefficients for the candidate variables are compared between the different periods using a Z-test initiated by Clogg et al. (1995). The Z-statistic for each comparison is reported in the last column. Time-series averages of the coefficients are multiplied by 100 and the associated time-series t-statistics are reported in parentheses. IVOL is estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns on the Fama and French (1993) 3 factors. The several candidate explanatory variables are estimated as described is sections 3.3, 3.4 and 3.5. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.*

Variable	U.S. Sample 2000-2021			Chinese Sample 2000-2021			
	Coeff.	Std. Error	Fraction	Coeff.	Std. Error	Fraction	Z-statistic
Panel A							
Skew	-0.850	0.259	4.10%***	-8.532	2.715	12.68%***	2.82
Coskew	-0.112	0.263	0.54%	8.236	3.716	-12.24%**	-2.24
E(IS)	-1.852	2.072	8.94%	0.199	1.613	-0.30%	-0.78
LagRet	-0.619	0.855	2.99%	-20.371	5.263	30.28%***	3.70
Amihud	-1.269	0.507	6.12%**	0.989	1.328	-1.47%	-1.59
ZeroRet	-0.412	0.300	1.99%	-0.961	0.717	1.43%	0.71
MEBE	-0.218	0.098	1.05%**	-7.446	1.925	11.07%***	3.75
Residual	-15.399	3.509	74.28%***	-39.390	7.049	58.55%***	3.05
Total	-20.731***	5.498	100.00%	-67.276***	9.374	100.00%	4.28
Panel B							
Skew	-1.598	0.372	4.88%***	-8.200	1.611	12.87%***	3.99
Coskew	-0.177	0.199	0.54%	6.459	2.539	-10.13%**	-2.61
E(IS)							
LagRet	-0.811	1.108	2.48%	-15.741	3.849	24.70%***	3.73
Amihud	-1.349	0.638	4.12%**	-0.583	0.952	0.91%	-0.67
ZeroRet	-0.463	0.281	1.41%*	-1.763	0.567	2.77%***	2.06
MEBE	-0.127	0.116	0.39%	-5.618	1.293	8.81%***	4.23
Residual	-28.235	4.573	86.19%***	-38.287	5.443	60.07%***	1.41
Total	-32.759***	5.452	100.00%	-63.733***	7.723	100.00%	3.28

6. Robustness Check

In section 5.1, this study investigated the occurrence of the IVOL puzzle among U.S. and Chinese stocks, using a portfolio strategy indicated by Ang et al. (2006), where stocks were sorted on their historical IVOL estimate relative to the Fama and French (1993) 3 factors (FF-3). However, since the FF-3 model contains relatively few risk factors compared to other asset pricing models, the FF-3 model may not provide the most accurate estimate for the IVOL estimates. Therefore, to check the robustness of the results of section 5.1, this section will repeat the value-weighted 1/0/1 trading strategy, as discussed in section 3.2, but with estimated IVOL estimates relative to the Carhart (1997) 4-factor model, the Fama and French (2015) 5-factor model (FF-5), and the Fama and French (2018) 6-factor model (FF-6). Based on this analysis, it can be determined whether the IVOL puzzle still occurs when IVOL is estimated relative to asset pricing models other than the FF-3 factor model.

To be able to calculate the IVOL estimates relative to the different factor models, the following time-series regressions will first be conducted:

$$r_t^i - rf_t = a_t^i + \beta_{Mkt}^i Mkt_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \beta_{UMD}^i UMD_t + \varepsilon_t^i, \tag{16}$$

$$\begin{aligned} r_t^i - rf_t = a_t^i &+ \beta_{Mkt}^i Mkt_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \beta_{RMW}^i RMW_t \\ &+ \beta_{CMA}^i CMA_t + \varepsilon_t^i, \end{aligned} \tag{17}$$

$$\begin{aligned} r_t^i - rf_t = a_t^i &+ \beta_{Mkt}^i Mkt_t + \beta_{SMB}^i SMB_t + \beta_{HML}^i HML_t + \beta_{RMW}^i RMW_t \\ &+ \beta_{CMA}^i CMA_t + \beta_{UMD}^i UMD_t + \varepsilon_t^i, \end{aligned} \tag{18}$$

where equation (16) corresponds to the Carhart (1997) 4-factor model, equation (17) to the Fama and French (2015) 5-factor model and equation (18) to the Fama and French (2018) 6-factor model. UMD_t is the return difference of past winning stocks relative to past losing stocks on day t, RMW_t is the difference in return between high (robust) operating profitability company stocks and low (weak) operating profitability company stocks on day t and CMA_t is the return difference of company stocks that is invest conservatively relative to company stocks that invest aggressively on day t. These daily factors are also obtained from the Kenneth R. French Data Library. The other factors have been discussed earlier in section 3.1 and are not differently estimated for this analysis. All the factor loadings, used to compute the residual estimate (ε_t^i) for firm i on day t for the different models, are, as done in section 3.1, estimated using an expanding monthly window. The three IVOL estimates for a specific month are defined as the standard deviation over the daily residuals from equation (16), (17) and (18) at the end of the relevant month using the following equation:

$$IVOL_t^i = \sqrt{Var(\varepsilon_t^i)}. \tag{19}$$

As in section 3.1, a minimum of ten daily observations is required to obtain a monthly IVOL estimate.

Table 11 reports the results of the robustness check. Panel A shows the results of the 1/0/1 strategy for the U.S. sample and Panel B shows the results for the Chinese sample for portfolios sorted on the three different estimates of IVOL. The results in Panel A show very similar results to the results of Table 2 in section 5.1 for all three different estimates of IVOL. All the long-short portfolios (P5-P1) show an negative average monthly excess return, significant at the 1% level. The long-short portfolios sorted by IVOL relative to the Carhart 4 factor model, FF-5 factor model and FF-6 factor model show an average monthly return of -1.22%, -1.27% and -1.26%, respectively, which are nearly equal to the average monthly return of the value-weighted long-short portfolio from Table 2 (-1.28%). Panel A's findings suggest that the IVOL puzzle still occurs in the investigated U.S. sample when IVOL is estimated relative to asset pricing models other than the Fama and French 3 factors.

Panel B reports also similar results for the three different estimates of IVOL compared to the results of the first row in panel B of Table 4. Also for the Chinese sample, all the long-short portfolios report an negative average monthly excess return, significant at the 1% level (-1.06% for the Carhart 4 factor model and -1.07% for both FF-5 factor and FF-6 factor model). Also these returns are similar to average monthly return of the value-weighted long-short portfolio for the Chinese sample in Table 4 (-1.05%). Additionally, the results of Panel B indicate that the IVOL puzzle still occurs in the examined Chinese sample when asset pricing models other than the Fama and French 3 factor model are used to estimate IVOL.

To conclude, the analysis in this section shows that the IVOL puzzle still occurs among the whole investigated U.S. and Chinese stocks samples, regardless of which asset pricing model is used to estimate IVOL, as all the long-short portfolios report significantly negative returns. Therefore, the analysis confirms that the main results, regarding the occurrence of the IVOL puzzle in the whole investigated U.S. and Chinese sample, are robust to using other asset pricing models to estimated IVOL.

Table 11: Portfolios sorted by IVOL relative to Carhart 4 factor model, Fama and French 5 factor model and Fama and French 6 factor model for U.S and Chinese stocks

*The table reports the results from value-weighted quintile portfolios sorted on the IVOL estimates of the past month. The IVOL estimates are estimated as the standard deviation of the monthly residuals from a regression of daily stock excess returns relative to the Carhart (1997) 4 factors, Fama and French (2015) 5 factors and Fama and French (2018) 6 factors. Portfolio P1 (P5) is the portfolio of stocks with the lowest (highest) idiosyncratic volatilities. The Mean Excess Returns are measured in monthly percentages for all three strategy's. The t-statistics are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively. The sample period is January 1982 to December 2022 for U.S. stocks and January 2000 to December 2021 for Chinese stocks.*

Portfolio Rank	Mean Excess Returns IVOL relative to Carhart	Mean Excess Returns IVOL relative to FF-5	Mean Excess Returns IVOL relative to FF-6
Panel A: Value-weighted portfolios sorted by IVOL for U.S. stocks 1982-2022			
P1	0.81%	0.82%	0.84%
P2	0.77%	0.75%	0.72%
P3	0.73%	0.67%	0.68%
P4	0.38%	0.43%	0.42%
P5	-0.41%	-0.45%	-0.42%
P5-P1	-1.22%***	-1.27***	-1.26***
	(-4.42)	(-4.55)	(-4.54)
Panel B: Value-weighted portfolios sorted by IVOL for Chinese stocks 2000-2021			
P1	0.72%	0.71%	0.72%
P2	0.78%	0.90%	0.87%
P3	0.86%	0.85%	0.88%
P4	0.56%	0.57%	0.56%
P5	-0.34%	-0.37%	-0.35
P5-P1	-1.06%***	-1.07%***	-1.07%***
	(-3.27)	(-3.33)	(-3.31)

7. Conclusion and Discussion

The main goal of this study was to investigate the following research question: *To what extent do the various candidate explanations for the idiosyncratic volatility puzzle explain the idiosyncratic volatility puzzle nowadays*? To answer the research question, the IVOL puzzle has been approached from different perspectives, where this study was not expected to be able to explain the whole IVOL puzzle but to provide new insights into explaining the IVOL puzzle.

First, the study investigated the occurrence of the IVOL puzzle in the U.S. based on a portfolio trading strategy following the methodology of Ang et al. (2006) and examined the occurrence of the puzzle over time and the difference in occurrence between a developed and an emerging market. The results showed significantly negative returns on the long-short portfolios, which conclude that the IVOL puzzle occurs in the U.S. sample, which is in accordance with the results found by Ang et al. (2006). In addition, the rest of the results showed that the occurrence of the puzzle differs by time period, indicating a time-varying occurrence of the puzzle, which is in line with Fu (2009) and Brandt et al. (2010). The portfolio trading strategy also showed that there is a difference between the occurrence of the IVOL puzzle between the Chinese stock market (emerging) and the U.S. stock market (developed), which may be explained by the fraction of retail traders active on the relevant stock market or the differences in short-sell constraints, or both. Furthermore, the robustness test revealed that the main findings, concerning the occurrence of the IVOL puzzle across the entire U.S. and Chinese sample examined, are robust to using other asset pricing models to estimate IVOL.

Next, the decomposition method of Hou and Loh (2016) was used to examine what fraction of the IVOL puzzle could be explained by the several candidate explanations from the literature and to investigate if the candidate explanations could explain the whole IVOL puzzle together in the U.S. sample. The study finds that most of the existing explanations for the puzzle can only explain less than 10% of the IVOL puzzle, which is in line with the results of Hou and Loh (2016). However, when all the explanations are taken together, the study finds that explanations based on market frictions and variables related to other explanations can explain a considerable fraction of the puzzle, whereas Hou and Loh (2016) found that explanations based on investors' lottery preferences could explain a large part of the puzzle. Explanations based on market frictions together explain 18-21%, the candidate explanations related to other explanations explain approximately 10-14% and the explanations related to the lottery preferences of investors contribute to about 7-10% in explaining the IVOL puzzle for U.S. stocks in this study. However, even when all investigated explanations considered, the biggest part of the IVOL puzzle remains unexplained.

Finally, the explanatory power of the candidate explanations for the IVOL puzzle was examined to detect a possible difference over time in the U.S. sample and to detect a possible difference between a developed and an emerging stock market, making use of a Z-test initiated by Clogg et al. (1995). The results show that the magnitude of the IVOL puzzle tends to differ over time and between a developed and an emerging stock market. Furthermore, the several significant Z statistics tend to indicate that the explanatory power of the part of the candidate explanations that is related to IVOL to predict subsequent stock returns differs significantly through time and between a developed and emerging stock market. Also, the fractions of the IVOL puzzle that can be explained by the candidate variables varies considerably over time and between developed and emerging stock markets.

To answer the main research question: *To what extent do the various candidate explanations for the idiosyncratic volatility puzzle explain the idiosyncratic volatility puzzle nowadays*?, this study provides significant evidence that the idiosyncratic volatility puzzle still occurs nowadays but tends to differ in occurrence through time and between developed and emerging markets. Furthermore, the several explanations from the literature for the IVOL puzzle are unable to explain the complete puzzle using more recent data, where explanations based on market friction contribute the most in explaining the puzzle when more recent data is considered. Additionally, the explanation for the IVOL puzzle related to the bid-ask spread of a stock, shows the most promising results in explaining the puzzle nowadays. The study observes significant differences in the explanatory power of the explanations through different dimensions, suggesting that the IVOL puzzle is most likely to be explained by different explanations for each dimension. These differences may also indicate the explanation for the complete IVOL puzzle being dimensionally bound, indicating that the explanation of the entire IVOL puzzle will most likely differ per dimension and there will probably not be an explanation that covers the entire puzzle for each dimension.

This study contributes to the literature as it is one of the first studies to investigate the explanatory power of the several candidate explanations between different time periods and between developed and emerging stock markets. Since the study shows that the explanatory power of the variables for the IVOL puzzle tends to significantly differ per dimension, both traders and researchers can take this into account for in the future, where other researchers can build forward on these new insights to possibly find other explanations for the puzzle. Furthermore, the study investigates the IVOL puzzle for a more recent time frame and uses an explanation that has been proposed after the paper of Hou and Loh (2016) was published.

Unfortunately, the study also comes with some limitations that should be discussed. First, the study draws conclusions for developed and emerging markets based on two countries compared. It may be that the U.S. and China give a distorted picture of the comparison between developed and emerging countries, as a comparison between more countries could potentially lead to a different result. Second, a z-test initiated by Clogg et al. (1995) was used to assess the difference in explanatory power of the candidate explanations through different dimensions. However, the z statistics could be biased, since the stage 3 coefficients are set relative to the Total IVOL coefficient, which could cause a significant difference in the coefficients by construction. Finally, this study looks at the effect of IVOL on subsequent excess returns on a stock, whereas studies such as Hou and Loh (2016) and Anneart et al. (2022) look at the effect of IVOL on adjusted returns. This could possibly lead to the fact that the comparison regarding the fractions that can be explained by the candidate explanations is not fully justified and this study may observe other results when adjusted returns were taken into account.

For future research it would be interesting to use the decomposition method of Hou and Loh (2016) to decompose the IVOL puzzle for more emerging stock markets, as this study only looks at one emerging stock market, making it difficult to generalize the results to the rest of the emerging stock markets. In addition, it may be of interest for future research to investigate why and what causes the difference in explanatory power of multiple candidate variables for the IVOL puzzle per time period and stock market. If it can be found what the drivers of this differences are, this could possibly help to solve the full IVOL puzzle. Finally, it would be interesting to perform the same comparison analysis between China and the U.S. when more reliable data becomes available for Chinese stocks. For example, the candidate variable Spread could explain most of the IVOL puzzle in the U.S. sample in this study, but this variable was not available for Chinese stocks. It is possible that the comparison of the explanatory power of the candidate variables between the two markets can be more accurately compared if all candidate variables are available for both samples and for a longer period of time.

References

Amihud, Y. (2002). Illiquidity and stock returns: cross-section and time-series effects. *Journal of Financial Markets*, *5*(1), 31–56. https://doi.org/10.1016/s1386-4181(01)00024-6

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2006). The Cross-Section of Volatility and Expected Returns. *Journal of Finance*, *61*(1), 259–299. https://doi.org/10.1111/j.1540-6261.2006.00836.x

Ang, A., Hodrick, R. J., Xing, Y., & Zhang, X. (2009). High idiosyncratic volatility and low returns: International and further U.S. evidence. *Journal of Financial Economics*, *91*(1), 1–23. https://doi.org/10.1016/j.jfineco.2007.12.005

Annaert, J., De Ceuster, M. J., & Van Doninck, F. (2022). Decomposing the idiosyncratic volatility anomaly among euro area stocks. *Finance Research Letters*, *47*, 102672. https://doi.org/10.1016/j.frl.2022.102672

Avramov, D., Chordia, T., Jostova, G., & Philipov, A. (2013). Anomalies and financial distress. *Journal of Financial Economics*, *108*(1), 139–159. https://doi.org/10.1016/j.jfineco.2012.10.005

Bali, T. G., Cakici, N., & Whitelaw, R. F. (2011). Maxing out: Stocks as lotteries and the cross-section of expected returns. *Journal of Financial Economics*, *99*(2), 427–446. https://doi.org/10.1016/j.jfineco.2010.08.014

Barberis, N., & Huang, M. (2001). Mental Accounting, Loss Aversion, and Individual Stock Returns. *Journal of Finance*, *56*(4), 1247–1292. https://doi.org/10.1111/0022-1082.00367

Barberis, N., & Huang, M. (2008). Stocks as Lotteries: The Implications of Probability Weighting for Security Prices. *The American Economic Review*, *98*(5), 2066–2100. https://doi.org/10.1257/aer.98.5.2066

Bekaert, G., Harvey, C. R., & Lundblad, C. T. (2007). Liquidity and Expected Returns: Lessons from Emerging Markets. *Review of Financial Studies*, *20*(6), 1783–1831. https://doi.org/10.1093/rfs/hhm030

Black, F. (1972). Capital Market Equilibrium with Restricted Borrowing. *The Journal of Business*, *45*(3), 444. https://doi.org/10.1086/295472

Blitz, D., & Van Vliet, P. (2007). The Volatility Effect. *The Journal of Portfolio Management*, *34*(1), 102–113. https://doi.org/10.3905/jpm.2007.698039

Boyer, B. H., Mitton, T., & Vorkink, K. (2007). Idiosyncratic Volatility and Skewness: Time-Series Relations and the Cross-Section of Expected Returns. *Social Science Research Network*. https://doi.org/10.2139/ssrn.972668

Boyer, B. H., Mitton, T., & Vorkink, K. (2010). Expected Idiosyncratic Skewness. *Review of Financial Studies*, *23*(1), 169–202. https://doi.org/10.1093/rfs/hhp041

Brandt, M. W., Brav, A., Graham, J. R., & Kumar, A. (2010). The Idiosyncratic Volatility Puzzle: Time Trend or Speculative Episodes? *Review of Financial Studies*, *23*(2), 863–899. https://doi.org/10.1093/rfs/hhp087

Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance*, *52*(1), 57–82. https://doi.org/10.1111/j.1540-6261.1997.tb03808.x

Chabi-Yo, F., & Yang, J. (2010). Default Risk, Idiosyncratic Coskewness and Equity Returns. *Social Science Research Network*. https://doi.org/10.2139/ssrn.1572661

Clogg, C. C., Petkova, E., & Haritou, A. (1995). Statistical Methods for Comparing Regression Coefficients Between Models. *American Journal of Sociology*, *100*(5), 1261–1293. https://doi.org/10.1086/230638

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics*, *33*(1), 3–56. https://doi.org/10.1016/0304-405x(93)90023-5

Fama, E. F., & French, K. R. (2015). A five-factor asset pricing model. *Journal of Financial Economics*, *116*(1), 1–22. https://doi.org/10.1016/j.jfineco.2014.10.010

Fama, E. F., & French, K. R. (2018). Choosing factors. *Journal of Financial Economics*, *128*(2), 234–252. https://doi.org/10.1016/j.jfineco.2018.02.012

Fama, E. F., & MacBeth, J. D. (1973). Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy*, *81*(3), 607–636. https://doi.org/10.1086/260061

Frankfurter, G. M., & McGoun, E. G. (2001). Anomalies in finance. *International Review of Financial Analysis*, *10*(4), 407–429. https://doi.org/10.1016/s1057-5219(01)00061-8

Fu, F. (2009). Idiosyncratic risk and the cross-section of expected stock returns. *Journal of Financial Economics*, *91*(1), 24–37. https://doi.org/10.1016/j.jfineco.2008.02.003

George, T. F., & Hwang, C. (2013). Analyst Coverage and the Cross Sectional Relation Between Returns and Volatility. *Social Science Research Network*. https://doi.org/10.2139/ssrn.1787152

Goyal, A. K., & Santa-Clara, P. (2003). Idiosyncratic Risk Matters! *Journal of Finance*, *58*(3), 975–1007. https://doi.org/10.1111/1540-6261.00555

Han, Y., Hu, O., & Huang, Z. (2023). A tale of idiosyncratic volatility and illiquidity shocks: Their correlation and effects on stock returns. *International Review of Financial Analysis*, *86*, 102517. https://doi.org/10.1016/j.irfa.2023.102517

Han, Y., & Lesmond, D. A. (2011). Liquidity Biases and the Pricing of Cross-Sectional Idiosyncratic Volatility. *The Review of Financial Studies*, *24*(5). https://doi.org/10.2139/ssrn.1571758

Hou, K., & Loh, R. (2016). Have we solved the idiosyncratic volatility puzzle? *Journal of Financial Economics*, *121*(1), 167–194. https://doi.org/10.1016/j.jfineco.2016.02.013

Johnson, T. P. (2004). Forecast Dispersion and the Cross Section of Expected Returns. *Journal of Finance*, *59*(5), 1957–1978. https://doi.org/10.1111/j.1540-6261.2004.00688.x

Kohers, G., Kohers, N., & Kohers, T. (2006). The risk and return characteristics of developed and emerging stock markets: the recent evidence. *Applied Economics Letters*, *13*(11), 737–743. https://doi.org/10.1080/13504850500407210

Kumar, A. (2009). Who Gambles in the Stock Market? *Journal of Finance*, *64*(4), 1889–1933. https://doi.org/10.1111/j.1540-6261.2009.01483.x

Leippold, M., Wang, Q., & Zhou, W. (2022). Machine learning in the Chinese stock market. *Journal of Financial Economics*, *145*(2), 64–82. https://doi.org/10.1016/j.jfineco.2021.08.017

Lesmond, D. A. (2005). Liquidity of emerging markets. *Journal of Financial Economics*, *77*(2), 411–452. https://doi.org/10.1016/j.jfineco.2004.01.005

Li, W., & Wang, S. Q. (2010). Daily institutional trades and stock price volatility in a retail investor dominated emerging market. *Journal of Financial Markets*, *13*(4), 448–474. https://doi.org/10.1016/j.finmar.2010.07.003

Lintner, J. (1965). Security Prices, Risk, And Maximal Gains From Diversification. *Journal of Finance*, *20*(4), 587–615. https://doi.org/10.1111/j.1540-6261.1965.tb02930.x

Liu, H., Chen, Y., Wan, W., & Zhang, Q. (2021). A novel explanation for idiosyncratic volatility anomaly: An asset decomposition perspective. *Economics Letters*, *206*, 109994. https://doi.org/10.1016/j.econlet.2021.109994

Liu, H., Li, Q., & Zeng, Y. (2018). Firm Valuation Over Lifecycle: A Perspective on Growth Option. *Asia-Pacific Journal of Financial Studies*, *47*(5), 720–743. https://doi.org/10.1111/ajfs.12236

Merton, R. C. (1987). A Simple Model of Capital Market Equilibrium with Incomplete Information. *Journal of Finance*, *42*(3), 483–510. https://doi.org/10.1111/j.1540-6261.1987.tb04565.x

Sharpe, W. F. (1964). Capital asset prices: a theory of market equilibrium under condition of risk. *Journal of Finance*, *19*(3), 425–442. https://doi.org/10.1111/j.1540-6261.1964.tb02865.x

Stambaugh, R. F., Yu, J., & Yuan, Y. (2015). Arbitrage Asymmetry and the Idiosyncratic Volatility Puzzle. *Journal of Finance*, *70*(5), 1903–1948. https://doi.org/10.1111/jofi.12286

Wong, P. (2013). Earnings Shocks and the Idiosyncratic Volatility Anomaly in the CrossSection of Stock Returns. *Working Paper. University of South Carolina*.

Xu, Y., & Malkiel, B. G. (2004). Idiosyncratic Risk and Security Returns. *Social Science Research Network*. https://doi.org/10.2139/ssrn.255303

Zhong, A. (2018). Idiosyncratic volatility in the Australian equity market. *Pacific-basin Finance Journal*, *50*, 105–125. https://doi.org/10.1016/j.pacfin.2017.06.010