

Ambiguity Aversion and the Stock Market Participation: Empirical Evidence

Submitted by Ruogu Zhang to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Finance
In April 2015

This thesis is available for Library use on the understanding that it is copyright material
and that no quotation from the thesis may be published without proper
acknowledgement.

I certify that all material in this thesis which is not my own work has been identified and
that no material has previously been submitted and approved for the award of a degree
by this or any other University.

Signature: **Ruogu Zhang**

Acknowledgements

First and foremost, I express my deepest sense of thanks to my supervisors Professor Richard Harris and Dr. Constantinos Antoniou. Undertaking this PhD has been a truly life-changing experience for me, and it would not have been possible to do without their dedication, inspiration and encouragement.

Furthermore, I also owe a deep gratitude to Professor Ian Garrett and Dr. Rajesh Tharyan, for examining my thesis. Their scholarly advice enhanced this thesis and helped me to bring this project to a new level.

I would also like to thank my family and friends. My parents, Yanxia Wen and Diefeng Zhang, have supported me through the whole PhD journey. Their selfless dedication granted me the opportunity to progress this research, and I am forever in debt for it. In addition, my friends Polina, Parisa, Fiona, Padraic, Mara, Olga, Wenjing, Haneen, Lu Zou and Natalia are like my families here. Their encouragements helped me through some of the hardest time, and kept me going by giving me the joy of life.

Finally, I am extremely thankful to Professor Kamstra, who kindly shared his data. His generosity is truly essential to the achievements of this study.

Abstract

Theoretical models predict that ambiguity is an asset pricing factor in addition to risk, however few of them have been tested in the real market. This thesis tests one of the hypotheses that, investors' propensity to invest in stocks is reduced when ambiguity in the marketplace increases. The hypothesis is tested by using equity fund flows and households' equity holding as measurements of the market participation, and using dispersion in analysts' forecasts about aggregate returns as measurement of ambiguity. The results confirm this hypothesis, since the increases in ambiguity are significantly and negatively related to equity fund flows, as well as the likelihood that the average household invests in equities. Moreover, the results also find that the fund flows in non-dividend paying stocks are more sensitive to the changes in ambiguity, and investors transfer capital from the equity market into more liquid asset classes during high-ambiguity periods.

In addition, this thesis also tests whether there is heterogeneity in individuals' ambiguity aversion, and examines the psychological roots of ambiguity aversion. FNE theory explains ambiguity aversion as the result of fearing negative evaluation from others. It predicts that married households are more ambiguity averse; while households with higher income and education, or households that are more mature, are less ambiguity averse. On the other hand, self-evaluation theory explains ambiguity aversion as the result of minimizing anticipated regret. It predicts that households that are more optimistic, or have less income, are less ambiguity averse; while households that have negative market experience, or have higher income, are more ambiguity averse. The results show that married households, or households with high income / negative market experience, are more ambiguity averse; and households that are more optimistic / more mature, are less ambiguity averse. Therefore, both theories have successful predictions, suggesting that the ambiguity aversion is the combined result of the two motivations.

Keywords: Stock market participation; ambiguity aversion; fund flows; household portfolio

Chapter1. Introduction	1
1.1 Introduction	1
1.2 Background	1
1.3 Statement of research objectives	4
1.4 Purpose of study	5
1.5 Methodology of the study	7
1.6 Conclusion.....	10
Chapter2. Literature review.....	12
2.1 Introduction	12
2.2 Classic risk-return trade off and its problems	14
2.2.1 Risk models	14
2.2.2 Problems with risk models	16
2.2.3 Ambiguity: the missing parameter	18
2.3 Ambiguity: concept and theoretical works.....	21
2.3.1 Origins from Knight	21
2.3.2 Risk and ambiguity: SEU and Non-SEU	23
2.3.3 Modeling ambiguity	25
2.3.4 Measuring Ambiguity	26
2.4 Ambiguity and financial market.....	28
2.4.1 Source of ambiguity in financial market	28
2.4.2 Behaviour under ambiguity	30
2.4.3 Ambiguity Premium.....	34
2.5 Conclusion.....	45
Chapter3. Ambiguity and Stock Market Participation	48
3.1 Introduction	48
3.2 Literature review	50
3.2.1 Non-participation puzzle and possible explanations	50
3.2.2 Ambiguity and non-participation: reviews, models and hypothesis	52
3.2.3 Measuring ambiguity and market participation empirically	56
3.3 Methodology	59
3.3.1 Measuring ambiguity with dispersion	59
3.3.2 Measure of mutual fund flow and mutual fund exchange.....	61
3.3.3 Control variables: conditional volatility.....	63
3.3.4 Other control variables for mutual fund flows	64
3.3.5 Macro economic variables.....	65
3.3.6 Regressions.....	66
3.4 Results analysis	68
3.4.1 Ambiguity and Equity Fund Flows	68
3.4.2 Ambiguity and Different Equity Styles.....	69
3.4.3 Ambiguity and non-Equity Fund Flows.....	71
3.4.4 Ambiguity and macro variables	72

3.4.5 Robustness.....	72
3.5 Limitations and future developments	76
3.6 Conclusion.....	77
Chapter4. Heterogeneity in Household Ambiguity Aversion	79
4.1 Introduction	79
4.2 Literature review	82
4.2.1 Heterogeneity in ambiguity aversion and market participation.....	82
4.2.2 Psychological roots of ambiguity aversion	85
4.2.3 Empirical factors for ambiguity aversion	90
4.2.4 Factors for individual household participation.....	95
4.2.5 Research questions	97
4.3 Data and methodology	97
4.3.1 Survey of consumer finance	97
4.3.2 Household's participation variables from SCF	98
4.3.3 Ambiguity aversion variables.....	99
4.4 Regression results and analysis	101
4.4.1 Testing ambiguity and individual stock market participation	101
4.4.2 Testing heterogeneity in ambiguity aversion	102
4.4.3 Robustness.....	109
4.5 Limitations and future developments	111
4.6 Conclusion.....	112
Chapter5. Conclusion.....	115
5.1 Introduction	115
5.2 Empirical findings.....	115
5.3 Implications.....	117
5.3.1 Theoretical implications	117
5.3.2 Policy and practical implications	119
5.4 Limitations of study	120
5.5 Suggestions for further research.....	121
5.6 Conclusion.....	122
References	123
Figure 3.1: Net flows and net exchanges for the equity asset class.	133
Panel A. Net flows	133
Panel B. Net exchanges.....	133
Figure 3.2: Ambiguity	134
Panel A. Quarterly ambiguity.....	134
Panel B. Changes in monthly interpolated ambiguity.....	134
Figure 3.3: Ambiguity and Risk over Time.	135
Figure 3.3: Continued.....	135
Table 3.1: Classification of mutual funds	136
Table 3.2: Summary of Statistics	137

Table 3.3: Correlations	138
Table 3.4: Ambiguity and Equity Fund Flows and Exchanges	139
Table 3.5: Ambiguity and Different Style Equity Funds	140
Table 3.6: Ambiguity and non-Equity Mutual funds	141
Table 3.7: Ambiguity and macro economic variables.....	142
Table 3.8: Quarterly Regressions	143
Table 3.9: Additional Control Variables	144
Table 3.10: Different Risk Measures	145
Table 3.11: Different ν values for the ambiguity measure.....	146
Table 4.1: Ambiguity and household stock market participation.....	147
Table 4.2: Robustness for ambiguity and household stock market participation.....	148
Table 4.3: Ambiguity and household stock market participation: The effect of income	149
Table 4.4: Ambiguity and household stock market participation: The effect of age	150
Table 4.5: Ambiguity and household stock market participation: The effect of education ..	151
Table 4.6: Ambiguity and household stock market participation: The effect of marriage....	152
Table 4.7: Ambiguity and household stock market participation: The effect of social capital	153
Table 4.8: Ambiguity and household stock market participation: The effect of optimism level	154
Table 4.9: Ambiguity and household stock market participation: The effect of life-time experienced return.....	155
Table 4.10: Ambiguity and household stock market participation: The effect of smoking ..	156
Table 4.11: Ambiguity and household stock market participation: the joint test.....	157
Table 4.12: Robustness for Table 4.3-4.10	158
Appendix	159

Chapter 1. Introduction

1.1 Introduction

This thesis reports on the results of an empirical study about the impact of ambiguity on the US stock market. The purpose of this study is to investigate the influence of market ambiguity on stock market participation, and to identify the assets that are more ambiguous, and the investor groups that are more ambiguity averse. The first chapter of the thesis outlines the purpose of this study, and provides the context of the problem.

1.2 Background

Modern risk-theory has been widely applied in the financial sector and it has made an enormous impact on the global economy. According to the risk-return trade-off theory, investors can either use simple risk models such as CAPM for basic asset pricing, or choose more sophisticated ones for the derivatives. However, in 2008 the financial crisis proved that some of the risk-based pricing models can be faulty: research about the crisis shows that many of these failures come from unexpected market condition changes. For example, the pricing of Mortgage Backed Securities (MBS) is based on the assumption that the default risks of many geographically diversified mortgages are independent or, at least, are at low correlations. This assumption had been correct for a long time until the crisis brought a liquidity shock to the market, after which the correlation between mortgages defaults suddenly deteriorated. The expensive lessons from the crisis thus suggested it is important to admit that our information is imperfect and that the estimations of the parameters could be imprecise. However, since the traditional risk models require a perfect market condition and fully-rational investors' behaviour, such models are not capable of handling the situation.

Given the limited capacity of risk models, many sources suggest that ambiguity-based models might be a solution. Frank Knight (1921) was the first to point out that the pure risk models may fail to handle unpredictable changes, and he suggested that investors may think differently under ambiguity compared to risk. In his opinion, risk

refers to the events where changes are predictable, while under ambiguous events changes are not predictable. Knight's theory was backed up by Ellsberg (1961)'s study, which summarised that ambiguous events are events with low quality or conflicting information, or, more precisely, events where the probability of the outcomes is not known. Therefore, under ambiguity investors have to rely on their estimations while making their decisions. For example, to decide the fair price of the stock of a company which has conflicting analysts' forecasts, investors must generate a subjective opinion about different forecasts. Ellsberg's experiment shows that individuals are ambiguity averse, as they always make their decisions based on pessimistic results when they are not sure about precise probability, and thus they heavily undermine their predictions about ambiguous outcomes.

Such unique behaviour makes ambiguity aversion distinct from risk aversion, and there are studies which have developed detailed behavioural models for ambiguity. For example, the worst-case-scenario model assumes extreme ambiguity aversion, and it suggests that people try to maximize their utility under the worst possible case when they are engaged with ambiguity. On the other hand, the alpha-max-min model allows both ambiguity aversion and ambiguity loving behaviour, and it sees the final decision as a combination of the worst and the best result, which is weighted by the level of investors' ambiguity aversion. The worst-case-scenario model, thus, can be seen as an extreme version of the alpha-max-min model, where investors are absolutely ambiguity averse with their focus on the worst result.

There are two main-stream explanations of the psychological roots of ambiguity aversion: a Fear of Negative Evaluations by others (FNE) and a fear of self-evaluation. FNE assumes that the fear of negative evaluations from others is the main reason for ambiguity aversion: while the failure of a risky decision can be explained as pure bad luck, failure on an ambiguous decision makes people doubt their ability to make a judgement. Thus, it predicts that when people are grouped they become more ambiguity averse, as they are more concerned about others' opinions. Meanwhile, if people have a higher social status, i.e. better education or higher income, they are less affected by FNE and therefore become less ambiguity averse. The self-evaluation theory, on the other hand, suggests that people dislike ambiguity because they are afraid they will regret the decision they have made. This theory predicts that those who are more optimistic, or have lower income, are less ambiguity averse.

Due to its nature, ambiguity can exist in many areas of the market: investors can be ambiguous about the quality of information or confused about the factors pertaining to which model to use, or not sure about which model they can rely on. Thus, there are many possible ways to measure ambiguity. For example, Zhang (2006) focuses on the quality of information and he argues that some fundamentals, such as young firm age, no dividend payments and poor analyst coverage, can be indicators of ambiguity since they suggest low information quality. Zhang (2006) found that market returns are positively related to these measures thus implying ambiguity premium. On the other hand, Anderson, Ghysels and Juergens (2009) used a level of conflict in the available information to measure ambiguity. They built a beta-weighted average of analysts' forecasts on future market returns, and found that market return is positively related to this measure. In addition, they showed that under a given risk being controlled, portfolios with a higher ambiguity can outperform those with a lower ambiguity by up to 3% per year, which is a significant premium for the ambiguity.

Despite the findings of the ambiguity premium, previous researchers have reached little agreement on the mechanism of how these premiums are generated in the market. Many different reasons have been suggested for the existence of the ambiguity premium. Firstly, investors can be ambiguous about their expected market returns (Anderson et al., 2009; Zhang, 2006). Secondly, investors can be ambiguous about market volatility (Epstein and Schneider, 2006). Thirdly, ambiguity aversion can result non-trade and thus bring a liquidity pressure (Routledge and Zin, 2009; Ozsoylev and Werner, 2009). Fourthly, ambiguity can be related to investors' sentiments and can therefore affect their trading (Antoniou, Doukas and Subrahmanyam, 2014; Baker and Wurgler, 2006). Finally, ambiguity aversion may cause limited market participation (Easley and O'Hara, 2009), which would result in a price premium.

1.3 Statement of research objectives

The stock market has a long-lasting puzzle of the participation rate: in the last century the global stock market had on average a 3.5% annual premium over government bonds (Dimson, Marsh, and Staunton, 2006). Based on this factorable return, the expected-utility model predicts a very high willingness for risk-averse investors to participate in the stock market. However, stock market participation was very low compared to the predicted level: during the period 1982-1995 the US Consumer Expenditure Survey found that two thirds of all households did not invest in stocks. The study of Campbell (2006) shows that even at the eightieth percentile of wealth, almost 20% of households have no public equity. Williamson (1994) and Allen and Gale (1994) suggested that liquidity needs and transaction costs can reduce stock market participation. Yaron and Zhang (2000) suggested that the high fixed cost of entering the stock market also limits participation. Hsu (2012) argues that a household with low human capital has less diversified demand for stocks. Haliassos and Bertatut (1996) state that borrowing constraints and minimum investment requirements can also reduce market participation. However, none of these assumptions are strong enough to explain the non-participation puzzle completely.

On the other hand, many studies suggest that non-participation is driven by ambiguity aversion (Dow et al., 1992; Mukerji and Tallon, 2001; Easley and O'Hara, 2009; Epstein and Schneider 2010, Werner, 2001; Takashi, 2011). However, few of them have been empirically tested and, hence, there is a gap in the knowledge about how ambiguity affects the market participation and what the scale of the impact is. In addition, the ambiguity-based non-participation models assume that some stocks are more sensitive to ambiguity; however, the characters of these stocks are not described. Zhang (2006) found that some of the fundamental variables, such as firm age or analyst coverage, represent a high ambiguity level and indicate a higher return. Therefore, it is worth investigating empirically as to whether some of the fundamental characters can indicate ambiguity and, thus, affect the stock participation rate.

Another gap in the knowledge is the investor's ambiguity aversion level. While many models assume that investors have the same ambiguity aversion level, Easley and O'Hara (2009)'s model assumes heterogeneity in individual ambiguity aversion. Their model shows that such difference in the level of ambiguity aversion can affect willingness to participate in the market: the most ambiguity averse ones are more likely

to become sellers in the market, while the least ambiguity averse ones are buyers. Moreover, the divergence in their level of ambiguity aversion can determine whether the market will be in a 'no-trade' equilibrium. However, this model has not yet received empirical support, therefore, it is important to find evidence to prove its assumptions.

The final gap in the knowledge is an empirical study of the psychological reasons for the ambiguity aversion. As stated, there are two main-stream theories explaining ambiguity aversion: FNE and self-evaluation. FNE suggests that people are ambiguity averse, because they have a fear that failing in their decision making under ambiguity would render a negative evaluation from others. Therefore, people tend to make decisions which are easier to justify rather than making optimal ones. Researchers have found evidence that people are more ambiguity averse when making decisions as a group; however, there is no supporting evidence of this fact from the market. Self-evaluation theory assumes that people are afraid of future self-regret when making decisions under ambiguity, and they try to minimise such regret. This theory suggests that optimistic people are less ambiguity averse, but it has very little support from real market data. Therefore, although there are theoretical explanations for ambiguity aversion, the empirical evidence from the real market is still very important.

1.4 Purpose of study

According to the previous studies (Dow et al., 1992; Mukerji and Tallon, 2001; Easley and O'Hara, 2009; Epstein and Schneider 2010, Werner, 2011; Takashi, 2011), ambiguity aversion could be a reason for the limited stock market participation, and, thus, bring forth a price premium similar to the risk premium. Although these models claim that ambiguity should be an asset-pricing factor in addition to risk, there is a lack of empirical support to prove this. Therefore, the first purpose of this study is to use a quantitative method and empirically test the ambiguity's influence on stock market participation, and try to identify whether there is any asset class which carries a higher ambiguity premium than the others. As a result, the first research question which needs to be answered is:

1. Does ambiguity aversion affect stock market participation at all? And if it does, what is the scale of the impact? Is there any asset class which is affected by ambiguity aversion more than others?

Based on the previous studies, ambiguity can be measured empirically by using the divergence of opinions among analysts (Anderson et al., 2009). In addition, Kamstra, Kramer, Levi and Wermers (2014) report that as a direct measure of stock market participation the mutual fund flow can also reflect the changes in the investors' risk aversion level, due to withdrawals of money from mutual funds when their risk aversion level becomes higher. Consequently, it is assumed in this thesis that if ambiguity aversion can affect the stock market, there will be a similar effect. Therefore, the analysts' divergence will be used as a proxy of ambiguity and the mutual fund flow as a proxy of stock market participation in order to provide the empirical answers to the first questions.

Furthermore, if ambiguity aversion can in fact affect the stock market, then its micro mechanism needs to be understood. Models such as Easley and O'Hara (2009)'s work suggest heterogeneity in the individual ambiguity aversion level, and further reveal that such heterogeneity can affect individual market participation. There are various theories to explain such heterogeneity in ambiguity aversion by its psychological roots, such as negative evaluations by others (FNE) and fear of self-evaluation. However, these assumptions have not been empirically tested by using the financial market data either. Therefore, the second purpose of this study is to explore the following questions by using an empirical approach:

2. Does ambiguity affect the stock market participation at household level? If so, do different investors have a different level of ambiguity aversion? What are the reasons for it?

In answering these questions, the data of the Survey of Consumer Finance (SCF) is used. SCF reports the stock holding status of American households, and, therefore, it can be used to test the ambiguity's influence on individual stock market participation. In addition, according to FNE and self-evaluation theory, social status such as age, marital status and income, could indicate a different ambiguity aversion level. Based on the data from the SCF, households are divided into sub-groups according to these indicators and are tested on whether they have a different ambiguity aversion level. This will answer the research questions of the second part.

To sum up, in this thesis the research interest is in using empirical measures of ambiguity and stock market participation to test how, and to what degree, ambiguity aversion can affect stock market investors. It also empirically explores which asset class

bears a higher ambiguity premium, and what type of investors are more ambiguity averse.

1.5 Methodology of the study

This research applies a quantitative method to investigate the impact of ambiguity aversion on stock market participation. Chapter 3 uses a mutual fund flow and exchange as measurements of the market participation, and the ambiguity is measured by the divergence of analysts' opinions. The proxy for ambiguity is constructed by following Anderson et al. (2009), where it is measured by a divergence of forecasts. The forecast used is the Survey of Professional Forecasters (SPF), which has been produced since 1968. The forecasters are majorly from large financial institutions and the report is officially released by the Federal Reserve. Therefore, it reflects professional opinions and can be widely applied by the market. SPF forecasters give their forecasted results for economic variables such as the GDP, GNP and deflator from the previous quarter to the next five quarters of the US market, which are then used to produce the forecasted future stock market return.

In this study mutual fund flow is used as an empirical measure, to measure the stock market participation. The net mutual fund flow measures the net value of all cash inflow and outflow through the mutual funds, and thus provides a straight measurement of stock market participation. This data was provided by the Investment Company Institute (ICI) from 1985 to 2010. For each type of fund, the monthly data comprises sales, redemption, exchanges in, exchanges out and total net assets, which are used to calculate the mutual fund flow and exchange. Funds are divided into five categories by the asset class: equity, hybrid, corporate fixed income, government fixed income and money market. The main focus of this study is the equity fund category and, since the ambiguity measure is based on the US stock market only, the equity fund group must be modified to contain domestic stocks only. Therefore, the funds that are invested mainly outside the States - global equity, international equity, regional equity and emerging markets, are deleted. As a result, the final components of the equity fund groups are: aggressive growth, growth, sector, growth and income, income equity.

To ensure that the ambiguity measure does not just capture the risk, the model also contains conditional volatility as a measure of the risk. In order to net out the

influence of other factors that might affect mutual fund flow, the regression also includes the following control variables: lagged fund exchange/flows, saving rate, capital gain, advertising cost of the funds, past return of the funds, past return of the stock market and dummies for end-of-the-year effect. Capital gains and advertising costs data is provided by Kamstra et al. (2011); while saving rate is taken from the Federal Reserve data and the past stock market return is calculated from The Centre for Research in Security Prices (CRSP)'s market return index. Since the regression errors are serially correlated, the Ordinary Least Square (OLS) regression is not viable, and thus the regression method used is the Generalized Method of Moments (GMM) with monthly data, ranging from the year 1985 to 2010. The regression model is firstly tested on the equity-based mutual funds, and then on other asset classes. In addition, to examine whether some stocks are more affected by ambiguity aversion, the regression for the five fund styles is also run within the equity fund family. Additionally, it is also reasonable to assume that ambiguity can be predicted by some of the fundamental economical variables. Therefore, Chapter 3 also tests a regression model of the one-quarter-ahead ambiguity measure on the economical variables from Sibley, Xing and Zhang (2013).

To ensure the results are robust, three different methods are applied: firstly, the model with quarterly data is tested instead of monthly data to assure the interpolation has no impact on the conclusion. Secondly, a variable of median value of analyst forecast and a market sentiments proxy from Baker and Wurgler (2006) are added, therefore the conclusion made does not simply catch the effect of market sentiments. In addition, to ensure that the ambiguity measure is not just capturing the risk, the model is tested with the different risk measures including GARCH predicted volatility, realized volatility and VIX volatility index. Finally, since the ambiguity measure in this thesis is a weighted-average of the analysts' opinions, it is necessary to test the model's robustness against different weighting parameters. Therefore the model is tested again with different ambiguity series, which are generated by using a wide range of weighting parameters v .

Chapter 4 uses the data from the Survey of Consumer Finance (SCF) in order to further investigate the influence of ambiguity aversion on the individual household's stock market participation. This survey is conducted every three years by the Board of Governors of the Federal Reserve, and it gathers the demographic and financial information from a large number of different households. In Chapter 4, the data from 8

different surveys which concluded in 2010 is used. Since 1989, the SCF data files applied the multiple-imputations technique and, therefore, it contains five data sets (multiple-imputations is a technique commonly used to deal with missing information on the individual items in survey data, which employs multivariate statistical methods to impute the missing data, resulting in multiple complete data sets). As a result, in the regression the repeated imputation inferences (RII) technique is applied according to Rubin (1987), which combines the results obtained independently on each of the separate imputates using the multiple imputation combining rules.

The stock market participation of individual households is measured by their holdings of equity assets. If any household reports a direct stock holding or an equity-based fund holding, the dummy variable of the market participation is set to 1, otherwise it is 0. To ensure the results are consistent with Chapter 3, the ambiguity aversion is measured by using the method according to Anderson et al. (2009), which is the beta-weighted average of the analysts' forecasts. In addition, since the survey is conducted every three years, ambiguity is also measured by the average ambiguity between year $t-1$ and $t-3$. The control variables for the individual participation include household age, marital status (dummy is set to 1 if a household is reported as married, otherwise it is 0), an income (natural logarithm of income deflated into 2010 dollars), education level (dummy is set to 1 if the household has a college degree or higher, otherwise it is 0), anchoring of past prices (equals to the value of average Dow index price of the last year of the survey divided by the historical high price of Dow index till the last year of the survey), race (dummy is set to 1 if the household is white and 0 for the rest), retirement status (dummy is set to 1 if the respondent is retired, otherwise it is 0) and risk attitude (dummy is set to 1 if the household is not willing to take any financial risks; 2 if they can take average financial risks expecting to earn average returns; 3 if they can take substantial financial risks expecting to earn substantial returns; and 4 if they can take financial risks above average expecting to earn above average returns). After constructing the variables, firstly it is tested to see whether the ambiguity still reveals a significant impact on the individual level of the market participation by using a linear logistic model to examine the probability that a household owns equities conditionally on the predictors stated above. To check the robustness of this result, a similar process to Kumar (2009) is used, which drops households with a reported annual income less than 1000 dollars, and households with a head who is less (more) than 24 (75) years old.

The second task of Chapter 4 examines the psychological root of ambiguity aversion: if the fear of negative evaluation (FNE) is the main reason, then households with lower income (income < \$15000 in 1993 dollar) / higher education level (have college degree) / less social capital (not participating in volunteering activities) will be more ambiguity averse, as well as married households. On the other hand, a household with a higher income (income > \$125000 in 1993 dollar), older age (age > 65) will be less ambiguity averse, because they are more resistant to FNE. Meanwhile, according to the self-evaluation theory, households that are more optimistic (believe the US economy will improve in the next 5 years) will be less ambiguity averse, and households that experienced the negative life-time stock market return (have negative value of cumulated CRSP index return from the age of 10 of the household's head to the year when the survey was conducted) will be more pessimistic and, therefore, more ambiguity averse. Moreover, the self-evaluation theory makes a contrasting prediction on the relationships between ambiguity and income: low income households will regret it more if they miss the reward from ambiguous stocks. Therefore, they are less ambiguity averse. In addition, based on laboratory experiments previous studies suggest smoking as an indicator of ambiguity aversion (research from Sutter, Rutzler and Trautmann (2010) shows that impatient people tend to consume more alcohol and tobacco, demonstrate less saving behaviour and a lower ambiguity aversion level) and therefore, this assumption is tested in the real market condition here. All the parameters above are tested by using the same regression method, and to ensure the robustness of the result a joint test is also run that combines all the dummy variables related to the ambiguity aversion according to the assumptions.

1.6 Conclusion

The first part of the chapter presents the concept of ambiguity aversion and the potential existence of the ambiguity premium in the stock market in order to provide background knowledge for the research. The second part reviews the previous literature; it discusses the impact of ambiguity on market participation and the knowledge gap in empirical studies. The third section presents the purpose of this thesis, which is to explore the ambiguity's impact on stock market participation empirically, and highlight research questions. The last section discusses the methodology used to develop the answers to the research questions; it also provides the details about the data source, the

construction of variables, the regression method and robustness testing. The next chapter reviews the literature introducing the limitation of traditional risk models, and thus providing the motivation for the ambiguity research. Also it reviews the concept of ambiguity aversion and its potential impact on the market and gives evidence for the ambiguity premium.

Chapter2. Literature review

2.1 Introduction

For a long time the risk management concept has been an important topic in modern financial theory, and has lead to many highly developed models and theories. However, the recent financial crisis revealed that traditional risk theory may have missed some important issues, which may lead to serious results. For example, the pricing of a Mortgage Backed Security (MBS) product may rely on the assumption of independence of defaults for many individual mortgages in different locations. However, when the market collapsed and liquidity deteriorated, the correlation between these defaults suddenly increased, and hence brought about the catastrophic chain effect.

The assumption of a limited correlation between the mortgage defaults was predicted from the historical data, and was seen as being correct for a long time. It was not until the change in the market happened that investors finally realised that their models were wrong. As Jorion (2009) argued, failures in the risk models are normally due to *known unknowns*, such as model risk, liquidity risk and counterparty risk. However, during the 2008 crisis, risk models majorly failed because of *unknown unknowns*. That is, the risk management system is not able to foresee the structural and regulatory changes in the capital market. For example, a sudden restriction on short-sales can destroy risk-hedging plans based on short-selling. Moreover, to deal with counterparty risk it is not enough to know only about the counterparty; investors have to know about their counterparty's counterparties too. In Lehman's case, knowledge about the entire global financial network is required to foresee its total risk level. Therefore, the question is, can we really predict such changes to make the risk models perfect? In other words, can we truly gain the information that precisely reflects the market condition?

From Frank Knight (1921)'s point of view, this is impossible. Furthermore, Knight believed that in general people did not differentiate between the concept of risk and that of ambiguity. The 'risk' is usually described as the situation, where changes are foreseeable, or where the results may be unknown, but where the distribution of possible outcomes is informed. The 'risky' outcome could be described by applying the

law of numbers, and can be easily discounted to a current cost or hedged. In such a case, any abnormal returns will disappear in the long-term, because all the data has been discounted perfectly to the price. Since, in reality, the abnormal return did exist, and sometimes lasted for a long time, Knight believed that some of the dynamic changes could not be anticipated, insured or hedged against, and he called it ambiguity. When such ambiguity exists, people have no valid information base from which to make their decisions, and Knight suggested that they would have a special process of decision making, called the *estimation*. In addition, he argued that the process of ‘estimating’ is very different from the process of making decisions for the risky events.

Knight (1929)’s argument was supported mainly by the following studies, which showed that decision making under ambiguity is different from decision making under risk. Ellsberg (1961)’s experiment showed that individuals revealed preference to the risky event over the ambiguous event, and thus provided evidence of ambiguity aversion. Later studies further developed different models of decision making that involved ambiguity aversion, such as, for example, the worst case scenario model and the alpha-max-min model. Also there are studies which discuss the potential influence of ambiguity aversion in financial markets and in the corresponding pricing premium.

Although there is much theoretical literature about the ambiguity concept, one of the main issues is the lack of empirical support. Thus, this thesis aims to test the ambiguity’s impact on financial markets empirically in order to find evidence that would support the existence of the ambiguity premium. To back up such an empirical study, the first part of this chapter reviews the limitations of the traditional risk theories, and explains the motivation for the research in ambiguity. The second part introduces the concept of ambiguity including the definition, potential viable measurements, and behavioural models for ambiguity aversion. The third section further discusses the impact of ambiguity aversion in financial markets as well as the factors that may lead to the ambiguity premium.

2.2 Classic risk-return trade off and its problems

2.2.1 Risk models

For modern financial economic theory, the risk-return trade-off is always recognised as a first fundamental law. It has a wide application from the explanation of an individual investor's decision-making process to the management of an assets portfolio. The key reason for the risk being a pricing factor is due to the assumption of a risk aversion. A risk aversion, as Friedman and Savage (1948) defined, is that, with a given identical amount of return, investors should always choose the one with the lowest risk. As Katt (1965) observed, the person with a higher level of risk aversion will require a higher expected return or a lower price for the same level of risk. Therefore, when expressed in terms of wealth, risk aversion results in a risk premium and brings the trade-off.

Based on this trade-off, some of the most important financial theories have been developed, such as the modern portfolio theory and the CAPM model. The Markowitz portfolio theory suggests that given a certain level of expected return, investors can reduce the risk of their portfolio by diversification. So the optimal portfolio for individual investor depends on the assessment of the risk and their risk aversion level. This theory uses variance as an official proxy of the risk and this is a fundamental assumption. Also it is the first framework that introduces the risk-return trade-off in asset pricing when for a given expected risk the optimal expected return can be calculated. However, the Markowitz theory does not describe a clear relationship between the risk and the expected return for an individual asset, since the optimal solution depends on the available securities.

Based on the Markowitz theory, the CAPM (Sharpe, 1964) model further states that after eliminating the diversifiable risk, the return of any individual asset should be in line with its systematic risk, i.e.

$$E(R_i) = R_f + \beta(E(R_m) - R_f) \quad (2.1)$$

where R_f is the risk-free rate, and β stands for the sensitivity of the excess expected asset return to the excess expected market return. It describes the relationship between the asset risk and the expected returns for both portfolio and individual assets, and is perceived as one of the most fundamental risk-based pricing models. This model presents any risky asset's price as the result of three determining factors: level of

systematic risk, price of systematic risk, and a risk-free rate. The CAPM is very simple and the model is straightforward, because investors can estimate the price of a single asset only by using its correlation with the market risk, rather than its individual risk. Moreover, Merton (1973) introduced the ICAPM model, where the conditional expected market excesses return is positively related to the conditional variance of the market price:

$$E_t(R_{t+1}) = \alpha + \beta Var_t(R_{t+1}) \quad (2.2)$$

α is the abnormal return and should be equal to zero in a complete market, and β is the coefficient of the relative market level of the risk aversion.

Within both CAPM and ICAPM, beta is a key pricing factor for the individual assets, but in reality this is very difficult to measure. For example, for new listed stocks, due to a lack of historical data it is impossible to estimate a reliable beta. In addition, since beta can only be calculated from the historical data, it cannot reflect changes in the market conditions. Fama and French (1993) believe that beta is not the only measure of the risk and they improved the CAPM into a three factor model. They argue a size and book to market value ratio are also important factors for measuring risk. Therefore, their new model is:

$$E(R_i) = R_f + \beta(E(R_m) - R_f) + b_s SMB + b_b HML + \varepsilon \quad (1.3)$$

Besides the traditional part, the difference between the return of small and big companies (SMB) and high book to market to low book to market (HML) are also included. Since the coefficient of b_s and b_b can either be positive or negative, the model is much more flexible than normal CAPM.

Although the different types of the CAPM models use different measures and proxies of the risk, they all share some common assumptions:

1. The market is perfect, so there is no transaction cost and all information is available to all investors at the same time.
2. Investors are rational and risk-averse, thus they are looking to maximise utility.
3. Investors are only price takers; their transactions do not affect the market price.
4. Investors can lend and borrow unlimited amounts under the risk-free rate of interest.
5. Investors have homogeneous expectations.

However, in the real market these assumptions can hardly be met. Firstly, perfect market information is not available. Secondly, the investors are not the only price takers, and large demands can effectively change the market price. Thirdly, the trading costs and taxes are not zero, and unlimited borrowing facilities are not available for investors. Finally, the investors may not be perfectly rational: their expectations can be different, and they can sometimes make mistakes. Therefore, in reality, the application of risk models have many problems, which are discussed in detail in the next section.

2.2.2 Problems with risk models

As stated above, due to the strict assumptions in the model, the risk-return trade-off described by CAPM is difficult to find in practice. Although Ghysels (2005) found a significant positive relation when mixed data sampling is used; other studies, such as French, Schwert, and Stambaugh (1987) and Baillie and DeGennaro (1990), observed a positive relationship between the conditional variance and the conditional expected return but it is not significant. However, Campbell (1987) and Nelson (1991) found a significant but negative relationship. Moreover, Glosten, Jagannathan, Runkle (1993), and Harvey (2001), and Nelson (1991) observed both positive and negative relationship depending on how the average return and variance are measured.

The main explanation for such a problem of risk-return trade-off is that it requires a perfect market. In a perfect market, people are rational and information is always completely reflected by price. However many studies show that, in reality, the market is imperfect. For example, Shefrin (2002) argued that investors have a biased reaction to earning estimation, which could result in a market momentum. Hvidkjaer (2006) believe that the traders in the market can be categorised into two different groups, i.e. rational traders and noise traders, and with the existence of noise traders the market is always imperfect. Some researchers still argue that even though there are such irrational investors, the arbitragers will eventually eliminate their influence on the price. However, due to the fact that arbitrage is not costless and risk-free, it is much less efficient than expected. Jones and Lamont (2002) suggest that the cost of short stocks can be very expensive and thus the stock can be overpriced due to a lack of short-sale. DeLong, Shleifer, Summers, and Waldmann (1991) further state four different constraints for arbitragers and they prove that noise traders can add more risk to

arbitrage and can limit the behaviour of arbitragers. Therefore, the real market price, in fact, is always influenced by certain unpredictable factors.

Thus the traditional asset pricing models resulted an important dilemma for investors: to achieve a risk-return trade-off, investors need precise knowledge about the distribution of outcome, but due to the unpredictable factors in the market this is often impossible. Firstly, investors' information can be ambiguous. For example, for small companies and new industries there is always a lack of accurate fundamental information and thus the future income stream does not have a valid base to be calculated. In such a case, since the mean and variance of return cannot be calculated precisely, any model based on those parameters does not work. The results of Olsen and Troughton (2006) prove this. The results of their survey show that investors think the quantitative part is of little help in small company valuation, and believe personal judgement is more important when evaluating small firms. Moreover, even for general stocks there is much ambiguous information. One of the most obvious examples is the analysts' forecasts. Analysts barely reach agreement about the same stock, and, in a short time, even the reports about the same company from the same people may vary significantly. Therefore, if any investment decision is made using such forecasts, the outcome will be uncertain. In fact, Stickel (1991), Chan, Jegadeesh, and Lakonishok (1996), and Gleason and Lee (2003) document that stock prices start drifting right after the analyst forecast revisions.

Secondly, the procedures involved in building models can also be ambiguous. For example, it is well known that stock price volatility is related to political events. However, to describe the political situation quantitatively is very difficult. Therefore, it is controversial which independent variables should be included into the regression when one would like to introduce a political influence into the stock return models. Another example is the Fama-French (1993) three factors model. As they admit, in reality beta may not be a very good measurement of risk, and the variables in the traditional model need improvement. However, it is still controversial which measurement of risk would be the best to use.

Thirdly, investors' behaviour can be ambiguous due to the fact that they are not perfectly rational. There is a lot of evidence of biased decision making situations and irrational behaviours:

1. Investors cannot interpret past information fully and correctly, thus they can have heterogeneous beliefs. Hirshleifer and Teoh (2003) state that individual investors may have limited attention to detail and so they may omit certain important aspects of financial statements such as stock options that are not disclosed directly.

2. Investors cannot update the new information perfectly to receive different signals, and thus have un-unique prior beliefs. Hong, Scheinkeman and Xiong (2008) show that, if the market contains both informed and naive investors, the divergence in the advisors' opinions will result in an upward bias of price. This is because, in the beginning stage of the bubble, the advisors had the motivation to give the upward biased price estimation. When a new industry arises, the advisors with better knowledge usually have the estimated figures higher than the 'old-fags'. However, at the beginning it is very difficult for the investors to distinguish these experts from the others. Thus, the 'experts' would like to raise all their expectations to help investors identify their type. An informed investor will recognise this signal and will take a discount, but a naive investor will take them at face value and therefore causes the bubble in stock prices.

2.2.3 Ambiguity: the missing parameter

As stated above, in the real market the risk-return trade-off may not always hold since the situation is ambiguous rather than simply risky: investors cannot have precise information to estimate the expected return. However, the traditional asset pricing models do not consider such a situation, as they assume that investors will make the same decision for both risky and ambiguous situations.

Savage (1954) argues that even though the situation is ambiguous and investors are not sure about the distribution, they can always assign a subjective prior belief and stick to it. In such cases, he believes that the decision making process becomes exactly the same as with a risky situation. For example, assuming there is a black box containing an unknown number of red and black balls, and investors make a decision based on the probability of drawing a red ball out of the box. Savage's theory believes that investors can assume a certain figure for this probability, and make decisions deductively and consistently based on this prior belief.

Although Savage (1954)'s model may sound theoretical, empirical studies suggest otherwise. Ellsberg (1961) is the first who confirmed a different behaviour

when people are facing ambiguity. He designed an experiment to distinguish cases with precise and imprecise information: Consider two boxes with 100 red and black balls in total. The first one containing 50 red balls and 50 black ones; the second box has no further information. When a ball is randomly picked out from box one and it is red, the situation is called red 1, if it is black then the situation is called black 1, similarly for box two. The following questions are posed to the respondents:

1. Which is more likely to happen, red 1, black 1, or they are equal?
2. Which is more likely to happen, red 2, black 2, or they are equal?
3. Which is more likely to happen, red 1, red 2, or they are equal?
4. Which is more likely to happen, black 1, black 2, or they are equal?

The results of Ellsberg's test showed that in question 1 and 2 people responded that the chance should be equal. However, in question 3 and 4 they showed preference for the second box. Obviously, their prior beliefs were violated: from question 3, the preference of box two implied that people believed there are more red balls than black balls in the box. Meanwhile, choosing box two in question 4 means people also believed that, there were more black balls in the same box. Thus it can be seen that the decision making process is different when the information is ambiguous.

The next part of Ellsberg's work has some more interesting results regarding ambiguity. As shown in Table 2.1, consider the box containing 30 balls of red and another 60 balls of black and yellow in total. Picking up one ball randomly and the individuals are informed of the following bets:

Case 1: if the outcome is red then get a reward of £1, otherwise nothing.

Case 2: if the outcome is black then get a reward of £1, otherwise nothing.

Case 3: if the outcome is red or yellow then get a reward of £1, otherwise nothing.

Case 4: if the outcome is black or yellow then get a reward of £1, otherwise nothing.

Table2. 1 Ellsberg paradox

This table reports the Cases in Ellsberg (1961)'s experiment. The experiment starts with a box of 30 balls of red and another 60 balls of black and yellow in total, but the numbers of black or yellow balls are unknown. A ball is randomly picked out of the box; before its true colour is revealed, participants are asked to choose from the following bets to earn their rewards: Case 1: if the outcome is red then you get a reward of £1, otherwise nothing. Case 2: if the outcome is black then you get a reward of £1, otherwise nothing. Case 3: if the outcome is red or yellow then you get a reward of £1, otherwise nothing. Case 4: if the outcome is black or yellow then you get a reward of £1, otherwise nothing.

	Red: $p(R)=1/3$	Black: $p(B) = ?$	Yellow: $p(Y) = ?$
Case 1	1	0	0
Case 2	0	1	0
Case 3	1	0	1
Case 4	0	1	1

The participants were firstly asked to choose between Case 1 and 2 and then choose again between 3 and 4. Intuitively, if any investor prefers Case 1 rather than 2, then he should prefer 3 to 4. However, the majority of the participants reversed their choice to avoid ambiguity. Therefore, Ellsberg (1961)'s result denies Savage (1954)'s theory that investors behave indifferently towards ambiguity and risk, and Ellsberg's result is backed up by many other studies. Bossaerts, Ghirardato and Guarnaschelli and Zame (2010) improved Ellsberg's experiment by transferring the stakes into real money gain for the subjects and again showed that ambiguity can affect the investor's decision. In particular, they found that under ambiguity, there could be a point where an ambiguity-averse investor refuses to respond to market price change. Similarly, the study of Auman (1997) shows that if investors do not have enough information, they will become 'bounded rational', and so would rather have a satisfactory choice than an optimal choice, which leads to bias in the decision-making process. Gerdjikova (2006) has a similar result from the case-based study theory. The case-based study theory assumes that under ambiguity investors could only learn from past experience by evaluating a past act's performance. The result shows that with ambiguity the investor, who follows the case-based study process, in general has a suboptimal option rather than an optimal one. This result is also supported by their experiments and empirical studies.

Therefore, given the existence of ambiguity can cause ambiguity aversion (which is distinct from risk aversion), it is also reasonable to assume that ambiguity is another pricing factor, in addition to risk. However, unlike risk, ambiguity in the stock

market can be presented in many different ways. For example, investors can be uncertain in many parameters, such as expected returns, or expected volatility, or expected excess return. Thus ambiguity can affect the stock market in various ways, and even the measurement of ambiguity is very controversial. Anderson, Ghysels and Juergens (2009) used the divergence of analysts' opinions as a proxy of ambiguity, and found that the excess market return is positively correlated to this measure. In addition, their results show that controlling for risk, stock portfolios that have a higher ambiguity yield up to 3% of annual return over low ambiguity portfolios. Zhang (2006) argues that the age of the firm and the analyst coverage can also present the level of ambiguity, and found that such measurements can predict the stock returns. Other potential measures of ambiguity are included: political stability, investor sentiment and bid-ask spread.

To sum up, due to imperfect information, the market is sometimes ambiguous rather than risky, and investors behave differently towards ambiguity compared to risk. Moreover, the empirical evidence suggests that people aim to avoid ambiguity and thus this could result in a price premium. Therefore, the next section explores the concept of ambiguity further.

2.3 Ambiguity: concept and theoretical works

As stated above, the evidence shows that ambiguity aversion differs to risk aversion since there is a different decision making process is employed. Thus it is important to understand how to define the concept of ambiguity as well as how to measure and model it. Therefore, this section provides a detailed review of ambiguity.

2.3.1 Origins from Knight

Ambiguity, or the so-called Knightian uncertainty, has its origins in the work of Frank Knight in 1921. He believed that people misunderstood the concept of risk and ambiguity. 'Risk' is usually described as the situation in which changes are foreseeable, or in which the exact results of those may be unknown but the distribution of possible outcomes is informed. In such a case, the 'risky' outcome can be calculated by using expectations due to the law of numbers, and can be easily discounted to a current cost or hedged against. Thus a market with perfect knowledge only has a pure risk, and all

economic activities always move towards a zero abnormal return, as the perfect competition will discount every future event into cost.

Since, in reality, the abnormal return does exist and sometimes lasts for a long time, Knight believed that some dynamic changes could not be anticipated, insured or hedged against, and he called this ‘ambiguity’. In fact, as was concluded in the last section, in reality the market where everybody is exposed to both risk and ambiguity is imperfect. Knight further argued that in cases where people have no valid information basis they use a special process of decision making, namely estimation. As both Knight and Ellsberg agree, ambiguity exists in cases where there are un-measurable probabilities, and Table 2.2 shows all the possible attributes of an event before it actually happens:

Table 2.2 Divisions of Events

This table reports all possible situations of the knowledge about a future event. The trails of possible results are the results that could happen in this event, e.g. toss a coin and the trails are ‘head’ and ‘tail’. The distribution of possible results gives the probability distributions for all possible results, e.g. toss a perfect coin and the distribution of ‘head’ and ‘tail’ should be 50% and 50%, however if the coin is imperfect the distribution is unknown. The final result is the actual outcome of the event, such as ‘toss a coin and get a head’.

Type of situation	The trails of possible results	The distribution of possible results	Final result
1	known	known	known
2	known	known	unknown
3	<i>known</i>	<i>unknown</i>	known
4	<i>known</i>	<i>unknown</i>	unknown
5	<i>unknown</i>	<i>known</i>	known
6	<i>unknown</i>	<i>known</i>	unknown
7	unknown	unknown	known
8	unknown	unknown	unknown

Intuitively, it is impossible for anyone to know the real distributions without being informed by the set of possible results. Therefore, the distribution of possible results can be known only if the trails of possible results are known. So the only possible cases left are: 1, 2, 7, 8. Furthermore, Case 1 is the event that is certain since everything is known.

Case 2 matches the traditional description of the risk, where the distributions are perfectly known. As defined by Knight, both Case 7 and Case 8 belong to the ambiguity events, since their distributions are unknown and thus a precise prediction is impossible.

The significant difference between the Case 7 and Case 8 is whether the trails of the possible outcomes are informed. In terms of this, Case 8 is even more ambiguous than Case 7, since the possible outcomes are unknown. From the original words of Knight, they should both be considered as ambiguity events; however, their natures are very different. In fact, Case 8 is now defined as ‘unawareness’, as people are unaware of what could happen. The research of unawareness is a separate section of financial and economic studies. Therefore, in this thesis ambiguity is defined as the situation where the outcome is known but the probability is unknown.

2.3.2 Risk and ambiguity: SEU and Non-SEU

To state the influence of ambiguity, it is firstly necessary to go through the traditional decision making process. The risk-return trade-off is based on the Savage (1954)’s subjective expected utility (SEU) theory: Considering a bet on a series of events $g = \{X_1 \text{ on } E_1, \dots, X_n \text{ on } E_n\}$, in which the events E_i is mutually exclusive, and $\{E_1, \dots, E_n\}$ is a complete collection of the event. Then there is utility function X_i and a subjective probability distribution μ of events E_i . Therefore investors’ subjective expected utility (SEU) of events g is:

$$SEU_g = \sum_{i=1}^n U(X_i) \mu(E_i) \quad (2.4)$$

The model is very important because in most of the financial theories this is how the expected return is measured. It implies that the utility $U(X_i)$ of a certain gain X_i is certain, and independent with the probability distributions. The final outcome is a weighted-average of individual events. In many theories the outcome is weighted by the statistical probability, so the expected outcome is the probability-weighted average. Savage assumes that although the actual distribution is unknown, people can always use a prior belief by applying the Bayesian process from prior observations, so behaviour under ambiguity should be identical with the risk.

However, Ellsberg (1961)’s experiments demonstrated that in reality people did behave differently when the distributions are unknown and violate the SEU. As a result,

many researchers tried to develop new utility models which could describe behaviour under ambiguity, and were named NON-SEU models. A good example of such a model is from Chen and Epstein (2002), who documented a way through the backward stochastic differential utility (BSDU) model. The model allows an agent to have multiple prior beliefs other than a single one in SEU. To define the ambiguity, let there be a state space (Ω, F) . A decision maker maps Ω into outcome set X . Then according to the multi-variable models, the utility of an act f is:

$$U(f) = \text{MIN}_{Q \in P} \int u(f) dQ \quad (2.5)$$

Where $u: X \rightarrow R^1$ is the utility index and P is the subjective set of the probability measures from space (Ω, F) . The model becomes the traditional SEU model if P is a singleton. The other assumptions are:

1. Time $t, t \in [0, T]$
2. The probability space (Ω, F, P) .
3. d -dimensional standard Brownian movement $W_t = (W_t^1, W_t^2, \dots, W_t^d)$ based on space (Ω, F, P) .
4. A Brownian filtration $\{F_t\}$, where F_t is generated by $\sigma(W_s : s < t)$

Then let the $\theta = \theta(t)$ be a R^d valued process, which satisfies:

$$E[\exp(\frac{1}{2} \int_0^T |\theta_s|^2 ds)] < \infty \quad (2.6)$$

It is called the density generator, and process z_t^θ is a P-martingale and:

$$\frac{dQ^\theta}{dP} = z_T^\theta \quad (2.7)$$

Therefore, given a set of density generator, there is a set of prior

$$P^\Theta = \{Q^\theta : \theta \in \Theta\} \quad (2.8)$$

Now an unambiguous event u can be defined:

$$u = \{B \in F_t : Q(B) = P(B) \text{ for all } \theta \text{ in } P^\Theta\} \quad (2.9)$$

And all the other events are called ambiguous. In this model many differences with SEU model can be seen. For example, the expected outcome for the ambiguous events is no longer simply the probability weighted average. There are many other NON-SEU

models, and they all reveal a different decision making process than the traditional SEU model.

2.3.3 Modeling ambiguity

After the acknowledgement of Ellsberg's experiment, many studies start to catch more features of people's behaviour under ambiguity. Olsen and Troughton (2006) have summarised some main aspects observed from previous literature:

1. Ambiguity can affect decision making and, in general, people are ambiguity averse.
2. Risk aversion and ambiguity aversion are not highly related.
3. If there are any kinds of difficulties in estimating the value or distribution of the outcome, people tend to ask for a premium for the asset.
4. Ambiguity averse people tend to put more emphasis on the negative information when they are engaged with ambiguity.

Based on these facts, there are quite a few models trying to explain how people make their decisions under ambiguity, and there are two mainstream types of such models. The first type is called **alpha max-min model**, where people combine their initial judgments with estimations, and the weight depends on the level of the ambiguity and their attitude towards the ambiguity. An example is from Kelsey, Ford and Pang (2013), who believe that the final expected utility is a weighted average of maximum utility, minimum utility and expected utility. The formula is presented below:

$$E' = \delta(\alpha \text{Max}(w) + (1 - \alpha)\text{Min}(w)) + (1 - \delta)E(w) \quad (2.10)$$

Here the Max (w) and Min (w) are the max and min utility of an event W, and E (w) is the initial expected utility. $\delta \in [0,1]$ is the level of the ambiguity, the lower the δ , the more the individual is certain about the situation and sticks to his initial judgment. $\alpha \in [0, 1]$ is the attitude towards the ambiguity; a higher α means the individual puts more weight on a good outcome and is thus more optimistic. Further is stated that α is a character of the individual, which depends on the environment, and the Bayesian Updating rules should be applied here. In addition, α , δ should be measured experimentally. This model is the one that is closest to Knight's original belief that people are 'estimating' under ambiguity. It also has the major advantage of allowing an

agent to seek the ambiguity, since such a situation can sometimes be observed in the market (Viscusi & Chesson, 1999).

Another model is the max-min utility model, or what is better known as the **worst case scenario model**. It assumes that when engaged with ambiguity, ambiguity averse people act as they would when facing the worst case scenario, and they try to maximise their utility in such a situation. An example of such a model is from Caballero and Krishnamurthy (2007). Firstly, a situation w is defined, which has a ‘real’ probability θ^w , and the agent has its own ideal about the probability, which is θ_w . Let $\theta_w^w = \theta_w - \theta^w \in K = [-K_1, K_2]$, then K captures the range of err θ_w^w and so can characterize the ambiguity level. Again by defining an action C , which can be described by matrix C_i , the utility of C under situation w is $U^w = U^w(C_i)$. Then in order to find the optimal action under the ambiguity, the following equation needs to be solved:

$$\max_{C_i} \min_{\theta_w^w \in [-K_1, K_2]} E[U^w(C_i) | \theta_w^w] \quad (2.11)$$

This ‘worst-case scenario’ approach can be seen as a modification of the previous model under extreme pessimism, where $\alpha=0$ and the highest level of the ambiguity is $\delta=1$, thus the individual totally gives up his initial opinion and becomes totally pessimistic. In such a manner it causes the agent to behave symmetrically as so that their ambiguity aversion level stays the same (although the worst case can still be different depending on the individual position). Besides, this model is fundamental for many other models of ambiguity, for example, the smooth decision model.

2.3.4 Measuring Ambiguity

So far some models reviewing the behaviour under ambiguity have been already seen, however, in order to apply them, the level of ambiguity needs to be measured first. Unfortunately, this is something that both Knight and Ellsberg missed, as they left no mathematical definition. Due to its nature of being non-quantitative, unlike risk, ambiguity is very difficult to measure directly. Later studies suggest that many different methods, which are majorly based on the idea of describing the quality of information, can be used to measure ambiguity. An example is from Epstein and Schneider (2008). Assume an agent has a signal but he is not sure about it, and, as a result, the agent does not update his belief following the normal Bayesian process, but with multi-priors. Let

θ be a parameter that an investor needs to learn, and assume the signal s is related to the parameter with a family of likeli-hood:

$$S = \theta + \varepsilon, \varepsilon \sim N(0, \sigma_s^2), \sigma_s^2 \in [a, b] \quad (2.12)$$

When $a=b$, this case becomes a normal Bayesian case with a single likeli-hood, and the quality of the information is measured by $1/\sigma_s^2$. However, under the ambiguity situation the level of the ambiguity is captured in the range of $[1/b, 1/a]$.

Following this theoretical approach, many studies try to measure the ambiguity with proxies empirically. Anderson et al. (2009) suggest measuring the ambiguity with the divergence of analysts' opinions while using volatility for the risk. They assume that an agent has some imperfect knowledge about the market's excess return, so it is distributed with the informed variance σ^2 and the unknown mean μ . Thus the market risk is measured by σ while the ambiguity is measured by the agent's belief about the variance c . More precisely, let $\hat{\mu}$ be the best estimator of μ , and the level of the ambiguity can be measured by $E(\mu - \hat{\mu})^2$. In practice, this is expressed by a weighted variance of analysts' opinions. Let x be a series of a forecast about the market return at a time t , which has f_t forecasts in total, and rank the forecasts from high to low. The weight of the i th lowest forecast is:

$$W_{it}(v) = \frac{i^{v-1}(f_t+1-i)^{v-1}}{\sum_{j=1}^{f_t} j^{v-1}(f_t+1-j)^{v-1}} \quad (2.13)$$

where v is the parameter to decide the weight of each forecast. Thus the disagreement can be measured by the beta-weighted variance of the forecast series x :

$$ambiguity_t(v) = \sum_{i=1}^{f_t} W_{it}(v) [x_{it+1|t} - \sum_{i=1}^{f_t} W_{it}(v) x_{it+1|t}]^2 \quad (2.14)$$

On the other hand, whether the ambiguity can be measured by one single variable is controversial, since even the risk cannot be measured by the beta only. Therefore, similar to the Fama-French three factors model, some researchers also suggest the multi-variable models. Zhang (2006) documented other proxies, such as firm size, firm age and analyst coverage, dispersion in analyst forecasts, return volatility, and cash flow volatility.

Although it seems that those empirical proxies are quite different, they still serve the same purpose for the theoretical path, which is to catch the information quality. Intuitively, young and small companies have a lower information quality due to the lack

of records. And, as Anderson et al. (2009) argue, agents are likely to form their beliefs based on the professional forecasts, and if such forecasts have any arguments among themselves, the agent receives ambiguous information. In fact, a sophisticated agent may admit that his model is not totally reliable. Therefore, it is also reasonable to assume that the model explanation power is another kind of proxy for the ambiguity measure.

2.4 Ambiguity and financial market

2.4.1 Source of ambiguity in financial market

As previously discussed, the stock market can sometimes be ambiguous rather than simply risky. The following section provides a detailed discussion about the possible reasons that lead to market ambiguity.

a. Information ambiguity

The first and most common way to be involved in market ambiguity is by receiving ambiguous information. As stated above, by operating under imperfect knowledge, all investors are exposed to ambiguity, and this can be recognised in many different ways. Zhang (2006) argues that the information quality in the market is ambiguous. He noticed that investors always under-react to the new information such as, for example, earnings announcements. And he further assumed that investors would under-react even more with a higher level of information ambiguity: when ambiguity is involved, investors tend to overweigh their initial information and underweigh new messages. So for the stocks under high ambiguity, a strategy of buying assets with good news and selling assets with bad news should work well. He suggests that there are some proxies for the information ambiguity, like firm size, firm age, analyst coverage, dispersion in analyst forecasts, return volatility, and cash flow volatility. The evidence he found suggests that the initial market reaction to new public information is incomplete, so bad news predict relatively lower future returns and good news predict relatively higher future returns. Moreover, the level of incompleteness of the market reaction increases with the level of the information ambiguity, which implies that high ambiguity leads to the situation where the agent is to under-react more with the new information. Therefore, higher information ambiguity produces relatively lower future returns following bad

news and relatively higher future returns following good news respectively. This is inconsistent with the assumption that ambiguity is a cross-sectional risk factor, and thus supports the view that the ambiguity is a unique factor in determining stock returns.

Epstein and Schneider (2008) also believe in information ambiguity. In particular, they divide the information into tangible information such as dividends, and intangible information such as news and forecasts. The assumption is that the tangible information is certain and intangible information is uncertain. For example, people can always have a precise number of annual dividends by reading the earnings report. However, the forecast of the next annual return could be very useful or simply pointless, as it is totally depends on the analysts but very hard to confirm. They also believe that people would apply the worst case scenario formula to such a situation. As a result, prices would respond stronger to the bad news than to the good news. Their results further state that the expected excess return of the market should decrease with future information quality, as the ambiguity adverse investors would require a compensation for the ambiguity they have to hold. In other words, the investor would ask for a high ambiguity premium with a poor quality of information when the fundamental is more volatile. To prove this fact, Epstein and Schneider (2008) tested the model, which contains ambiguity averse factors in the traditional Bayesian model for the US market after 9.11. They found that the ambiguity model works more efficiently.

b. Model ambiguity

The second important source of ambiguity has appeared from the models. There are two types of problems with the models. The first one is the flaws in modelling or variable selecting process. An example of flaws in modelling is the recent crisis caused by MBS. Those MBS securities were assumed to be safe and highly rated; because traditionally people assumed that the default of two mortgages that geographically diversified in America is independent or, at least, very lowly correlated. However, when the aggregated level of liquidity was drained, such correlation suddenly rocketed and the previous risk models became no longer reliable.

Even where the agent is able to construct a correct model, in practice it may be very difficult to find a precise proxy for the variables and so the variable selecting process is flawed. For example, beta is a widely applied measure of risk, but also known

for its poor performance. Therefore, when CAPM is applied, an agent may anticipate a biased result. Another example is, as documented by Arin et al. (2007), there being 17 proxies for the degree of political freedom in previous studies, and thus it is very difficult for investors to select the right ones. His results show that due to such model ambiguity, the relationship between the market return and political events are very weak. In addition, many models come with very strict constraints, which, in practice, can never be fulfilled; for example the assumption of rational behaviour. These disadvantages of the models may have the result of a sophisticated investor becoming uncertain about the predicted results.

The second problem with models is choosing between different models. In many cases there are multiple ways of solving the same problem, e.g. to estimate an optimal global market portfolio return, one approach is to use the ICAPM model, where the weight is simply the percentage of each country's return in the global capital market. Another way is to generate benchmark weights from a mean-variance optimisation with sample estimates of the mean and covariance matrix of asset returns as inputs. Apparently both of these approaches have disadvantages: the modelling approach only works when all the constraints of ICAPM are held. On the other hand, the data approach totally relies on the past return data and ignores everything else. Therefore, a sophisticated investor may accept neither of them unconditionally, but would take a weighted average of them. However, the actual weight depends on the individual's decision making and cannot be measured precisely. Baele, Pungulescu and Horst (2006) studied the degree of home bias and found that investors do combine those two models and the weight varies through time. Therefore, if investors apply different models, they also bring ambiguity to market.

2.4.2 Behaviour under ambiguity

a. Empirical evidence of ambiguity aversion

The previous section provided a detailed discussion about the reasons that may result in market ambiguity, however it is also very important to understand how investors would react to these ambiguous situations: if most people are ambiguity neutral, it simply has no effect on assets return. The reviews of the ambiguity models made an important prediction that people are ambiguity averse, and there is empirical evidence to support

this: Ellsberg (1961)'s experiment is the first, which revealed that people are in general ambiguity averse. Actually, Ellsberg observed strong pessimism when decisions were made under ambiguity: from the quotes someone explained,

' ... if the stakes are important... I must consider how much I might expect to lose, without being unreasonable.... So in the situation where I cannot really judge confidently among a whole range of distributions...pushing me towards actions whose expected value are relatively insensitive to the particular distribution of that range...that strikes me as a sensible, conservative rules to follow. What's wrong with it?' (Ellsberg, 1961, p.663)

This is very solid empirical evidence for ambiguity aversion. However, the result was observed in an experimental environment, therefore whether it could be held in the market is still unknown.

Dimmock, Kouwenberg, Mitchell and Peijnenburg (2013) also observed ambiguity aversion under experimental conditions. By using an internet survey which implements on more than 3,000 respondents in the American Life Panel (ALP), they found strong heterogeneity in ambiguity aversion: a large fraction of respondents (52%) is ambiguity averse, a small fraction (10%) ambiguity neutral, and the remainder (38%) ambiguity seeking. Their experiment follows the classic Ellsberg urn problem, and asks respondents to choose between lotteries with known probabilities, versus lotteries with unknown probabilities. Although all of the respondents are eligible to win real monetary incentives (a total of \$23,850 was paid to 1,590 of the respondents), and thus their research shows some convincing results about an individual's economic preferences, it is still not clear how ambiguity aversion will affect the real market.

Anderson et al. (2009) found that the disagreements between analysts are positively related to the expected excess stock return. They explained that such disagreement was priced because the agents are aware of the model ambiguity. The result reveals that ambiguity aversion is very significant and also, during the market time when ambiguity is unusually large, the market excess return also goes up. In addition, the constructed high-ambiguity portfolio outperformed the low-ambiguity portfolio by 3% annually, given the same level of risk.

Although ambiguity is a possible missing parameter in the asset pricing model, empirical support is important for further research. This is also one of the main

motivations for this thesis: to find empirical evidence for ambiguity's influence on the financial market.

b. Biased decision from ambiguity and market anomalies

Another important question regarding ambiguity is about how it affects people's decision making, as this has an influence on trading and, hence, prices. The models of ambiguity, which have been previously discussed, show that decision making under ambiguity is quite different from the risk situation, as people violate the SEU model. Therefore, there are models based on the assumption that a decision made under ambiguity is biased and does not achieve the normal optimal goal. For example, Epstein and Schneider (2008) provide some details about such decision-making processes. Firstly, before information arrives, if investors are anticipating low quality information, they avoid any consumption plans following this information. Secondly, after ambiguous information arrives, the agents respond asymmetrically, i.e. the bad news would affect the investors' decision more than good news. They interpret this applying the worst case scenario model and like to maximize their minimum utility. Similarly, Zhang (2006) documents that if the situation is ambiguous, an agent becomes pessimistic and undervalues the new signal until the final result is revealed. He states that in the market investors can overweigh their own information and under-react in relation to the new information, which leads to an under-reaction to the events such as a better earning announcement.

Biased decision making can produce many anomalies in the financial market. The first type is biased portfolio choice. As French and Poterba (1991) documented, that although the US stock market comprises less than 48% of the global equity market, US investors devoted 94% of their capital into the domestic market, and such phenomenon persists even in times when most direct obstacles to foreign investment have disappeared. They believe such a biased decision is caused by ambiguity aversion, as the investors have insufficient information from foreign companies. Similarly, Uppal and Wang (2003) stated that due to ambiguity aversion, investors tend to choose stocks that they are familiar with, i.e., a local company, rather than an optimal stock. Both of the two biases (home bias and familiarity bias) lead to under-diversification and thus increase unnecessary risk.

The second type of anomaly that ambiguity could cause is the biased asset price. Mehra and Prescott (1985) find that it is difficult to simultaneously explain large historical average equity returns and the small average risk-free rate by means. For the period from 1889-1978 they observe an annualised return of 7% from S&P 500, and an annualised risk-free rate, as represented by t-bills, around 1%. The result indicates a historical equity premium of roughly 6% per year; however, a pure risk model will not imply average annual excess returns above 0.4%, unless implausibly high levels of risk aversion are tolerated. Chen and Epstein (2002) argue that such a premium is partly caused by ambiguity aversion, as investors cannot eliminate their multi-priors by simply learning from the past. Therefore, the excess returns of stocks are not only a pure risk premium, but also a partly ambiguity premium. Meanwhile, Caskey (2009) believes that ambiguity averse investors will choose to trade on aggregate level of information, and therefore cause market momentums. (A market momentum refers to the fact that stocks which have outperformed the benchmark are more likely to keep outperforming in the future, and *vice versa* to the under-performing stocks.) Similarly, Daniel, Hirshleifer and Subrahmanyam (1998) argue that price momentum will be stronger when there is high ambiguity. The reason is that when the level of ambiguity is high, information can hardly be justified. As a result, self-attribution bias and over-confidence is more difficult to control, which keeps investors' attention on the previous winning / losing stocks.

Finally, bias from ambiguity can cause liquidity issues. Easley and O'Hara (2010) stated that if the level of ambiguity is high enough, an ambiguity averse agent may stop trading, which can cause results ranging from limited market participation to market crash. Similarly, Both Kelsey et al. (2013) and Routledge and Zin (2004) believe that the biased decision from ambiguity can make it more difficult for the market-dealer to recognize informed traders. Since the dealer makes profit from noise traders and loses money to informed traders, under ambiguity the dealer will widen his bid-ask spread for compensation and, thus, cause a liquidity crisis in the market.

2.4.3 Ambiguity Premium

As stated above, people in general avoid dealing with assets with ambiguity or respond more slowly to ambiguous information, which implies the possibility of an ambiguity premium. However, since the ambiguity has different sources, the ambiguity premium might not work in the same way as a risk premium. Olsen and Troughton (2006) suggest that the abnormal part of the risk premium, such as a risk premium puzzle and a small company effect, could come from the ambiguity. They interpret the risk as ‘a function of the possibility of incurring a loss’, while ambiguity is ‘lack of information about and lack of confidence in estimating future distributions of returns or incurring a loss’. Their results show that more than 70% of professional investors would not consider just quantitative figures to estimate the risk, but also have many other non-quantitative factors as their concern. Plus, more than 90% of the investors agree that for small or extremely volatile companies, personal judgement plays a much more important role than quantitative results. Therefore, for such ambiguity-associated securities people may expect an excess return. However, as stated above, ambiguity has various sources and may affect investment decision in many different ways, so the following part of the chapter is to discuss these possible impacts.

a. Premium from ambiguity in expected return

One main reason for the ambiguity premium is that investors have imperfect information about market returns, and this results in ambiguity aversion behaviour and a premium. One example is from Anderson et al. (2009), who documented a model with the theoretical impact on ambiguity premiums:

Firstly consider the following states, which people believe follow the process:

$$dx_t = a_t dt + \Lambda_t dB_t \quad (2.15)$$

where B_t is a vector of independent standard Brownian motions and $a_t = a(x_t)$ and $\Lambda_t = \Lambda(x_t)$ are functions of the current state. The risk free rate is $\rho_t = \rho(x_t)$ and the price for the k th asset on the market is p_{kt} :

$$dp_{kt} = d_{kt} p_{kt} dt + \xi_{kt} p_{kt} dB_t \quad (2.16)$$

where $d_{kt} = d_k(x_t)$ is scalar and $\xi_{kt} = \xi_k(x_t)$ is vector. d_t and p_t are vectors, the k th elements are d_{kt} and p_{kt} . ξ_t is a matrix whose k th row is ξ_{kt} . The first asset is interpreted as the market and the wealth of investor y_t is:

$$dy_t = (\psi_t' \lambda_t y_t + p_t y_t - c_t) dt + \psi_t' \xi_t y_t dB_t \quad (2.17)$$

Here $\lambda_t = d_t - \rho_t$ is the excess return, ψ_t is the portfolio weigh vector, whose k th element presents the weight of investment on k th asset, c_t is consumption. However, as the investors perceive the ambiguity in their model, they only treat those formulas as references models. Now an agent believes that the conditional mean of the state is $a_t - \Delta_t g_t$ instead of a_t , and the conditional expected return is $d_t - \eta_t g_t$ rather than d_t . $\Delta_t = \Delta(x_t)$, $g_t = g(x_t, y_t)$ and $\eta_t = \eta(x_t)$ have the same dimension with Λ_t, B_t, ξ_t respectively. Thus in reality the true price of assets is:

$$dx_t = (a_t - \Delta_t g_t) dt + \Lambda_t dB_t \quad (2.18)$$

$$dp_{kt} = (d_{kt} - \eta_{kt} g_t) p_{kt} dt + \xi_{kt} p_{kt} dB_t \quad (2.19)$$

$$dy_t = (\psi_t' \lambda_t - \psi_t' y_t \eta_t g_t + p_t y_t - c_t) dt + \psi_t' \xi_t y_t dB_t \quad (2.20)$$

Anderson et al. (2009) assume that agents know Δ_t and η_t , but do not know g_t . Therefore, rather than being Bayesian and using a distribution of g_t , agents solve a robustness problem with g_t , which is a function of exogenous state and wealth. Therefore in the paper they apply the worst case scenario model by penalising the deviation from the reference model with the quadratic term $(g_t' g_t)/2\phi_t$, where $\phi_t = \phi(x_t, y_t)$ is a function of exogenous wealth and states. According to the worst case scenario model, the investors maximize their utility by solving:

$$E_0 \int_0^\infty \exp(-\delta t) \left[\frac{c_t^{1-\gamma}}{1-\gamma} + \frac{1}{2\phi_t} g_t' g_t \right] dt \quad (2.21)$$

where γ is the parameter for the risk aversion and δ is a time discount rate, E_0 which stands for the expectation at time zero. As a result, there is a positive premium to the excess return under the ambiguity:

$$E_t r_{et+1} = \gamma V_t + \theta M_t \quad (2.22)$$

In the model $E_t r_{et+1}$ is the excess market return at time t , while θ measures the level of ambiguity aversion and M_t measures the level of the ambiguity in the market. M_t can be measured by the disagreement with analysts' opinions. However, since the analysts have a different influence, it is unreasonable to apply a flat weight. Anderson et al. (2009)'s empirical study suggest the a flat-weighted model does not have any predictive ability on the returns. The alternative method is to use a beta-weighted way. They also found that the ambiguity and the risk do not have any significant relationship.

b. Premium from ambiguity in market volatility

So far some brief ideas about ambiguity in the market and the ambiguity premium have been discussed above. The following facts provide a brief summary:

1. People reveal ambiguity aversion in the market
2. Ambiguity can either come from low quality information or from an imperfect model
3. Part of the equity premium should be the ambiguity premium
4. Ambiguity premium is more related to excess return rather than return
5. Ambiguity premium can vary through time

As documented, ambiguity can come from multiple sources; therefore ambiguous expected return is not the only ambiguous parameter. Epstein and Schneider (2006) argue that information can be divided into tangible and intangible information and people only feel uncertain about intangible information, such as return forecasts. Consequently, the interest arises in some other vital 'intangible' parts, such as volatility. It is well-known that volatility is essential for modern financial operations. The value-at-risk is one of the most important approaches of estimating risk exposure and the estimation of volatility plays an important role in it. Also, in the dynamic risk management process, an agent not only has to estimate historical volatility, but also has to forecast it for the future, or more precisely, prepare conditional volatility estimation. However, the stock's volatility varies over time, and Campbell et al. (2006) showed that from 1962 to 1997 the firm level of stock volatility had significant changed. They found the number of stocks needed to achieve a certain level of diversity within a portfolio was increased. Therefore, any forecasting method has a chance of failing on certain stocks. Intuitively, if an agent is ambiguity averse and feels uncertain about his

ability to forecast volatility in the future, they avoid the stock or requires a premium for it.

In practice, there are a lot of methods of estimating volatility, but before considering them, some fundamental settings of return and variance need to be confirmed. Let S be a stock in the market, and at time t its price is $S_t > 0$. Then the return at time t is:

$$r_t = \ln \left(\frac{S_t}{S_{t-1}} \right) \quad (2.23)$$

Since r_t usually is seen as a random variable, people may be more interested in its time series, its mean value of μ and its standard deviation σ . In practice μ is usually treated as zero, so σ is more important for describing the returns. However, compared to returns, volatility is unobservable: the integrated volatility σ , is latent. Therefore, the market is in general more interested in a conditional volatility, which is $var(r_{n+1}|F_n)$ and provided the information up until now (F_n), the variance of the financial return r_{n+1} over the next period. This forecast is based on the current information set, so the assumptions about the returns are crucial. For example, if investors assume the returns are independent and identically normally distributed, the volatility can be expressed as below, where μ is the mean return over the period T and r_t is return at time t :

$$\hat{\sigma} = \sqrt{\frac{1}{T-1} \sum_{t=1}^T (r_t - \mu)^2} \quad (2.24)$$

Some more advanced ways to express market volatility include applying stochastic models. For example, the Heston model, where the conditional variance is in the form of:

$$d\sigma_t = \theta(\omega - \sigma_t)dt + \xi\sqrt{\sigma_t}dB_t \quad (2.25)$$

where ω is the long term mean volatility, θ stands for the rate that the volatility reverts towards its long-term mean, and ξ stands for the volatility of the volatility process, and dB_t is Gaussian with mean zero and unit standard deviation. The key assumption for this model is that the variance at least has the following features:

1. Tends to revert to its long-term mean ω
2. Has a volatility related to its standard deviation

Another popular model is the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) model, which constructs the forecasted volatility as:

$$\sigma_t^2 = \omega + \beta(L)\sigma_{t-1}^2 + \alpha(L)\eta_t^2 + \varepsilon_t \quad (2.26)$$

Comparing it to the previous model, GARCH assumes that the volatility of the variance is related to past variance rather than the square root of variance in the Heston model. Therefore, the forecast of conditional volatility shares one common problem with: the resulting volatility measures are only valid under the assumptions of the models and it is uncertain whether any of the assumptions provide a good description of actual volatility. It is important to note that investors are interested in the true integrated volatility σ^2 . Thus, due to the imperfect information set F_n , any conditional volatility forecast cannot perfectly represent the integrated volatility. In fact, the result from the models can be expressed as: $\hat{\sigma}_t^2 = \hat{\sigma}_{modeled}^2 + \varepsilon_t$, where $\hat{\sigma}_t^2$ is the true volatility and ε_t is the error term caused by the model which varies through different stocks. Hence investors face the ambiguity about the distribution of the returns because of the model ambiguity, and the lower the explanation power of the model applied, the higher is the ambiguity.

Since such model ambiguity arises from the nature of imperfect information, it is reasonable to assume that all existing models include ambiguity. So without loss of generality, some specific models, such as GARCH, can be used as a reasonable proxy for ambiguity in volatility. As Anderson et al. (2009) proved, the ambiguity in the expected return (which is measured by the divergence of analyst's opinion) is related to the excess market return: a high-ambiguity portfolio yields 3% higher annual returns compared to a low-ambiguity portfolio, given the risk to be the same. Their assumption is that investors see the market return with some unknown mean μ and known standard deviation σ ; the degree of divergence of opinion presents the level of ambiguity μ . Thus it can also be assumed that the investors are ambiguous about the distribution σ , due to the model uncertainties.

Therefore, without loss of generality, the predicted value $\hat{\sigma}$ by GARCH, the investors see the result as the true value of σ plus an error $\varepsilon \sim N(0, \sigma_{GARCH}^2)$. σ_{GARCH}^2 will vary through different stocks, as GARCH has a different performance with stocks. The smaller the σ_{GARCH}^2 , the less ambiguous the result will be. However, since σ_{GARCH}^2 is unobservable in practice the other proxies need to be used.

One possible way of doing this is to construct the interested company X with two time series of volatility, the realised one and the GARCH forecasted one, then a linear regression is run between these two. The r-square of the regression should be a reasonable measure of the overall performance of the GARCH model on stock X . The higher the r-square, the better the predictive ability the GARCH has, and the smaller the σ_{GARCH}^2 should be. Thus investors can obtain the information about the σ_{GARCH}^2 from the r-square, and then decide the level of ambiguity on the stock X and consequently the ambiguity premium.

c. Ambiguity premium from market liquidity

Recent studies show a strong link between ambiguity and the market liquidity and the liquidity premium. It is possible that, part of the liquidity premium is the result of ambiguity aversion.

From Easley and O'Hara (2009)'s model, the most optimistic investors provide the ask price and the least optimistic investors provide the bid price. As a result, the bid-ask spread is, in fact, the ambiguity premium, and, therefore, the ambiguity level should be positively related to the bid-ask spread. Similarly, Routledge and Zin (2009) also suggest that ambiguity can increase the bid-ask spread and can reduce the probability of trade.

From Ozsoylev and Werner (2009), ambiguity aversion can cause arbitragers to choose not to trade and thus reduce market liquidity. They assume the market has three participants: risk-averse informed traders, risk-neutral arbitragers and noise traders who trade randomly. The market contains one risky asset and one risk-free asset. The ambiguity arises from the pay-off of the risky asset, and the trading volume of the noise traders. At time 0 both informed traders and arbitragers are uncertain about the pay-off of the risky asset and so they have multi-beliefs. At time 1 the informed traders receive private signals which reveal the exact distribution, while the arbitragers stay uncertain and try to extract information from a market price. Since arbitragers are ambiguous, they are assumed to follow the max-min utility model. Thus if the market price is lower than the minimum expected price over his set of priors, an arbitrageur will take a long position; and if the market price is higher than the maximum expected price over the set of priors, the arbitrageur will take a short position. However, if the price is between the

maximum and minimum price, the arbitragers will choose not to trade, as the max-min utility model indicates that any transaction would not be profitable. Thus, Ozsoylev and Werner (2009)'s results show that ambiguity about the means of the information and a random asset supply create a range of price, where arbitragers will not trade and this implies an illiquid market.

The above study further compared the impact of ambiguity over the mean and the variance of the returns: higher ambiguity over the means leads to a higher liquidity risk (the probability of the market being illiquid) and a lower trading volume, but no effect on the market depth. On the other hand, higher ambiguity over the variance has no effect on the liquidity risk, but can cause a lower trading volume and a lower market depth. Ozsoylev and Werner (2009) also found that ambiguity can increase the probability of excess volatility.

Eichberger and Spanjers (2007) model ambiguity and market liquidity by using another approach, but they obtain similar results. Eichberger and Spanjers (2007) assume that ambiguity arises from the uncertainty of individuals and aggregate liquidity demand from consumers who do not know their own liquidity needs, in the first stage. In the second stage they learn their liquidity needs and the proportion of consumers with high liquidity needs become publicly available. Consumers with a higher demand for liquidity strictly prefer to hold a bank deposit which has a yield of risk-free rate, while the rest would accept illiquid assets. The market contains many small investors with one unit of initial wealth, which they can decide to hold as a deposit or as more illiquid assets. The return on the bank deposit depends on the bank's reserve rate and the expected rate of withdrawal in the future. In the period 0 both individual and aggregate liquidity needs are unknown. In period 1 the consumers privately learn their liquidity needs and the proportion of consumers with high liquidity needs become common knowledge. In period 2 the return on the asset or on the bank deposit is paid out. The two sources of the ambiguity are the unknown individual and aggregate liquidity needs of the consumers. Even though the consumers do not know their own liquidity needs in the beginning, they have to decide on what to do with their initial wealth in the period 0. In different scenarios they can choose proportions of holding their cash or investing it in illiquid assets.

d. Ambiguity premium and investor sentiments

Antoniou et al. (2014) carried out empirical studies and he found that investors can be pessimistic about the returns of companies with low information quality. The historical data implies that the excess return of small companies is too high for any pure risk measurements. Thus one possible explanation for this fact is that the price of a small company bears some premium for the ambiguity. The paper also assumes that a pessimistic mood changes over time, and will be more significant during a downward market. By applying a weighted-average of the analysis forecast as measurement of information quality, Antoniou et al. (2014) find that the forecast for small companies does have more bias as the investors receive a lower quality of information for small companies and are more uncertain about their beliefs. They further believe that the premium depends on the mood in the market. The de-trended market P/E ratio is picked as the indicator, and obviously a high P/E means that the market is more optimistic at that time. The results show that a small company's premium changes with the market sentiment, and only exists when the sentiment is low. So small companies are very likely to be undervalued when the market is more ambiguity averse. Similarly, Baker and Wurgler (2006) conclude that when a market sentiment is low, stocks that are 'highly subjective and hard to arbitrage', such as young stocks, small companies stocks, high volatility stocks, distressed stocks and no-dividend paying stocks, will have a positive excess return afterwards. And when the sentiment is high, the excess return becomes negative.

However, there is also a possibility that such a premium is negative because some of the investors are the ambiguity seekers. Brenner and Izhakian (2011) argue that the attitude to the ambiguity depends on the probability of loss which the investors are facing: if there is a high probability of loss, they will be the ambiguity loving, and if the chance of gain is high they will be ambiguity averse. Brenner and Izhakian (2011)'s research finds ambiguity loving behaviour in the market, and so the corresponding premium for the ambiguity is negative. It is plausible that under some circumstances, investors can see ambiguity as a chance. However for a long-term period and aggregated level of market, the ambiguity premium is much more likely to be positive as there are a lot of studies which support the fact that the majority of the population shows ambiguity aversion behaviour.

e. Ambiguity premium from limited market participation

Another important link between ambiguity aversion and the price premium is through limited market participation. Some models suggest that ambiguity aversion can limit investors' participation in the market and thus decrease the demand for ambiguous assets, and therefore makes the prices lower and results in a premium. An example is discussed by Easley and O'Hara (2009). Assuming there are two types of investors in the market, one type is sophisticated with a fraction of $1 - \mu$ of total investors and the type is not sophisticated with a fraction μ . The investors are $j=1,2,\dots,J$; they all have a CARA utility function with the risk aversion parameter set constant to 1: $u_j(w) = -\exp(-w)$, where w is the investor's wealth. The market has three types of assets: risk-free asset, with a constant price of 1 and net supply of 0; two types of risky assets with independent and normally distributed returns of $v^i, i = 1, 2$ and supply \bar{x}^i . The possible mean return and variance of the risky assets are: $\{\bar{v}_1^i, \dots, \bar{v}_N^i\}$ and $\{\bar{\sigma}_1^i, \dots, \bar{\sigma}_N^i\}, i=1,2$; $\theta^i = \{\theta_1^i, \dots, \theta_N^i\}$ is the set of all possible combinations of mean and variance of asset i . The sophisticated investors know about the true pay-off $\{\hat{v}^i, \hat{\sigma}^i\}$ and have EU utility, while the naive investors follow a max-min approach of all possible distributions. The investors have a budget constraint as: $w = m + p^1x^1 + p^2x^2$ and can go either long or short on each asset. If the investor chooses portfolio $\{m, x^1, x^2\}$ then his wealth in the next period would be $\tilde{w} = m + \tilde{v}^1x^1 + \tilde{v}^2x^2$.

Thus for the sophisticated investors, the problem is to solve:

$$(\hat{v}^1 - p^1)x^1 + (\hat{v}^2 - p^2)x^2 - \frac{1}{2}\hat{\sigma}^1(x^1)^2 - \frac{1}{2}\hat{\sigma}^2(x^2)^2 \quad (2.27)$$

And their demand for each asset is:

$$x_s^{i*}(p^i) = \frac{\hat{v}^i - p^i}{\hat{\sigma}^i} \quad (2.28)$$

For the unsophisticated investors, the problem is:

$$\max_{(x^1, x^2)} \min_{(\theta^1, \theta^2)} (\bar{v}^1 - p^1)x^1 + (\bar{v}^2 - p^2)x^2 - \frac{1}{2}\bar{\sigma}^1(x^1)^2 - \frac{1}{2}\bar{\sigma}^2(x^2)^2 \quad (2.29)$$

And their demand for each asset is:

$$x_U^{i*}(p^i) = \begin{cases} \frac{\bar{v}_{\min}^i - p^i}{\sigma_{\max}^i}, & \text{if } \bar{v}_{\min}^i > p^i \\ 0, & \text{if } \bar{v}_{\min}^i \leq p^i \leq \bar{v}_{\max}^i \\ \frac{\bar{v}_{\max}^i - p^i}{\sigma_{\max}^i}, & \text{if } \bar{v}_{\max}^i < p^i \end{cases} \quad (2.30)$$

If the market price of the asset is between the min and max of possible pay-off, the unsophisticated investors will not participate, and this decision is irrelevant to the variance. Also the unsophisticated investors only consider the maximum value of volatility if they do participate, since they always take the worst case scenario.

The difference between the position of risky holdings of two different types of investors, $x_S^{i*}(p^i) - x_U^{i*}(p^i)$, is always positive, thus the sophisticated investors always have more risky assets because they do not have ambiguity averse behaviour. Since per capita supply and demand must be matched:

$$\mu x_S^{i*}(p^i) + (1 - \mu)x_U^{i*}(p^i) = \bar{x}^i \quad (2.31)$$

If the market price is between the max and min of expectation of the unsophisticated investors, there is no participation by them and the entire asset class is owned by the sophisticated investors. Thus the market clear price is:

$$\hat{p}^i = \hat{v}^i - \frac{\hat{\sigma}^i - \bar{x}^i}{1 - \mu} \quad (2.32)$$

Since $\bar{v}_{\max}^i \geq \hat{v}^i > \hat{p}^i \geq \bar{v}_{\min}^i$ is the real constraint, for such situation to happen $\hat{p}^i \geq \bar{v}_{\min}^i$ need to be satisfied.

If there is participation from both types of investors, the clear price is:

$$p^{i*} = \frac{\mu \hat{\sigma}^i \bar{v}_{\min}^i + (1 - \mu) \sigma_{\max}^i \hat{v}^i - \bar{x}^i \sigma_{\max}^i \hat{\sigma}^i}{\mu \hat{\sigma}^i + (1 - \mu) \sigma_{\max}^i} \quad (2.33)$$

The participation of the naive investors happens only when: $\hat{p}^i = \hat{v}^i - \frac{\hat{\sigma}^i - \bar{x}^i}{1 - \mu} < \bar{v}_{\min}^i$.

Since the previous result shows the constraint for non-participation case which is $\hat{p}^i \geq \bar{v}_{\min}^i$, it can be seen that only one equilibrium can exist at the same time.

It is notable that in this model the naive investors will never go short (although they are allowed), since they require the price to be greater than \bar{v}_{\max}^i for going short.

However, by this time all sophisticated investors will go short and therefore none of them will take the long position. Also, the naive investor will only participate when:

$$\hat{p}^i = \hat{v}^i - \frac{\hat{\sigma}^i - \bar{x}^i}{1-\mu} < \bar{v}_{min}^i \quad (2.34)$$

If the naive investor has a really low anticipation of \bar{v}_{min}^i , the chances for such an investor to participate in the market are low .

To test the effect of ambiguity aversion on the market return, let us assume there are two risky assets: one is the stock of a large firm and another is the stock of a small firm. For the first asset both types of investors participate and for the second asset there is enough ambiguity to stop naive investors from participating. Thus the price for the first asset would be: $p_1 = \hat{v}^1 - \bar{x}^1 \hat{\sigma}^1$ and the price for the second asset would be: $p_2 = \hat{v}^2 - \bar{x}^2 \hat{\sigma}^2 / (1 - \mu)$. Now assume there is an outsider, a classic representative CARA investor, who observes the market price and analyses the information to form a ‘market belief’. Such an investor will have the following results:

$$\sigma^{A1} = \hat{\sigma}^1, \sigma^{A2} = \frac{\hat{\sigma}^2}{1-\mu}, v^{A1} = \hat{v}^1, v^{A2} = \hat{v}^2 \quad (2.35)$$

The market representative agent then calculates the return on each asset as: $\tilde{\gamma}^i = \frac{\tilde{v}^i}{p^i} - 1$, and the market return is $\tilde{\gamma}^M = \frac{\tilde{v}^1 \bar{x}^1 + \tilde{v}^2 \bar{x}^2}{p^1 \bar{x}^1 + p^2 \bar{x}^2} - 1$. Since the assumed agent observes $v^{A1} = \hat{v}^1, v^{A2} = \hat{v}^2$, which are correct, the estimation on the market return is therefore correct. And he also sees that $\sigma^{A1} = \hat{\sigma}^1$, which is again correct. However for the second asset, due to the non-participation of the naive ambiguity averse investors, the assumed CARA investor would miscalculate the variance as $\sigma^{A2} = \frac{\hat{\sigma}^2}{1-\mu}$. Therefore, when the assumed agent tries to apply the CAPM model, it would not work perfectly because there is now an overestimation of the second asset’s true variance. Therefore, the agent’s calculation of beta would be different from the real beta, and lead to result:

$$\tilde{\gamma}^i = \alpha^i + \beta^i \tilde{\gamma}^M \quad (2.36)$$

where:

$$\alpha^1 = \frac{-\mu \bar{x}^2 \bar{r}^1 \bar{r}^2 p^2}{p^1 [(\bar{x}^1)^2 \sigma^1 + (\bar{x}^2)^2 \sigma^2]} < 0, \alpha^2 = \frac{-\mu \bar{x}^1 \bar{r}^1 \bar{r}^2 p^1}{p^2 [(\bar{x}^1)^2 \sigma^1 + (\bar{x}^2)^2 \sigma^2]} > 0 \quad (2.37)$$

Thus from the outsider's view, there is a positive excess return with the ambiguity assets, which is affected by the non-participation of ambiguity averse investors, when the traditional CAPM model is used. This is due to the fact that ambiguity averse investors overweigh their portfolios of the non-ambiguous asset, which increases its price and lowers its return. On the other hand, since the ambiguous assets have a lower demand from the ambiguity averse investors, other investors will be able to buy them at a lower price and to gain a higher return.

2.5 Conclusion

The recent crisis has proved that traditional risk theories can be faulty, due to imperfect information. For such models as CAPM to work, some strict assumptions are required such as perfect market information and rational investor behaviour. However, in the real market investors normally have imprecise information when they are building up models, or selecting model input. In addition, the behaviour of investors is not perfectly rational.

Frank Knight (1921) believed that investors do not face a simple risk: he argued that risky events are events with a known probability of distribution, and if the distribution is unknown then the situation is ambiguous. For example, if a firm has a lack of historical data, or the opinions of different analysts towards the same company conflict with each other, the market is ambiguous rather than risky. Knight further suggested that investors deal with ambiguity differently to risk, as their decision making process is different. Savage (1954) disagreed with this idea, and suggested that investors can change an ambiguous event into a risky event by sticking to a single prior assumption. However, this theory was denied by Ellsberg (1961), whose results revealed that people change their prior assumptions when they are making decisions under the ambiguity. Moreover, Ellsberg (1961) found that people try to avoid ambiguity more than risk, thus he provided the evidence for ambiguity aversion. Therefore, similar to risk, it is also reasonable to assume an ambiguity-return trade-off, which suggests that the ambiguity should be an independent asset pricing factor.

Based on Knight (1921) and Ellsberg (1961)'s results, further studies constructed different models to explain ambiguity aversion. Most of the literature agrees on the fact that when investors are engaged with ambiguity, instead of being perfectly

rational and seeking for the optimised solution, they might be dominated by their emotions and make sub-optimal decisions. The two mainstream models are the ‘worst case scenario’ model and the ‘alpha-max-min’ model. The worst case scenario model assumes that investors are pessimistic towards the ambiguity, and thus they always choose to prevent loss in the worst possible case. The alpha-max-min model, however, allows investors to be both ambiguity averse and ambiguity loving. Under this model the final decision is the combination of the best and the worst outcomes, and the weighting depends on how optimistic the investors are.

Apart from distinct behaviour, the sources of ambiguity are also different from risk. For example, Zhang (2006) suggests that ambiguity may come from a low quality of information. He argues that young firms or firms with a lack of analysts’ coverage are ambiguous. Zhang (2006) found that investors require a higher return on these stocks. On the other hand, Arin et al. (2007) believe that ambiguity may also come from the model fitting process. They state that there are 17 suggested proxies for the degree of politicians’ freedom in previous studies, and they showed that due to model ambiguity, the relationships between market return and political events are very weak. Moreover, investors can even be ambiguous about which model they have to rely on. For example Baele, Pungulescu and Horst (2006) studied the degree of home bias and found that investors did combine different models to estimate the global market return, and the weights of different models are time-varying.

Such unique features of ambiguity have resulted in difficulties in empirical studies. In particular, measuring ambiguity is always a controversial process. Zhang (2006) suggests that some fundamental variables such as firm age, analysts’ coverage and dividend payment are good indicators of ambiguity. On the other hand, Anderson et al. (2009) use the divergence of analysts’ opinions as measure of the ambiguity. They applied a beta-weighted variance of analysts’ opinions to measure stock market ambiguity and found that the excess market returns are positively correlated with this variable. In addition, Anderson et al. (2009) found that under a given risk being controlled, a high-ambiguity portfolio can outperform a low-ambiguity portfolio by 3% annually which means a significant ambiguity premium.

Further studies suggest many different reasons for the ambiguity premium to exist. Firstly, investors can be ambiguous about expected market returns. (Anderson et al., 2009; and Zhang, 2006). Secondly, investors can be ambiguous about market

volatility (Epstein and Schneider, 2006; Campbell, et al. 2006). Thirdly, ambiguity aversion can result in no trade and thus bring a liquidity pressure (Routledge and Zin, 2009; Ozsoylev and Werner, 2009). Fourthly, ambiguity can be related to investors' sentiments and it can affect trading (Antoniou et al. 2014; Baker and Wurgler, 2006). Finally, ambiguity aversion may result in limited market participation (Easley and O'Hara, 2009). Although, there is sufficient theoretical knowledge which suggests the existence of an ambiguity premium, little of this knowledge has been empirically tested. Therefore, there is a motivation in the next parts of this thesis to study the impact of the ambiguity empirically, in order to test the existing models and provide evidence for the ambiguity-based pricing model.

Chapter3. Ambiguity and Stock Market Participation

3.1 Introduction

The previous chapter reviewed literature about ambiguity and its behavioural models. As it can be seen, the theories have been well developed and many classic laboratory experiments, such as Ellsberg (1961), have been performed. However, there are a lack of empirical studies regarding ambiguity's effect in the real market. Therefore, in this chapter, the relationship between ambiguity aversion and stock market participation is explored by using the mutual fund data. According to the ambiguity-based explanation of the participation puzzle, stocks are involved with both risk and ambiguity. And since most people are ambiguity averse, their participation should be negatively related to their ambiguity level. So the main purpose of this chapter is to test these models empirically and to find the evidence for ambiguity aversion in the real market.

As a start, the reasons for the limited market participation and the ambiguity-based behaviour model are reviewed here. After formulating the research questions, it is examined whether ambiguity aversion can reduce a capital entry in the equity market. The starting point for this analysis is that non-professional investors, whose major investment approach is through mutual funds, are known to have a low participation rate in the stock market. Two empirical proxies are used for market participation: mutual fund net flow, which is the net cash flow into equity funds; and mutual fund exchange, which is the switching of capital between funds of different asset classes managed by the same investment house. The first measure captures stock market participation in absolute terms, and the second measures stock market participation that is relative to other asset classes.

Following Anderson et al. (2009), the Survey of Professional Forecasters (SPF) data is used to build the ambiguity measure, which is a weighted divergence of analysts' opinions from SPF. According to Ellsberg (1961), ambiguity can be objectively measured by the quality of information, or by the degree of conflict. Therefore, when divergence amongst analysts' opinions towards the future stock markets is high, ambiguity is also likely to be high, since the fundamental information for future

economy is conflicting. In this chapter it is tested whether such a measure can explain the changes in market participation after controlling past fund returns (Ippolito, 1992; Sirri and Tufano, 1998), capital gains (Kamstra et al., 2011), past flows (Ben-Rephael, Kandel, and Wohl, 2011), seasonal effects (Kamstra et al., 2011), advertising expenses (Gallaher, Kaniel and Starks, 2006), past market returns (Ben-Rephael, et al., 2012) and savings (Kamstra et al., 2011). In particular, conditional volatility is used as proxy of risk to make sure the results are not simply caused by risk aversion. The result shows that, with the listed controlling variables, the ambiguity level is significantly and negatively related to both fund flow and fund exchange, which supports the assumption that ambiguity aversion can reduce stock market participation.

In addition, the fund data is further divided by equity type in order to see whether the holding of a certain stock is more sensitive than others. The result shows that the effect of ambiguity is concentrated in funds which are classed as ‘aggressive growth’ and ‘growth’. These funds invest in more ambiguous, non-dividend paying firms. Furthermore, the result shows that while ambiguity has negative influence on the equity fund exchange, it increases the net exchange for money market funds. Based on the Investment Company Institute (ICI)’s definition the money market funds invest in low-risk, high-grade assets that receive full principal and interest within 90 days on average. Thus a plausible explanation for this result is that investors seek a safer and more liquid asset class when faced with higher ambiguity in the expected equity returns. For checking the robustness of this chapter’s result, quarterly regression is used instead of monthly regression and different risk measures such as realised volatility, implied volatility and GARCH predictions are tested, and the result is still held.

Previous studies built ambiguity-based theoretical explanations for the limited-participation puzzle, for example Epstein and Schneider (2010), Dimmock, Kouwenberg, Mitchell and Peijnenburg (2013). This chapter empirically supports these theories and shows that ambiguity aversion is an important factor in the stock market participation decision-making process. In addition, as discussed in Chapter 2, ambiguity models can resolve many market puzzles including the familiarity-bias (Uppal and Wang, 2003), the price momentum (Caskey, 2009) and the equity premium puzzle (Chen and Epstein, 2002). In this chapter it is shown that the divergence of analysts’ opinions from Anderson et.al (2009) could be a useful empirical measurement and may help to test these theories with financial data. Furthermore, the result also contributes to

the studies of mutual fund flows as it shows that the fund flows are negatively correlated to the uncertainty of future market returns.

3.2 Literature review

The previous chapter provides an introduction into the limitation of traditional risk models, and thus offers a motivation for research of ambiguity. The concept of ambiguity aversion and its potential impact on the market have also been reviewed. This section of the chapter reviews this topic even further and reveals a link between ambiguity aversion and market participation, and introduces background knowledge about the mutual fund flow in order to construct control variables.

3.2.1 Non-participation puzzle and possible explanations

In the last century the global stock market had a 3.5% annual premium over government bonds on average (Dimson, Marsh, and Staunton, 2006), which led to a remarkable risk-adjusted return for equity investors. Based on this high risk-return trade-off, the expected-utility model predicts a very high willingness for the risk-averse investors to participate in the stock market. However, stock market participation is very low when comparing it to the predicted level: during the period 1982-1995 the US Consumer Expenditure Survey found that two thirds of all households do not invest in stocks. The study of Campbell et al. (2006) shows that even at the eightieth percentile of wealth, almost 20% of households have no public equity. Such results are difficult to explain using the basic expected-utility model, and this phenomenon is widely known as the non-participation puzzle.

The non-participation puzzle is widely debated in the literature and there are various explanations of this puzzle. The main-stream theories focus on the cost of participation: Williamson (1994) and Allen and Gale (1994) suggest that liquidity needs and transaction costs can reduce the stock market participation. Yaron and Zhang (2000) suggest that the fixed cost of entering the stock market for the first time is too high and that limits the participation. This point of view is also supported by Vissing-Jorgenson (2002) and Guiso, Haliassos and Jappelli (2003). On the other hand, there are studies explaining the puzzle from consumers' demand and market limitations: Hsu (2012)

argues that a household with low human capital has less diversification demand for stocks, and Haliassos and Bertatut (1996) suggest that borrowing constraints and minimum investment requirements can also reduce market participation. However, none of these assumptions are strong enough to explain the non-participation puzzle completely, thus some other studies suggest more behavioural-based explanations.

One behavioural explanation is that non-participation is driven by ambiguity aversion (Dow et al., 1992; Mukerji and Tallon, 2001; Easley and O'Hara, 2009; Epstein and Schneider, 2010; Werner, 2011; Takashi, 2011). As stated in Chapter 2, ambiguity was initially developed by Knight (1921), and it describes a situation where the probabilities associated with future states are unknown. Ellsberg (1961)'s result was the first that stated that people are particularly averse to ambiguity rather than to pure risk only. This result was confirmed later by many studies in experimental economics and psychology. Various research results starting from Knight (1921) and Keynes (1921), followed by Ellsberg (1961) and, more recently, Ahn et al. (2009) show that situations which involve ambiguity are treated differently from those that involve risk. Hsu et al. (2005) and Levy, Snell, Nelson, Rustichini and Glimcher (2010) present evidence that ambiguous situations produce a unique neurological fingerprint, suggesting that ambiguity aversion has a root in the fundamentals of human cognition. According to the ambiguity-based explanation of the non-participation puzzle, stocks involve both risk and ambiguity. Since the majority of people are averse to ambiguity, their willingness to invest in stocks is lower than is implied by the EU model. The next section of this chapter reviews details about the relationship between the ambiguity aversion and the market participation puzzle.

3.2.2 Ambiguity and non-participation: reviews, models and hypothesis

The key of asset pricing is to evaluate future incomes with various possibilities. The expected utility (EU) theory suggests the utility of risky events can be presented by a weighted sum of each outcome's utility multiplied by their probabilities. In general, let us consider a bet on a series of events $g = \{X_1 \text{ on } E_1, \dots, X_n \text{ on } E_n\}$, where events E_i are mutually exclusive, and $\{E_1, \dots, E_n\}$ is a complete collection of all the events. If there is a utility function U for the gain X_i and probability distribution P for event E_i , individuals will have an expected utility of g :

$$EU_g = \sum_{i=1}^n U(X_i) P(E_i) \quad (3.1)$$

In real life, however, the probabilities are very likely to be unknown. Savage (1954) believes that the problem can be solved by generating a subjective prior distribution that leads to the subjective expected utility (SEU):

$$SEU_g = \sum_{i=1}^n U(X_i) P_{\text{subjective}}(E_i) \quad (3.2)$$

Although the assigned probabilities are subjective, SEU assumes these subjective beliefs will be treated the same as real distributions. Thus each $P(E_i)$ will not be changed unless it can be updated by a Bayesian process. SEU is deeply embedded in the asset pricing theory since it is a default way to calculate the expected returns. Under the SEU theory, decisions for risky and ambiguous events are identical, because individuals have the same subjective priors. However, as reviewed in Chapter 2, Ellsberg (1961) showed that ambiguous events lead to different behaviour as the priors are situational in the presence of the ambiguity. Moreover, his results suggest there is a unique premium for ambiguity aversion.

Easley and O'Hara (2009) have a detailed model which describes how ambiguity aversion could affect demand and pricing in the stock market. Assume there are two types of investors in the market. One is the sophisticated type with a fraction of $1 - \mu$ of total investors and the rest is the unsophisticated type of investors with a fraction μ . The investors are $j=1,2,\dots,J$; all have CARA utility function with the risk aversion parameter set constantly equal to 1: $u_j(w) = -\exp(-w)$, where w is the investor's wealth. The market has three types of assets: risk-free asset, or money, with a constant price of 1 and net supply of 0; two types of risky assets with independent and

normally distributed returns of v_i , $i = 1, 2$ and supply \bar{x}_i . The possible mean return and variances of the risky assets are: $\{\bar{v}_{i1}, \dots, \bar{v}_{iN}\}$ and $\{\bar{\sigma}_{i1}, \dots, \bar{\sigma}_{iN}\}$, $i=1,2$; $\theta_i = \{\theta_{i1}, \dots, \theta_{iN}\}$ is the set of all possible combinations of mean and variance of asset i . The sophisticated investors know about the true pay off $\{\hat{v}_i, \hat{\sigma}_i\}$ and so they have a EU utility, while the naive investors follow a max-min approach for all possible distributions. The investors have a budget constraint of: $w = m + p_1x_1 + p_2x_2$ and can go either long or short on each asset. If the investor chooses portfolio $\{m, x_1, x_2\}$ then his wealth in the next period would be $\tilde{w} = m + \tilde{v}_1x_1 + \tilde{v}_2x_2$. Thus for the sophisticated investors, the problem is to solve:

$$(\hat{v}_1 - p_1)x_1 + (\hat{v}_2 - p_2)x_2 - \frac{1}{2}\hat{\sigma}_1(x_1)^2 - \frac{1}{2}\hat{\sigma}_2(x_2)^2 \quad (3.3)$$

And their demand for each asset is:

$$x_{is}^*(p_i) = \frac{\hat{v}_i - p_i}{\hat{\sigma}_i} \quad (3.4)$$

For the unsophisticated investors, the problem is:

$$\max_{(x_1, x_2)} \min_{(\theta_1, \theta_2)} (\hat{v}_1 - p_1)x_1 + (\hat{v}_2 - p_2)x_2 - \frac{1}{2}\bar{\sigma}_1(x_1)^2 - \frac{1}{2}\bar{\sigma}_2(x_2)^2 \quad (3.5)$$

And their demand for each asset is:

$$x_U^{i*}(p_i) = \begin{cases} \frac{\bar{v}_{i\min} - p_i}{\sigma_{i\max}}, & \text{if } \bar{v}_{i\min} > p_i \\ 0, & \text{if } \bar{v}_{i\min} \leq p_i \leq \bar{v}_{i\max} \\ \frac{\bar{v}_{i\max} - p_i}{\sigma_{i\max}}, & \text{if } \bar{v}_{i\max} < p_i \end{cases} \quad (3.6)$$

Therefore, if the market price of the asset is between the min and max possible pay-off, the unsophisticated investors do not participate, and this decision is irrelevant to the variance. Also, the unsophisticated investors only consider the maximum value of volatility if they do participate, since they always take the worst case scenario.

The difference between the position of risky holdings for two different types of investors, $x_s^{i*}(p_i) - x_U^{i*}(p_i)$, is always positive, thus the sophisticated investors always hold more risky assets, because they do not have ambiguity averse behaviour. Since per capita supply and demand must match, there is:

$$\mu x_s^{i*}(p_i) + (1 - \mu)x_U^{i*}(p_i) = \bar{x}_i \quad (3.7)$$

If the market price is between the max and min of expectations of the unsophisticated investors, there is no participation from them and the entire asset class is owned by the sophisticated investors. Thus the market clear price is:

$$\hat{p}_i = \hat{v}_i - \frac{\hat{\sigma}_i - \bar{x}_i}{1 - \mu} \quad (3.8)$$

Since $\bar{v}_{i_{min}} \geq \hat{v}_i > \hat{p}_i \geq \bar{v}_{i_{min}}$, the real constraint for the situation to happen is that $\hat{p}_i \geq \bar{v}_{i_{min}}$. If there is a participation from both types of investors, the clear price would be equal to:

$$p_{i*} = \frac{\mu \hat{\sigma}_i \bar{v}_{i_{min}} + (1 - \mu) \sigma_{i_{max}} \hat{v}_i - \bar{x}_i \sigma_{i_{max}} \hat{\sigma}_i}{\mu \hat{\sigma}_i + (1 - \mu) \sigma_{i_{max}}} \quad (3.9)$$

The participation of the naive investors happens only when: $\hat{p}_i = \hat{v}_i - \frac{\hat{\sigma}_i - \bar{x}_i}{1 - \mu} < \hat{v}_{i_{min}}$. Since the previous result shows the constraint of a non-participation case is $\hat{p}_i \geq \bar{v}_{i_{min}}$, it can be seen that only one equilibrium can exist at the same time.

It is notable that in this model the unsophisticated investors never go short (although they are allowed to do so), since they require a price greater than $\bar{v}_{i_{max}}^i$ to go short. However, by this time all of the sophisticated investors will have gone short and none of them will take a long position. Since the unsophisticated investors only participate when:

$$\hat{p}_i = \hat{v}_i - \frac{\hat{\sigma}_i - \bar{x}_i}{1 - \mu} < \bar{v}_{i_{min}} \quad (3.10)$$

and, in the case that $\bar{v}_{i_{min}}$ is very low, then the chances for them to participate in the market are low as well.

To check the effect of ambiguity aversion on the market return, assume there are two risky assets of one large firm's stock and one small firm's stock. For the large firm's asset (asset 1), both types of investors participate; and for the other asset (asset 2), there is enough of the ambiguity to stop the naive investors from participating. Thus the price for the asset one would be: $p_1 = \hat{v}_1 - \bar{x}_1 \hat{\sigma}_1$ and the price for the asset two would be: $p_2 = \hat{v}_2 - \bar{x}_2 \hat{\sigma}_2 / (1 - \mu)$. Now, let us assume there is an outsider, a classic representative of a CARA investor, who observes the market price and analyses the

information to form a ‘market belief’, and he would have the following calculated results:

$$\sigma_{A1} = \hat{\sigma}_1, \sigma_{A2} = \frac{\hat{\sigma}_2}{1-\mu}, V_{A1} = \hat{v}_1, V_{A2} = \hat{v}_2 \quad (3.11)$$

This assumed market agent then calculates the return on each asset as: $\tilde{r}_i = \frac{\tilde{v}_i}{p_i} - 1$, and the market return is $\tilde{r}_M = \frac{\tilde{v}_1 \bar{x}_1 + \tilde{v}_2 \bar{x}_2}{p^1 \bar{x}_1 + p^2 \bar{x}_2} - 1$. Since the assumed agent observes $V_{A1} = \hat{v}_1, V_{A2} = \hat{v}_2$, are correct, then the estimation on the market return is correct as well. The agent also sees that $\sigma_{A1} = \hat{\sigma}_1$, which is again correct. However, for the asset two, due to the non-participation of the naive ambiguity averse investors, the assumed CARA investor has miscalculated the variance as $\sigma_{A2} = \frac{\hat{\sigma}_2}{1-\mu}$. If the assumed agent tried to apply the CAMP model, it would not work perfectly, because now there is an overestimation of the asset two’s true variance. Therefore, the agent’s calculated beta would be different from the real beta, and results:

$$\tilde{r}_i = \alpha_i + \beta_i \bar{r}_M \quad (3.12)$$

$$\alpha_1 = \frac{-\mu \bar{x}_2 \bar{r}_1 \bar{r}_2 p_2}{p_1 [(\bar{x}_1)^2 \sigma_1 + (\bar{x}_2)^2 \sigma_2]} < 0, \alpha_2 = \frac{-\mu \bar{x}_1 \bar{r}_1 \bar{r}_2 p_1}{p_2 [(\bar{x}_1)^2 \sigma_1 + (\bar{x}_2)^2 \sigma_2]} > 0 \quad (3.13)$$

Thus from an outsider’s view, there is a positive excess return with ambiguous assets, which has resulted from the non-participation of the ambiguity averse investors, and a negative excess return for the non-ambiguous assets, when a traditional CAPM model is used. This is due to ambiguity averse investors overweighing their portfolios by non-ambiguous assets, and thus increasing its price and lowering its returns. On the other hand, since ambiguous assets have a lower demand from ambiguity averse investors, others can buy them at a lower price and gain a higher return. From the stated model, one can see that ambiguity aversion could lower the demand for ambiguous stock assets, and result in a premium. Moreover, it implies that different assets may have different levels of ambiguity, which may lead to a cross-section varying ambiguity premium. **Thus the research questions which will be answered by the empirical results of this chapter are:**

Does ambiguity aversion affect stock market participation at all? And if it does, what is the scale of the impact? Is there any asset class which is affected by the ambiguity aversion more than others?

3.2.3 Measuring ambiguity and market participation empirically

a. Ambiguity and SPF data

To answer the research questions empirically, the first step is to build an empirical ambiguity measure. According to Ellsberg (1961), although ambiguity is a subjective variable, a high ambiguity level could be objectively recognised by noting the situations where available information is unreliable or highly conflicting. Thus, this chapter relies on the measure proposed in the recent study by Anderson et al. (2009), which reflects dispersion in analysts' forecasts using the data from the Survey of Professional Forecasters (SPF) issued by the Federal Reserve.

The SPF data starts from 1968 and provides forecasts to many different horizons, based on the responses from large financial institutions. Each quarter the participants are asked about the forecasts for the previous, current and the next four quarters. The reason for forecasting the previous and current quarter is due to the fact that when the survey is conducted, these results are still not available. The variables used in this research are: forecasts of Gross National Product (GNP), before 1992Q1 and Gross Domestic Product (GDP) after 1992Q1; forecasts of the GNP deflator before 1992Q1 and the GDP deflator thereafter, and corporate profits after taxes. The number of forecasters participating in the SPF has varied: in its early stage, there were more than 100 forecasters, but this number declined through the 1970s and 1980s. After the SPF takeover in 1990 by the Federal Reserve Bank of Philadelphia the average (median) number of forecasters each quarter has been 36 (35), with a low of 29 and a high of 52.

Since the SPF contains the forecast of quarterly data such as GDP growth and inflation from different analysts, it can be used for modelling a forecast of real aggregate stock market return for each analyst by taking the forecasts of aggregate company profit, GDP growth and GDP deflator from SPF, and apply the Gordon Growth Model to generate the forecasted aggregate market return. The divergence among forecasts can thus be calculated as the variance in a beta weighted way in order to reduce the influence of the extreme forecasts. This is because such extreme forecasts are more likely to be ignored by investors and a few outlined results can have a very large impact on the normal variance estimations. When divergence among analysts regarding the future performance of stock markets is high, ambiguity is also likely to be high, since experts have conflicting views regarding the fundamentals of the economy.

As stated above, this measure of ambiguity suggested by Anderson et al. (2009) corresponds closely to the general definition of the ambiguity provided by Ellsberg (1961), as high ambiguity means that the factors are more difficult to justify, and the forecasts of returns among different analysts should be less consistent.

Meanwhile, since Ellsberg (1961) argues that ambiguity reflects poor information quality or conflicting information, it is reasonable to assume that the ambiguity measure is related to business environment factors. In fact, Anderson et al. (2009) states that the ambiguity variable is affected by the NBER dummy, which is an indicator of the US economy recession. Therefore, despite not being the primary interest of this research, it is also interesting to see whether the ambiguity measure can be forecasted by some other macro economical variables. Based on Sibley, Xing and Zhang (2013) the fundamental variables such as default rate, term spread, and unemployment could affect an investor's sentiment (Baker and Wurgler, 2006), thus it is assumed a possibility that such variables can affect how analysts perceive information, and make impact on the ambiguity measure.

b. Mutual fund flows

After choosing the ambiguity measure, the next step is to measure the stock market participation empirically. To observe the stock market participation, one way is to use the net mutual fund flow data, which measures the net value of all cash inflow and outflow through mutual funds. Therefore, the net flow of funds invested in the equity market should only give a direct measure of the stock market participation incremental. Previous research suggests that the fund flow can reflect investors' behaviour of avoiding disadvantages in the market. For example, Gibson, Safieddine and Titman (2000) found there are negative fund flows due to tax evasion before the capital gain is announced, and Karceski (2002) found the investors' return tracing behaviour on the underline asset market can be reflected in mutual fund flow as well. Therefore, the observations of ambiguity aversion behaviour on mutual fund-flow data are also anticipated. In addition, mutual funds meet the previous assumption of no short selling and mutual fund flows reflect majorly the investment decisions of individual investors, which is the main reason for the stock market non-participation: according to the Investment Company Institute (ICI) 90.4 million individuals or 44% of total US

households had 89% of the total mutual fund assets in 2011. The ICI estimated that around 68% of the mutual fund investors had over a half of their total assets in the mutual fund account.

The previous studies suggest that many factors can significantly affect a fund flow data. Ben-Rephael, Kandel, and Wohl (2011) found the fund exchange is linked to following-year returns, and Ben-Rephael, Kandel, and Wohl (2011) show that fund flow data is auto-correlated association with lagging market returns. Edwards and Zhang (1998) show a link between the fund flows and the aggregate returns of their underline market, and causality tests suggest that the asset returns cause the fund flows. Warther (1995) shows that there is no correlation between the fund flows and past aggregate market return. However, he does find that the fund flows are correlated with the contemporaneous aggregate returns, with the stock fund flows showing correlation with the stock returns and the bond fund flows showing correlation with the bond returns. Gallaher, Kaniel, and Starks (2006) argue that investors also react strongly to advertisements by fund, as they found that the fund flows are positively related to the adverting costs of the funds. Similarly, Huang, Wei, and Yan (2007) also believe that other information which helps to reduce the search costs increases the mutual fund flows.

In addition, Kamstra et al. (2011) show that the fund flows are influenced by capital gains and consumer behaviour, plus a need for liquidity at the end of the year. They argue that capital gains can result in an increase in a fund flow due to the reinvestment of dividends' distribution, and, when the saving rate is high, there are significantly higher capital flows into the mutual fund market. Also, since people have a need for consumption during Christmas and New Year time, they tend to withdraw money in November and December, and then bring the capital back into funds in January and February. Thus the model includes the following controlling variables: lagged value of fund flow data; fund return of the last year; the lagged stock market return; capital gains; saving rates are included at the end of the year (January, February, November, December); and advertising cost.

Furthermore, Kamstra et al. (2011) argue that changes in the level of risk aversion are negatively correlated to the mutual fund flows. According to their study, the investor's risk aversion level change in different seasons: people are less risk averse

in the sunny summer, and during the winter they become depressed and thus more risk averse. Their results show that there is a significant fund outflow from the stock market which runs into other safer assets, and the opposite way round during summer time. Thus it is assumed that changes in the risk have a similar impact, and therefore, including the changes of risk in the model nets out its effect. In addition, since it is assumed ambiguity causes limited participation, the changes in the market participation (which can be measured by the net flow data) should be correlated with the incremental ambiguity level. Thus, there is a need to test the hypothesis of whether ambiguity aversion leads to market non-participation as there should be a negative correlation between the changes in market ambiguity and the mutual fund flows.

3.3 Methodology

3.3.1 Measuring ambiguity with dispersion

Following Anderson et al. (2009), this chapter measures ambiguity by a beta-weighted average disagreement of forecasts. This measure of the ambiguity uses data from the Survey of Professional Forecasters (SPF), which has been run since 1968 and has a long time span. The forecasters in SPF majorly are from large financial institutions and the report is released by Federal Reserve officially, therefore it reflects professional opinions and can be widely applied to the market. The SPF forecasters give their forecasted results for economic variables such as GDP, GNP and deflator from the previous quarter to the next five quarters for the US market.

Based on the forecasted variables the aggregate corporate profit after tax and GDP deflator are used to work out the real aggregate corporate profit. Let π_t be the real aggregate corporate profit level at time t , P_t be a deflator at time t , τ_t be a nominal corporate profit at time t , the forecaster i 's expected real aggregate corporate profit for time $t+1$ at time t is:

$$E_{it}(\pi_{t+1}) = \frac{\tau_{t+1} * P_t}{E_{it}(P_{t+1})} \quad (3.14)$$

Then the Gordon Growth Model is applied to work out an aggregate level of the expected market return for $t+1$ at time t :

$$E_{it}(r_{t+1}) = E_{it}\left(\frac{\pi_{t+1}}{q_t}\right) + \xi_{it} \quad (3.15)$$

Here q_t is a value of all corporate values in the U.S, which is available from the Flow of Funds Accounts of The United States, prepared by the Board of Governors of the Federal Reserve System. ξ_{it} is a quarterly forecasted long-term rate of the real growth from forecaster i . The ξ_{it} between quarter m and n is:

$$\xi_{it} = \left(\frac{E_{it}\tau_n * E_{it}P_m}{E_{it}\tau_m * E_{it}P_n}\right)^{1/(n-m)} \quad (3.16)$$

Following Anderson et al. (2009) the longest forecast horizon used here is also four quarters, therefore $n-m=4$. Now, let X be a series of forecasts about the market return at time t , which has f_t forecasts in total and the forecasts are ranked from high to low. The weight of the i th lowest forecast is:

$$W_{it}(v) = \frac{i^{v-1}(f_t+1-i)^{v-1}}{\sum_{j=1}^{f_t} j^{v-1}(f_t+1-i)^{v-1}} \quad (3.17)$$

Here v is a parameter, which decides the weight for the extreme forecast: if $v=1$, it is flat weighted; the higher the value of v , the less weight the extreme forecast has. To estimate this parameter, Anderson et al. (2009) used a maximum likelihood approach. It is assumed that the excess market return is positively correlated to both ambiguity and excess volatility, and therefore, we can maximize the quasi-likelihood of the excess market return as:

$$r_{et+1} \sim N(b + \tau vol_t(w) + \theta amb_t(v), \quad \sigma^2 vol_t(w)) \quad (3.18)$$

where r_{et+1} is the excess market return, $amb_t(v)$ is the market ambiguity, $vol_t(w)$ is the excess market volatility (details will be introduced in section 3.3.3). Anderson et al. (2009) used data between the years 1968 and 2003 to obtain the parameters, and the recommended value of v is 15.346 (the robustness of the result is also tested by using different values of v) according to Anderson et al. (2009). Thus the disagreement can be measured by the beta-weighted variance of forecast series as below:

$$amb_t(v) = \sum_{i=1}^{f_t} W_{it}(v) \left[x_{it+1|t} - \sum_{j=1}^{f_t} W_{jt}(v) x_{jt+1|t} \right]^2 \quad (3.19)$$

Because the value of the ambiguity is relatively small, the original results are multiplied by 100. Since the ambiguity measure is produced quarterly and the fund flow data monthly, the quarterly ambiguity is decomposed into the monthly data by a linear method. To apply this method, the difference between ambiguity in quarter t and quarter $t+1$ is divided by 3, namely $(amb_{t+1} - amb_t)/3$, so the ambiguity level of the first month in quarter $t+1$ is: $amb_t + (amb_{t+1} - amb_t)/3$, and the ambiguity level of the second month in quarter $t+1$ is: $amb_t + 2 * (amb_{t+1} - amb_t)/3$. The models consider the change rather than the level of ambiguity, because the hypothesis here is that a degree of the equity market participation, as measured by total net assets held by mutual funds, is determined by the level of ambiguity. Therefore, fund flows, which represent the changes in total net assets, are determined by the changes in the ambiguity. In equilibrium, for a given level of the ambiguity, the fund flows are equal to zero, and so positive (negative) fund flows arise from a decrease (increase) in the ambiguity. Panel A in Figure 3.2 plots the quarterly ambiguity measure and Panel B the changes in the monthly interpolated series. Both measures produce spikes in the mid 1980s to mid 1990s and in the 2000s. Table 3.2 reports the summary of statistics for the constructed ambiguity series and for the changes in the monthly interpolated series. Both series are moderately positively skewed and leptokurtic.

3.3.2 Measure of mutual fund flow and mutual fund exchange

In this thesis the main source of the fund data is the Investment Company Institute (ICI), which provides detailed information about monthly fund flows and investment objectives ranging from January 1984 to December 2010. For each type of fund, the monthly data contains sales, redemption, exchanges in, exchanges out and total net assets. The funds are divided into five categories by asset class: equity, hybrid, corporate fixed income, government fixed income and money market following Kamstra et al. (2011). The main focus of this research is the equity fund category; however, since the ambiguity measure is based on the US stock market only, the equity fund group must be modified in order to exclude non-domestic stocks. Therefore, the funds that invest mainly outside the states: global equity, international equity, regional equity and emerging markets have been excluded in the analysis, and the final components of the equity fund group are: aggressive growth, growth, sector, growth and income, income equity. For the rest of fund families, international focused funds, such

as the long-term global bond, short-term global bond and other world bonds have also been selected. Table 3.1 reports the classification of the funds by the investment objective category.

[Table 3.1 here]

Following Kamstra et al. (2011), the net cash inflow into asset class i in month t is computed as:

$$Net\ Flow_{i,t} = \frac{Sales_{i,t} - Redmptions_{i,t} + Exchanges\ In_{i,t} - Exchanges\ Out_{i,t}}{TotalNetAssets_{i,t-1}} \quad (3.20)$$

Similarly, following Ben-Rephael et al. (2012), the net exchange into asset class i in month t is computed as:

$$Net\ Exchange_{i,t} = \frac{Exchanges\ In_{i,t} - Exchanges\ Out_{i,t}}{TotalNetAssets_{i,t-1}} \quad (3.21)$$

[Figure 3.1 here]

Figure 3.1 plots the net flows and exchanges for the equity group of funds. The net flows and exchange into equity had been much more volatile before 1993, with a large flow out of the equity asset class followed the October 1987 crash. Since 1994, net flows and net exchanges have been less volatile, but have also declined. Table 3.2 reports the summary of statistics for the net flows and exchanges for the equity asset class. The average net flow is 0.51%, representing a substantial increase in the total net assets over the sample, while the average net exchange is close to zero. Net exchanges are negatively skewed and strongly leptokurtic, while net flows have much lower skewness and kurtosis.

3.3.3 Control variables: conditional volatility

To ensure that the ambiguity measure is not just capturing the risk, a measure of conditional volatility is included in the model, following Anderson et al. (2009). Therefore, the weighted variance of past squared excess market returns is computed as proxy of risk. The weight for i^{th} lag is given by:

$$l_i(\omega) = \frac{(s+1-i)^\omega}{\sum_{j=1}^s (s+1-j)^\omega} \quad (3.22)$$

where s is the minimum number of trading days available for the previous 12 months over the entire sample, and the parameter ω determines the speed at which the weights decline as the lag length increases. In the empirical analysis Anderson et al. (2009) is followed and $\omega = 14.939$. (the estimating process of ω is discussed in section 3.3.1, with formula 3.18) The conditional variance is thus presented in the formula below:

$$cvar_t(\omega) = s \sum_{i=1}^s l_i(\omega) \left(r_{et,i} - \frac{1}{s} \sum_{j=1}^s r_{et,j} \right)^2 + 2s * \sum_{i=1}^{s-1} \sqrt{l_i(\omega) l_{i+1}(\omega)} * \left(r_{et,i} - \frac{1}{s} \sum_{j=1}^s r_{et,j} \right) \left(r_{et,i+1} - \frac{1}{s} \sum_{j=1}^s r_{et,j} \right) \quad (3.23)$$

where $r_{et,i}$ is the market excess return at i^{th} lag, which is computed as the daily Center for Research in Security Prices (CRSP) value-weighted index (series VWRETD) return minus the daily return of the three month T-bill. Figure 3.3 plots the monthly conditional variance together with the monthly ambiguity for the period from March 1985 to December 2010. It is clear that the two series capture very different dimensions of the market, with periods when the ambiguity is high but the conditional variance is low, and vice versa. Panel B of figure 3.3 further presents the data in a scatter plot, and shows that the ambiguity and the risk are not correlated. Table 3.2 reports the summary of statistics for the conditional variance. As expected, the conditional variance is highly positively skewed and leptokurtic. As with the ambiguity, the changes in monthly conditional variance are used in the regression, and Table 3.2 reports the descriptive statistics for this series. It can be seen that the changes in the conditional variance are also positively skewed and leptokurtic.

To make sure the result catches the effect of the ambiguity rather than the risk, several risk measures are used for the robustness check: realized volatility, GARCH

predicted volatility and market implied volatility. These are the risk measures which are applied widely and they can present market volatility from different angles. The realized monthly volatility is the sum of the squared daily returns for the month, based on the CRSP value-weighted return series (series VWRETD), and the implied market volatility is based on Chicago Board Options Exchange (CBOE) market volatility index (series VXO), which captures the expectations about the future volatility of aggregate market returns. There are two GARCH models used in this chapter. The first one is a simple GARCH model, and the second one is based on the studies of Glosten, Jaganathan, and Runkle (1993), which catches the asymmetry of volatility: considering the return series y_t , that follows a GARCH process. The conditional distribution of the series Y with time t is given as: $y_t | \psi_{t-1} \sim N(0, h_t)$, where ψ_{t-1} is the available information at time t-1, and in normal GARCH model the conditional variance h_t can be found as:

$$h_t = w + \sum_{i=1}^q \alpha_i y_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (3.24)$$

However there are differences in the influence of positive and negative return changes on the future volatility. Thus the model can be applied:

$$h_t = w + \sum_{i=1}^q (\alpha_i + I_{\varepsilon_{t-i} < 0} |\psi_i|) \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j} \quad (3.25)$$

where $\varepsilon_t = \sqrt{h_t} e_t$, $e_t \sim N(0,1)$. So there is an extra slope coefficient for each lagged squared error: the indicator $I_{\varepsilon_{t-i} < 0}$ is 1 if $\varepsilon_t < 0$ and 0 for the rest. The model in this chapter is estimated by using the same time period of baseline regression.

[Figure 3.3 here]

3.3.4 Other control variables for mutual fund flows

There are a number of other factors that have been proven to be important in explaining the mutual fund flows such as past fund returns (Ippolito, 1992; Sirri and Tufano, 1998), capital gains (Kamstra et al., 2011), past flows (Ben-Rephael et al., 2011), seasonal effects (Kamstra et al., 2011), advertising expenses (Gallaher, et al., 2006), past market returns (Ben-Rephael, et al., 2012) and savings (Kamstra et al., 2011). To capture serial correlation in the fund flows lagged monthly net flows and net exchanges for the past one, two and three months are included. Also included in the regression are the personal savings rate from the Bureau of Economic Analysis (series PSAVERT), and the capital gains data and advertising costs from Kamstra et al. (2011). The aggregate return of the

equity fund group over the previous 12 months is also added into the regression to capture a return-chasing behaviour. Following Ben-Rephael et al., (2011) and Oh and Parwada (2007), the regression further includes the aggregate market return over the last three months, from $t-3$ to $t-1$. In addition, transaction costs and liquidity needs are proposed as the reasons for the limited participation puzzle. Thus the measure of illiquidity proposed by Amihud (2002) is included in the model, which captures the responsiveness of prices to the trading volume. Following Amihud (2002), for each individual stock i the illiquidity measure in month t is defined as:

$$Iliquidity_{i,t} = \sum_{d=1}^j \frac{|ret_{i,d}|}{vol_{i,d}} \quad (3.26)$$

where j is the number of the available trading days in month t , and $ret_{i,d}$ and $vol_{i,d}$ are the daily return and volume of stock i in month t , respectively. Again the value weighted average is taken as a measure of market illiquidity level, and the rolling average is used over the previous three months, $t-3$ to $t-1$, in the regressions. Finally, dummy variables for the months of November, December, January and February are included to capture the year-end effect. Table 3.2 reports the summary of statistics for the control variables over the period from March 1985 to December 2010. Table 3.3 reports the correlations between the variables, and it can be seen that the changes in the ambiguity series is negatively correlated with both net fund flows and net fund exchanges.

[Table 3.2 here]

[Table 3.3 here]

3.3.5 Macro economic variables

As stated above, it is suspected that the ambiguity measure of Anderson et al. (2009) can be predicted by some of the fundamental macro variables. To test this hypothesis, based on Sibley et al. (2013) the following macro-economic variables are used: U.S. recession dummy (NBER), U.S. unemployment rate (Unemp), change in consumption (dCons), default spread (Def), term spread (Term), the value-weighted market return including dividends (VWRETD). The NBER recession dummy indicates the US recession period, which is a period between a peak and a trough based on the National

Bureau of Economic Research (NBER)'s business cycle data. The unemployment (Unemp) rate is taken from the Bureau of Labor Statistics Data, series ID: LNS14000000. The changes in consumption (Cons) are the changes in the consumption rate of real personal consumption expenditures per capita, which is constructed by dividing the CITIBASE series of the seasonally adjusted real consumption (excluding durables) by the number of monthly population according to the Bureau of Census, following Chen, Roll and Ross (1986). The default spread (Def) is defined as the difference in yields between Baa-rated corporate bonds and AAA-rated corporate bonds as it is performed in Fama and French (1989). The term spread (Term) is defined as the difference in yields between the 10-year Treasury bond and the 3-month T-bill as in Chen, Roll, and Ross (1986). *Vwretd* is the Centre for Research in Security Prices (CRSP)'s value-weighted stock market return index. The sample period is from Dec 1968 to Dec 2010, and all the variables are taken as a quarterly average to match the ambiguity data.

3.3.6 Regressions

After clarifying all the variables, the regression for the net flows is:

$$\begin{aligned}
 netflow_t = & a_0 + a_1\Delta amb_t + a_2\Delta cvar_t + a_3adv_t + a_4cap_t + a_5rfund_{t-12,t-1} \\
 & + a_6rmkt_{t-3,t-1} + a_7netflow_{t-1} + a_8netflow_{t-2} \\
 & + a_9netflow_{t-3} + a_{10}sav_t \\
 & + a_{11}llq_{t-3,t-1} + a_{12}Jan_t + a_{13}Feb_t + a_{14}Nov_t + a_{15}Dec_t + \varepsilon_t
 \end{aligned}
 \tag{3.27}$$

where adv_t is the aggregate cost of print advertising across all the funds divided by the previous year's total advertising costs, cap_t is the capital gains, sav_t is the personal savings rate, $rfund_{t-12,t-1}$ is the aggregate fund return for the previous year, $rmkt_{t-3,t-1}$ is the return on the value-weighted CRSP index over the last 3 months. $llq_{t-3,t-1}$ is the average market illiquidity during the previous three months, and Jan_t , Feb_t , Nov_t , and Dec_t are dummy variables that are equal to one in the respective month and zero otherwise.

For the net exchanges a similar model is constructed, but excluded are the savings variable and the seasonal dummy variables. The model for the net exchanges is therefore:

$$\begin{aligned}
 netexchange_t &= a_0 + a_1\Delta amb_t + a_2\Delta cvar_t + a_3adv_t + a_4cap_t + a_5rfund_{t-12,t-1} \\
 &+ a_6rmkt_{t-3,t-1} + a_7netexchange_{t-1} + a_8netexchange_{t-2} \\
 &+ a_9netexchange_{t-3} + a_{10}llq_{t-3,t-1} + \varepsilon_t
 \end{aligned}
 \tag{3.28}$$

Models 3.27 and 3.28 are constructed as a system of equations using the Generalized Method of Moments (GMM) (Hansen, 1982) with heteroscedasticity and autocorrelation consistent standard errors following the procedure of Newey and West (1987, 1994). In addition, the regression uses the Bartlett kernel, where the autocovariance lag structure is equal to $4*(N/100)^{2/9}$; N is the number of observations in the regression. The instruments in each regression model include the full set of independent variables. The hypothesis is that in both regressions the estimation of a_1 is negative and significant.

To examine the model's goodness of fit, the result of a Breusch-Godfrey autocorrelation test is reported. To perform the test, firstly the models shown in equations 3.26 and 3.27 are jointly estimated for all 5 fund families using GMM, and then the residuals are used to test whether they exhibit autocorrelation up to the sixth lag. For example, after the finishing estimating equation 3.27, there is residual series ε , and let $l_1 = lag(\varepsilon)$, $l_2 = lag2(\varepsilon)$, ..., $l_6 = lag6(\varepsilon)$. Then the following model is estimated with OLS method:

$$netflow_t = a_0 + a_1\Delta amb_t + \dots + a_{15}Dec_t + b_1l_1 + \dots + b_6l_6 + \varepsilon' \tag{3.29}$$

and it is tested whether: $b_1 = b_2 = \dots = b_6 = 0$.

Meanwhile, to test if ambiguity can be predicted by macro variables, this chapter uses a regression of the one-quarter-ahead ambiguity on the following macro economical variables: U.S. recession dummy (NBER), U.S. unemployment rate (Unemp), change in consumption (dCons), default spread (Def), term spread (Term), the value-weighted market return including dividends (VWRETD). To adjust for serial

correlation, the model also contains two lags for the one-quarter-ahead ambiguity. Therefore the regression is:

$$\begin{aligned}
 amb_{t+1} = & a_0 + a_1NBER_t + a_2unemp_t + a_3Dcons_t + a_4def_t + a_5term_t + a_6vwretd_t \\
 & + a_7amb_t + a_8amb_{t-1} + \varepsilon_t
 \end{aligned}
 \tag{3.30}$$

The model is also estimated using the Generalized Method of Moments (GMM) (Hansen, 1982) with heteroscedasticity and autocorrelation consistent standard errors following the procedure from Newey and West (1987, 1994).

3.4 Results analysis

This section reports the results of the empirical analysis. It firstly focuses on the equity asset class, and then considers the effects of the ambiguity on the non-equity fund flows and exchanges. Finally, the last part of the section conducts some of the robustness checks.

3.4.1 Ambiguity and Equity Fund Flows

Panel A of Table 3.4 reports the results of the estimating model (3.27) for the net flows for the equity asset class. The coefficient on the change in the ambiguity is negative and significant at conventional levels (-1.36, $p=0.00$). Therefore, in support of this study hypothesis, an increase in ambiguity is associated with a net outflow of capital from the equity mutual funds. In contrast, the changes in conditional variance do not have the statistically significant impact on net flows (-0.002, $p=0.41$). The savings variable a has significantly positive coefficient (0.03, $p=0.00$), which is consistent with a “free cash flow” effect on the fund flows. Consistent with the previous research (Kamstra et al., 2011) the lagged net fund flows from the previous month and three months are positive and highly significant, showing a strong autocorrelation in the flows.

[Table 3.4 here]

Panel B of Table 3.4 reports the estimation results from (3.28) for the net exchanges of the equity asset class. Similar to the net flows, the changes in the ambiguity are negatively associated with the net exchanges, and this relationship is statistically significant (-0.50 $p=0.01$). Again, the changes in the risk have a negative and significant impact (-0.002 , $p=0.02$). The one month and three months lag of the net exchanges is positive and statistically significant, reflecting a strong autocorrelation in this series.

These results suggest that an increase in the ambiguity has a negative and statistically significant impact on the net flows and the net exchanges, supporting this thesis hypothesis that an increase in the ambiguity leads to a reduction in the equity market participation. Moreover, while there is a clear link between the ambiguity and the net fund flows and exchanges, the impact of the risk is negative but not statistically significant for the fund flow regression. These results are consistent with Anderson et al. (2009), who shows that excess market returns have a strong positive association with the ambiguity, but much less related to the conditional variance. In addition, since the standard deviation of the ambiguity measure is 0.0013 and the average total net assets for equity funds is \$1.9 trillion; one standard deviation change in the ambiguity will on average yield a net flow of \$1.7 billion and a net exchange of \$0.05 billion.

3.4.2 Ambiguity and Different Equity Styles

The results in the previous section show that ambiguity aversion reduces stock market participation overall. However, since ambiguity varies across equities (Kelsey et al., 2013; Brenner and Izhakian, 2011; Antoniou et al., 2014), this effect is more significant in the funds that invest in more ambiguous stocks. This section tests this hypothesis by investigating the relationship between ambiguity and the fund flows for the five investment objective categories separately, namely: ‘aggressive growth’, ‘growth’, ‘sector’, ‘growth and income’ and ‘income equity’. According to the ICI definition, aggressive growth and growth funds invest in riskier, non-dividend paying stocks with a focus on capital gains. On the other hand, funds in the remaining categories focus on less risky, dividend-paying stocks (ICI Factbook, 2012).

It is noteworthy that a dividend policy is a signal about the stability of the expected profitability of the firm, as well as the indicator of the ambiguity level. The

study of Michaely et al. (1995) argues that firms are concerned about the penalties related to dividend omissions, and tend to pay dividends when the firms reach a mature stage. In this way firms are expected to be able make these payouts in the future consistently and thus avoid penalties incurred from dividend omissions. Therefore, dividend payers tend to be more mature and stable, and the information about them is supposed to be of better quality and less conflicting. This conclusion is also supported by other studies: Denis and Osobov (2008) show that dividend payers tend to be larger and more profitable companies; and Bossaerts, Ghirardato, Guarnaschelli and Zame (2010) also believe that companies at a growth stage are likely to show a significant ambiguity. Furthermore, firms that do not pay dividends are typically those with significant growth opportunities, and it is often difficult forecast their future development. Therefore, on average, the ambiguity is considerably higher for non-dividend payers which implies that the effect of ambiguity is stronger among the aggressive growth and growth categories.

Thus this section runs the models 3.27 and 3.28 again for different styles of equity funds, and the variables that control for the aggregate market characteristics (risk, returns and liquidity) are the same as those in the baseline model. The results are shown in Table 3.5, and, for brevity, the table reports the estimated coefficient on the changes in the ambiguity only. For the ‘aggressive growth’ and ‘growth’ categories, the coefficient of the changes in the ambiguity is negative and highly statistically significant for both fund flows and exchanges. In terms of economical significance, on average one standard deviation change in ambiguity will cause the net flow of ‘aggressive growth’ and ‘growth’ fund at \$7549 million and \$7318 million in respect. For net exchanges, on average one standard deviation change in ambiguity will cause ‘aggressive growth’ fund \$544 million and ‘growth’ fund \$390 million.

For the ‘growth and income’ and ‘income equity’ categories, the coefficient is negative but only significant for one of the two regressions. For the ‘sector’ category, the coefficient is insignificantly positive. In addition, regarding the absolute value of the estimations, the ‘aggressive growth’ and ‘growth’ categories are also much higher than the rest. Therefore, while the earlier results of this study show that an increase in the ambiguity leads to the flows and exchanges out of the equity asset class as a whole, it can be seen that within the equity asset class, the effect is concentrated on the funds that invest in more ambiguous, non-dividend paying assets.

[Table 3.5 here]

3.4.3 Ambiguity and non-Equity Fund Flows

This section examines the relationship between the changes in the ambiguity and the flows in the funds that invest in the non-equity asset classes, namely hybrid, government and corporate fixed income and money market. It is reasonable to expect that in response to an increase in the ambiguity in the stock market, investors will transfer funds into less ambiguous, non-equity investments.

Thus the models 3.27 and 3.28 are tested again for the different asset classes and the variables that control for the aggregate market characteristics (risk, returns and liquidity) are the same as those in the baseline model. Panel A (B) of Table 3.6 reports the estimated coefficient of the change in the ambiguity from the net flows (exchanges) model. For the net flows the coefficients on the ambiguity are not significantly different from zero. For the net exchanges, however, the coefficient for the money market asset class is positive and significant. Thus, as the ambiguity increases, the investors withdraw their capital from the equity funds and reinvest, at least partially, in the money market funds. According to the ICI definition, the money market funds invest in low risky, high-grade assets that receive full principal and interest within 90 days on average. Since in this study the ambiguity measure is based on the stock market's forecast of long-term growth, one possible explanation for this findings is that the investors seek safer assets with higher liquidity, when they are faced with a higher ambiguity in the expected stock returns. In terms of economical significance, on average one standard deviation of change in ambiguity will cause a \$29 million of capital running into money market funds. However, this conclusion needs to be treated with caution, as the tests show that the serial correlation in the regression remains significant.

[Table 3.6 here]

3.4.4 Ambiguity and macro variables

This section examines whether the ambiguity measure used in this thesis can be predicted based on Sibley et al. (2013) by the following macro economical variables: U.S. recession dummy (NBER), U.S. unemployment rate (Unemp), change in consumption (dCons), default spread (Def), term spread (Term), the value-weighted market return including dividends (VWRETD). Model (3.30) gives the regression to test this hypothesis, and Table 3.7 reports the result.

It can be seen that the NBER dummy is positively and significantly correlated to the one-quarter-ahead ambiguity ($p=0.01$), while the stock market return variable VWRETD is negative and significant at 10% level ($p=0.07$). Therefore, when a recession is current, or the current market return is low, investors are more likely to perceive a higher ambiguity in the future. In addition, since one standard deviation change in the ambiguity will on average yield a net flow of \$1.7 billion and a net exchange of \$0.05 billion, this result can also be useful for fund managers, as it will help them to forecast the period of a high ambiguity and prepare for the outflows in advance.

[Table 3.7 here]

3.4.5 Robustness

a. Quarterly regression

As previously discussed in the methodology section, the baseline regression (Table 3.4) uses the linear interpolation for the ambiguity measure to obtain monthly estimates and hence increase the power of this study's tests. This section estimates the models shown in the equations 3.27 and 3.28 using non-interpolated, quarterly data. Newey and West (1987, 1994)'s heteroscedasticity and autocorrelation consistent standard errors are in continuous use here.

The net flows and exchanges are calculated on a quarterly basis. The changes in the ambiguity and conditional variance are equal to $\Delta Qamb_t = amb_t - amb_{t-3}$, and $\Delta Qcvar_t = cvar_t - cvar_{t-3}$ respectively. Quarterly capital gains, savings and

advertising costs are equal to the sum of the monthly values over each quarter: $Qadv_t = \sum_{k=t-2}^t adv_k$, $Qcap_t = \sum_{k=t-2}^t cap_k$ and $Qsav_t = \sum_{k=t-2}^t sav_k$. Lagged market return, illiquidity premium and fund return are defined as previously. To control auto-correlation, the lagged value of last quarter and last year is added into regression. The quarterly regressions are in the form of:

$$\begin{aligned} netflow_t = & a_0 + a_1\Delta Qamb_t + a_2\Delta Qcvar_t + a_3Qadv_t + a_4Qcap_t \\ & + a_5rfund_{t-12,t-1} + a_6rmkt_{t-3,t-1} + a_7netflow_{t-3} \\ & + a_8netflow_{t-12} + a_9Qsav_t + a_{10}llq_{t-3,t-1} + a_{11}Dec_t + \varepsilon_t \end{aligned} \quad (3.31)$$

$$\begin{aligned} netexchange_t & \\ = & a_0 + a_1\Delta Qamb_t + a_2\Delta Qcvar_t + a_3Qadv_t + a_4Qcap_t \\ & + a_5rfund_{t-12,t-1} + a_6rmkt_{t-3,t-1} + a_7netexchange_{t-3} \\ & + a_8netexchange_{t-12} + a_9llq_{t-3,t-1} + \varepsilon_t \end{aligned} \quad (3.32)$$

The results for these quarterly regressions are shown in Table 3.8. Even though the number of observations in these models are reduced, it has been found that increases in the ambiguity are negatively and significantly related to both fund flows (Panel A: -1.48, $p=0.00$) and fund exchanges (Panel B: -0.68, $p=0.00$). In terms of economical significance, on average one standard deviation in change of ambiguity will cause the quarterly net flow to change \$5.4 billion, and quarterly net exchanges to change 0.21 billion.

[Table 3.8 here]

b. Sentiments

Another concern about this study's result of the ambiguity index is that it may be related to the investor's sentiments toward the stock market, which has been shown affect the equity investments through the mutual funds (Ben-Rephael et al., 2012). Therefore another robustness test has been run in order to control for this effect. In an expanded version of the baseline model the sentiment index of Baker and Wurgler (2006) and the median forecast from the SPF data have been controlled. These two variables represent investors' attitudes towards the stock market's performance, and therefore, ensure that

findings of this research do not simply reflect the effect of investors' sentiments on the market participation.

[Table 3.9 here]

The results from the estimation of this expanded model are shown in Table 3.9, in which the ambiguity variable continues to be negative and significant (coefficient: -1.56, p -value=0.00 for the net flow; and coefficient: -0.54, p -value= 0.01 for the net exchange, which means on average one standard deviation change in ambiguity will cause the net flow to change 1.93 billion and net exchange to change \$0.06 billion). The sentiment index of Baker and Wurgler (2006) is not significant, which is consistent with the previous studies of Ben-Rephael et al. (2012). The coefficient on the median SPF forecast is positive but also insignificant. Thus the results in Table 3.9 suggest that the baseline findings in Table 3.4 do not capture the effect of the investors' sentiments toward the stock market.

c. Alternative risk measures

In addition, to ensure the main factor that affects participation is the ambiguity rather than the risk, as an experiment the regression in Table 3.4 has been run with alternative measures of risk, including realised volatility, GARCH predicted volatility and market implied volatility. The realised volatility is calculated as the sum of daily squared market returns, using the CRSP value of the weighted return series. The GARCH predictions are the forecasts from a simple GARCH model, and an asymmetric GARCH model is based on Glosten, Jaganathan, and Runkle (1993). These models allow for asymmetries in the effect of positive and negative current returns on the future volatilities due to leverage effects. Both GARCH models are estimated within the full sample period of this study. The market implied volatility variable is the CBOE market volatility index (series VXO), which captures the expectations about the future volatility of aggregate market returns. Due to the data availability on the volatility index, the time period is 1986-2010. The same model from Table 3.4 is estimated for both net flows and net exchanges, replacing the original risk measure with these alternative definitions.

[Table 3.10 here]

The results are shown in Table 3.10. Briefly, this table reports the results for the coefficients on the change in the ambiguity and the risk. The risk is significant for the implied volatility and both GARCH predictions, but is negative and insignificant when it is measured using the realised volatility. The coefficient on the ambiguity is negative and significant in all four cases, suggesting that the findings of this research are robust to the different risk specifications, and, therefore, the main factor that affects market participation is ambiguity rather than risk.

d. Different weighting parameters of ambiguity

As has been previously discussed in this thesis, the ambiguity measure is constructed by a beta-weighted average of analysts' opinions to reduce the influence of extreme forecasts, and the weighting parameter ν is 15.346 according to Anderson et al. (2009). To examine whether the result is robust to this weighting parameter, the baseline regression is tested again for the equations 3.27 and 3.28; with different ambiguity series generated by $\nu=1, 5, 10,$ and 20 respectively. Table 3.11 shows the results of the regression and, for brevity, it reports an estimation of the changes in the ambiguity only. It can be seen that, while the analysts' forecast is not flatly weighted ($\nu \neq 1$), the changes in the ambiguity are negative and significant for both the net flow and net exchange among all three different ν values. Therefore, the results are robust to the different values of the weighting parameter ν . On the other hand, while the value of ν decreases and, therefore, the weight of extreme forecasts increases in the ambiguity measure, the significance of the ambiguity decreases as well. If $\nu = 1$ and the forecasts are flatly weighted, the estimation of the ambiguity becomes insignificant. This result is consistent with Anderson et al. (2009)'s findings, which state that in order to measure the ambiguity correctly the extreme forecasts must be diluted.

[Table 3.11 here]

3.5 Limitations and future developments

This chapter studies the impact of the ambiguity aversion on the stock market participation, through the mutual fund flows. The results shows that the fund flows and exchanges are negatively correlated to the changes in the market ambiguity, which is measured by the disagreement among the analysts' opinions from the SPF. This research provides the empirical evidence that supports the existence of ambiguity aversion in the stock market and the corresponding premium. However, there are still limitations in this study:

Firstly, a mutual fund only takes part in stock market investments, thus the impact of the ambiguity on mutual funds cannot fully explain the limited stock market participation. Secondly, a disagreement in the analysts' opinions is the only one of many ways to measure the ambiguity empirically, thus it may not fully present the real ambiguity level. Thirdly, there are still other factors which could contribute to the changes of the mutual fund flows and exchanges, such as the fast development of substitute products, for example, the exchange trade funds. Due to the limitation of the data source, such factors have not been presented in the model and their impact is unclear. Finally, this research is based solely on the American stock market; therefore, it is still questionable if the impact of ambiguity aversion on the market participation is a worldwide phenomenon.

Thus future studies can extend this chapter by a variety of approaches: firstly, alternative ambiguity measures such as firm size, firm age, and analyst coverage (Zhang, 2006) can be used in order to test the model developed in this chapter. This would provide a more well-rounded understanding of how ambiguity affects stock market participation. Secondly, the model can be extended into other markets outside the US. Finally, a further test can be conducted on other approaches of stock market participation in order to see whether the ambiguity solely affects the mutual fund flows or the market participation as a whole. The next chapter will conduct further research based on the Survey of Consumer Finance (SCF) to investigate the influence of ambiguity on the aggregate level of individual stock market participation.

3.6 Conclusion

Theories about the ambiguity indicate many possible applications in the financial markets such as home-bias (Uppal and Wang, 2003), price momentum (Caskey, 2009) and the equity premium puzzle (Chen and Epstein, 2002). Meanwhile, the limited stock market participation is a longstanding puzzle in the area of finance and there are many explanations including frictions-based and behavioural ones. In this chapter, one behavioural explanation for non-participation is the consequence of ambiguity aversion having been tested. According to the ambiguity-based explanation, stocks involve both risk and ambiguity, and since investors are ambiguity averse, their propensity to invest in stocks is lower than is predicted by EU models, and is negatively correlated to the level of the ambiguity in the marketplace. This measure is negatively and significantly correlated to the lagged value of the stock market return, and positively and significantly correlated to the lagged value of the US recession indicator.

Market participation is measured by flows of the capital in and out of the U.S. equity mutual funds. In this research the measure of the ambiguity is based on a recent study by Anderson, et al. (2009) and reflects the dispersion in analysts' implied forecasts about market returns. The results show that if other factors that affect the fund flows are controlled, increases in the ambiguity are significantly and negatively correlated to the equity fund flows and exchanges, and thus it supports the notion that limited stock market participation arises, because the stock market contains ambiguity, which is disliked by investors. The results of this chapter provide empirical evidence that supports the existence of ambiguity aversion in the financial markets, and point out that ambiguity aversion is an important factor in the decision-making process regarding stock market participation.

In addition, the initial interest of this research has been in the study of ambiguity heterogeneity in different asset classes. It was suspected that the non-dividend payers are more ambiguous than other stocks, since the dividend paying companies are normally more mature and stable companies. This assumption has been confirmed by the result, as the fund flows of the 'aggressive growth' and 'growth' groups are more sensitive to the changes of the ambiguity level compared to others. According to the ICI definition, the 'aggressive growth' and 'growth' funds invest in riskier, non-dividend paying stocks with a focus on capital gains, and thus are more ambiguous. Furthermore, it is found that, while high ambiguity reduces the capital within equity funds, the net

exchanges of the money market fund are positive. Based on the ICI definition, the money market funds invest in low risk, high-grade assets that receive full principal and interest within 90 days on average. Thus investors could switch to the short-term investments with a higher liquidity to avoid stock market uncertainty.

To test the robustness of this study result, four alternative risk measures have been applied to ensure the regression does not catch the influence of the risk only; and the market sentiment variables have been added to ensure the result is not just a pure effect of market sentiments. The regression with quarterly data has also been run in order to check whether interpolation would affect the conclusion, and to test the model with different values of the weighting parameter ν , which is used to construct the ambiguity measure. The results show that the conclusion of this chapter is robust to all of the tests.

Chapter 4. Heterogeneity in Household Ambiguity Aversion

4.1 Introduction

The previous chapter investigated the factors that affect stock market participation in aggregate level and the sensitivity of different asset classes to ambiguity. This chapter mainly focuses on the factors which can affect individual investors' ambiguity aversion, and investigates the reasons for ambiguity aversion behaviour.

Easley and O'Hara (2010)'s model argues that 'while ambiguity can affect stock market participation and stock price, the impact is a subject to the level of individual's ambiguity aversion'. Their work suggests that market participation relies on the most optimistic and the most pessimistic investors, when facing the same level of ambiguity. In other words, individual ambiguity aversions are heterogeneous, and can affect stock market participation. To explore the factors which determine individual ambiguity aversion further, one must first understand the psychology behind ambiguity aversion. Previous studies state that ambiguity aversion may come from the outcome generating process of perceived information, as people prefer familiar events over unfamiliar ones. Frisch and Baron (1988) argue that when people believe they have insufficient information, they will avoid making decisions. Chow and Sarin (2001), Fox and Tversky (1995) and Fox and Weber (2002) conclude that this happens more often if an alternative process with a higher perceived informational content is available. Therefore, ambiguity aversion is likely to be the consequence of avoiding the unpleasant experience of decision making under an ambiguous situation. While some studies believe that the fear of negative evaluation by others (FNE) is the cause of ambiguity aversion (Curley, Yates, and Abrams (1986)), others suggest that self-evaluation, anticipated cognitive dissonance or regret are the causes of the observed ambiguity aversion (Krahmer and Stone, 2010).

The theories above provide a good explanation of how ambiguity aversion could work on an individual level and what factors could determine ambiguity aversion; however these theories have not been empirically tested. Therefore, the purpose of this chapter is to test Easley and O'Hara (2010)'s model empirically, in order to examine

whether heterogeneity in ambiguity aversion could affect stock market participation, and also to explore the psychological reasons behind it.

These tests are run by using Survey of Consumer Finance (SCF) data from the Board of Governors of the Federal Reserve. This survey has been conducted every three years since 1989. It gathers demographic and financial information from a large number of different households. This chapter includes the tests which use data from 8 different surveys, from the year 1989 to the year 2010. The stock market participation is measured from the stock holding status of households from the SCF: if a household has either direct stock holding or stock-based funds, it is considered a participant in the stock market. To be consistent with the previous chapter, the dispersions of analysts' opinions are used (Anderson et al, 2009) as an ambiguity measure. In addition, since the SCF uses a multiple imputation technique to impute missing values, standard errors for multiple imputations (RII) are adjusted following the method of Rubin (1987).

As a further development of the results in the previous previous chapter, this chapter firstly tests if ambiguity affects individual stock market participation in general, rather than only via mutual funds. After controlling the variables that affect individual stock market participation, such as income, race, age, marriage, education and anchoring of past prices, the results show that ambiguity can still significantly reduce household stock market participation.

Furthermore, to test whether heterogeneity in ambiguity aversion exists in households as Easley and O'Hara (2010) suggested, several hypotheses are developed, based on the Fear of Negative Evaluation by others (FNE) and the self-evaluation theory. Since FNE suggests that negative evaluation from others is the main cause of ambiguity aversion, in this chapter the assumption is made that households with a higher income (>\$125000 income in 1993) and education level (with college degree) should be less ambiguity averse, since such households normally have a better social status and are less subject to others' opinions. For the same reason, it is assumed here that older people (>65 years old) are less ambiguity averse, because FNE is partly caused by a low self cognitive level, such as over-estimating awkwardness and feeling a lack of confidence or self-assurance. But when individuals become more mature, they improve their self cognitive level, and therefore suffer less from FNE. For example, Aydin (2008) finds that younger students have a higher level of FNE than mature students in a foreign language environment.

FNE also suggests that when people form a group, their decision making has more ambiguity aversion than the average of individuals, thus another assumption is made that married couples are more ambiguity averse. In addition, based on FNE, people with higher ambiguity aversion are less social; therefore it is assumed here that a lower social capital brings a higher ambiguity aversion level.

On the other hand, the self-evaluation theory believes that the negative feelings of regret are the main reasons for ambiguity aversion, thus those who are more optimistic are naturally less ambiguity averse. The level of optimism is measured by the response towards the future US economic situation, and those who believe the future will improve are considered more optimistic and are assumed to be less ambiguity averse. Similarly, people who experienced a low life-time stock market return before are more likely to be pessimistic towards the stock market and, therefore, become more ambiguity averse. In this chapter, based on Malmendier and Nagel (2009), a weighted life-time stock market return is produced with an assumption that those who have had a negative life-time return should be more ambiguity averse. It is also noteworthy that, in contrast to FNE, self-evaluation theory predicts that low income households should have less ambiguity aversion as they will regret it more if missing the rewards from ambiguous stocks, and *vice versa* to high income households. Finally, the research reference also points out that people who smoke are less ambiguity averse, so this hypothesis is also tested here.

The result obtained shows that older or more optimistic investors are less ambiguity averse, while investors with high incomes, married or who have experienced negative life-time stock market returns are more ambiguity averse. These results support Easley and O'Hara (2010)'s model empirically, and also show that ambiguity aversion is the combined result of both: fear of negative evaluations from others and regret from self-evaluations.

4.2 Literature review

4.2.1 Heterogeneity in ambiguity aversion and market participation

As Easley and O'Hara (2010) state, the presence of ambiguity renders the market price to be derived from fair pricing and thus can cause non-trade. They come to the conclusion that, instead of 'the price that would be received to sell an asset or transfer a liability in an orderly transaction between market participants at the measurement date' (definition of *fair pricing* from Financial Accounting Standards Board), ambiguity makes the prices become individual beliefs about the best and the worst case of outcomes.

Assume there is one risky and one risk-free asset in the market, and two periods 0 and 1. The risk-free asset always has a value of 1 per unit, and the risky asset has a price p_t , $t = 0,1$ at each period and an uncertain future pay off at the end of period 1, namely \hat{v} , which is believed to be normally distributed with a variance σ^2 . However, the agents (with a total amount of 1) disagree with the value of the mean, and for agent i the expected mean is \bar{v}_i and for at least two agents i, j we have $\bar{v}_i \neq \bar{v}_j$. Trader i 's endowment of the risky asset is \bar{x}_i . The per capita endowment of the risky asset is thus $\bar{x} = \sum_{i=1}^I \bar{x}_i$. At time 1 an unanticipated shock to trader's beliefs about the future value of the risky asset occurs and traders can re-trade. All traders have constant absolute risk aversion (CARA) utility of wealth, w , at the end of period 1 and at time 0 each trader maximizes the expected utility of wealth w , given beliefs and the time zero price of the risky asset. Traders are unaware of the possibility of a shock at time 1 and thus do not plan for it. Therefore, after the shock, the set of possible declines in the expected value of the risky asset is described by $1 - \alpha$ and $\alpha \in [\underline{\alpha}, \bar{\alpha}]$ with $0 < \underline{\alpha} < \bar{\alpha} < 1$. If trader i chooses the portfolio (x_{i0}, m_{i0}) at period 0, his future wealth after period 1 will be the random variable $\tilde{w}_i = \tilde{v}x_{i0} + m_{i0}$. So his expected utility from such random wealth is:

$$(\bar{v}_i - p_0)x_i - \frac{\sigma^2(x_i)^2}{2} + \bar{w}_i \quad (4.1)$$

where \bar{w}_i is the trader i 's initial wealth. Thus trader i 's period 0 demand for the risky asset has the standard form:

$$x_i^* = \frac{\bar{v}_i - p_0}{\delta^2} \quad (4.2)$$

In equilibrium, per capita demand for the risky asset must be equal to per capita supply, so:

$$\frac{1}{I \sum_{i=1}^I x_i^*} = \bar{x} \quad (4.3)$$

The average belief about the future mean value of the risky asset is:

$$\frac{\sum_{i=1}^I \bar{v}_i}{I} = \hat{v} \quad (4.4)$$

Therefore, the price and individual holding of the assets at period 0 are:

$$p_0 = \hat{v} - \delta^2 \bar{x}, \quad x_{i0}^* = \left(\frac{\bar{v}_i - \hat{v}}{\delta^2} \right) + \bar{x} \quad (4.5)$$

Thus, the period 0 price of the risky asset is the average mean future value reduced by a factor that compensates traders for holding the market risk. The individual holding is generated according to their relative optimistic level in relation to the average optimistic level. Now at period 1, let (t_i, m_i) be the trader i 's trade of risky asset and cash. The trader i 's period 1 budget constraint is then:

$$p_1 t_i + m_i = 0 \quad (4.6)$$

Thus the trader i 's future wealth, given his endowment and trade is then:

$$\widetilde{w}_{i1} = \widetilde{v} x_{i0}^* + t_i (\widetilde{v} - p_1) + m_{i0}^* \quad (4.7)$$

Therefore, for any expected future value of the risky asset \overline{v}_{i1} , that the trader i believes is possible, his expected future wealth is:

$$\widetilde{w}_{i1} = \overline{v}_{i1} x_{i0}^* + t_i (\overline{v}_{i1} - p_1) + m_{i0}^* \quad (4.8)$$

And the variance of his future wealth is:

$$\sigma^2(x_{i0} + t_i)^2 \quad (4.9)$$

If the trader wants to carry out a transaction t_i , it should satisfy at least two conditions:

1. It will add utility in any possible distribution of returns.
2. It will be better than any other possible transaction. So the trade t_i must firstly solve the following equation for all $\overline{v}_{i1} \in [\underline{\alpha} \overline{v}_i, \overline{\alpha} \overline{v}_i]$:

$$\bar{v}_{11}x_{i0}^* + t_i(\bar{v}_{11} - p_1) + m_{i0}^* - \frac{\sigma^2(x_{i0}^* + t_i)^2}{2} > \bar{v}_{11}x_{i0}^* + m_{i0}^* - \frac{\sigma^2(x_{i0}^*)^2}{2} \quad (4.10)$$

Meanwhile, for all $\bar{v}_{11} \in [\underline{\alpha}\bar{v}_1, \bar{\alpha}\bar{v}_1]$ there should be no trade t' which can result:

$$\begin{aligned} & \bar{v}_{11}x_{i0}^* + t'(\bar{v}_{11} - p_1) + m_{i0}^* - \frac{\sigma^2(x_{i0}^* + t')^2}{2} \\ & > \bar{v}_{11}x_{i0}^* + t_i(\bar{v}_{11} - p_1) + m_{i0}^* - \frac{\sigma^2(x_{i0}^* + t_i)^2}{2} \end{aligned} \quad (4.11)$$

Thus the condition for an agent i not to trade is that the price satisfies:

$$p_1 \in [\underline{\alpha}\bar{v}_1 - \sigma^2x_{i0}^*, \bar{\alpha}\bar{v}_1 - \sigma^2x_{i0}^*] \quad (4.12)$$

And for the entire market to have no trade is:

$$\cap_{i=1}^I [\underline{\alpha}\bar{v}_1 - \sigma^2x_{i0}^*, \bar{\alpha}\bar{v}_1 - \sigma^2x_{i0}^*] \neq \emptyset \quad (4.13)$$

This means no trade happened when:

$$\frac{\text{Max}_i \bar{v}_1}{\text{Min}_i \bar{v}_1} < \frac{1-\alpha}{1-\bar{\alpha}} \quad (4.14)$$

We can see here that the probability of no trade will depend on two conditions: the dispersion of opinions among investors and the level of ambiguity. The higher the dispersion, the less will be the probability of no trade, meanwhile if the ambiguity goes higher (which enlarges $\frac{1-\alpha}{1-\bar{\alpha}}$) the result is more likely to be a cessation in trading. Once such 'no trade' equilibrium is established regardless of the price change the total demand and supply will not match at all.

Furthermore, in this model ambiguity level is no longer the only factor which affects market participation; the participation level now also depends on the investors' dispersion $\frac{\text{Max}_i \bar{v}_1}{\text{Min}_i \bar{v}_1}$. If we assume that the level of ambiguity $[\bar{\alpha}, \underline{\alpha}]$ is absolute and objective, it will be the same to all investors. Thus market participation will rely on the most optimistic and pessimistic investors, which gives $\text{Max}_i \bar{v}_1$ and $\text{Min}_i \bar{v}_1$. As we can see, the agents with beliefs of $\text{Max}_i \bar{v}_1$ and $\text{Min}_i \bar{v}_1$ react totally differently towards the same ambiguity level because they set the bidding and asking prices respectfully. Thus, in other words, they have different ambiguity aversion levels, which determine their market participation rate.

4.2.2 Psychological roots of ambiguity aversion

As has been discussed above in the previous section, households may have heterogeneity in ambiguity aversion and thus may react differently towards the same level of ambiguity. Furthermore, since those with the belief of $\text{Min}_i \bar{v}_i$ are always the sellers of assets under ambiguous circumstances, they are the main reason for low market participation. This assumption is supported by the result of Dimmock, Kouwenberg, Mitchell and Peijnenburg (2013)'s experiment, which revealed that while the majority of investors are ambiguity averse, a small fraction are still ambiguity seeking. Therefore, it is important to understand which categories of investors could have a higher ambiguity aversion level than others. To find the factors that could lead to high ambiguity aversion, the following section explores the subtle psychological causes of ambiguity aversion.

A key point of the explanation of ambiguity aversion concerns the perceived informational content of the outcome generating process: people prefer to avoid making decisions when they think they have insufficient information (Frisch and Baron, 1988). As Chow and Sarin (2001), Fox and Tversky (1995), Fox and Weber (2002) believe, this particularly happens if an alternative process with a higher perceived informational content is available. In other words, people prefer familiar events over unfamiliar ones. In Ellsberg's (1961)'s case, this effect leads to a preference for the urn with a known probability of winning and where subjects feel more knowledgeable, and such behavior officially is described as an ambiguity aversion. In the section below the mainstream explanations of ambiguity aversion is reviewed, which will help to further derive the testing variables.

a. Negative evaluation by others

Curley, Yates, and Abrams (1986) argue that the fear of negative evaluation by others (FNE) is the cause of ambiguity aversion: the preference for the more informative process (i.e. the preference for known events or at least with known probability, which are risky) is driven by the expectation that one's actions or judgments may be difficult to justify in front of others. When people choose an ambiguous option and receive a bad outcome, they fear criticism from others as they fear that a failure in the right choice of ambiguous events could be an indicator of inferior decision making skills. On

the other hand, such criticism is easier to counter in a risky choice situation because a bad outcome is easier to explain if it is a result of bad luck rather than of having made a bad choice. This explanation is backed up by the empirical studies of Simonson (1989) and Lerner and Tetlock (1999). They found that when somebody is required to view an issue which is unknown and for which no prior commitment to any course of action exists, they simply make the decision which they believe is most easily justifiable to others rather than making the optimal one. In this way they can minimize the risk of being judged by others negatively regarding their decision-making qualities.

Curley, Yates, and Abrams (1986)'s experiments found that an increase in the number of people aware of the decision of a particular individual leading to an increase in ambiguity aversion stronger than any other factors they had tested. Also the studies of Tversky (1998), Fox and Weber (2002), and Heath and Tversky (1991), show that ambiguity aversion increases with the perception that others are more competent and more knowledgeable. Keller and Sarin (2007) also obtained similar results. Before the experiment, they assumed that a pair of people would reveal less ambiguity aversion in decision making than the average level of two separate individuals, due to information sharing. However, the result shows that the decisions made by the pair are more ambiguity averse, and hence this confirms that the FNE is a dominating psychological cause of ambiguity aversion.

Moreover, Bochner (1965) believe that due to FNE, individuals with a higher ambiguity aversion level are less likely to adapt to new technology or, even, are less social. On the other hand, there is evidence which presents the fact that social networking with other people enhances learning through ambiguous cases and helps to reduce ambiguity aversion. Foster and Rosenzweig (1995) show that a farmer reveals ambiguity aversion when adapt to new technologies; however, this can be reduced by a higher level of the social networking. For example, the new high-yield-variety rices are known to have high productivity but are sensitive to inputs such as water, fertilizer and sowing times, and the optimal level of inputs are subject to the quality of soil and to temperature, which is very ambiguous to individual farmers. As a result, farmers have significant ambiguity aversion in adapting to the new rice variety and prefer the traditional one, because they do not want to be the ones in the village to break the old tradition and fail, and therefore suffer from FNE. However, if it is arranged that the farmers are able to share their experience about the new rice variety with each other, they can update their knowledge quickly due to the much larger amount of trials. As a

result, they know they are much less likely to fail, and therefore it is much easier to justify their choice of choosing the new variety.

Finally, Gilbert (1998) finds that the results of FNE, such as shame, social anxiety and depression are highly related to feeling inferior and having a low social rank. McGuire and Gilbert (1998) further state that for households that have higher FNE, they tend to have a lower income, are less sociable and have a smaller living space. Therefore, since household high income and education is correlated to higher social rank, such households are supposed to suffer less from FNE and become less ambiguity averse.

Based on this theory, in this chapter several assumptions are made in relation to the possible testing variables of ambiguity aversion. Firstly, and intuitively from the results of Keller and Sarin (2007), married people should have a higher ambiguity aversion level than single ones. Secondly, based on McGuire and Gilbert (1998), it is reasonable to assume that the people with a higher social status are less sensitive to FNE. Therefore, investors with a higher income will be less ambiguity averse, and similarly those who have a higher education level. Thirdly, since more ambiguity averse individuals tend to be less open or sociable, it is assumed that investors who are more socially active should be less ambiguity averse. Finally, since older people have a more mature personality and a higher level of self-cognition, they should be less influenced by FNE. Chen, Kim, Nofsinger, and Rui (2003) used investors' ages as measure of investor sophistication, and found that older investors were more confident with their decisions and even had a higher likelihood of being over-confident. Thus it is expected that compared to younger people, older ones have a lower ambiguity aversion level.

b. Self-evaluation

Self-evaluation, anticipated cognitive dissonance, or regret, are also reasons for the observed ambiguity aversion. In other words, an evaluation of oneself can have a stronger effect than the fear of negative evaluation from others. The self-evaluation theory focuses on the 'anticipated regret' of individuals from future self-evaluation. It explains decision making as a process which results in minimizing self regret. The theory was originally developed by Ellsberg (1961), and Kraumer and Stone (2010) further model such 'regret' as an emotional gain after an agent finishes his choice.

Recalling Ellsberg's urns as an example: the agent has to choose between a risky urn (50 black and 50 white balls) and an ambiguous urn (100 balls, comprising an unknown number of black and white balls), and after his choice one ball is randomly picked from the chosen urn. If the ball is black, then the agent gets a \$100 reward, otherwise he gets 0. In this case, after the ball has been drawn, the agent compares the actual pay-off with the prior expected pay-off, which would have been the best, based on his ex-post knowledge. This prior expected pay-off is the agent's reference point, and, in addition to his real pay-off, the agent obtains an emotional pay-off, which is equal to the difference between his actual pay-off and the reference point. When the agent chooses the uncertain urn, the outcome shows the true information about the urn, and thus he revises his view of the urn's expected pay-off. If he wins the bet, it indicates that his actual choice has been the best choice, since his posterior belief is that there are more black balls than white balls in the ambiguous urn. Therefore, in this case, his reference point is the uncertain urn and he obtains an emotional gain, because he compares his pay-off of \$100 to the posterior mean of the uncertain urn, which is greater than \$50 and less than \$100. In the that case he loses the bet, his posterior belief is that there are more white balls than black balls. Accordingly, he believes that the risky urn would have been the best choice. Thus, in this case, his reference point is the risky urn, and he experiences regret because he compares his actual pay-off of \$0 to the expected pay-off of the risky urn of \$50.

When the agent chooses the risky urn instead, no outcome leads to the state that the agent will revise his assessment of either urn's expected pay-off. Therefore, regardless of winnings or losings, his choice is optimal because in either event his reference point is the risky urn, and he compares his actual pay-off to the posterior mean of the risky urn, which is \$50. However, if he chooses the uncertain urn, his regret in the event of a bad outcome is greater than his rejoicing in the event of a good outcome, because in the latter event he changes his assessment of the optimal choice in the light of his ex-post information, pushing up his reference point and diminishing his rejoicing. Consequently, the choice of the uncertain urn leads to higher expected regret and leads to the preference of riskier events over ambiguous ones, which we call ambiguity aversion. The worst-case scenario model can thus be seen as an extreme case of self-evaluation theory application, where the agent aims for minimum level of anticipated regret, because he is always prepared for the worst outcome. In Ellsberg's example, this

means that the agent will set his posterior choice of ambiguous urn as 0 black balls, therefore for any later outcome his regret will not be positive.

Such an explanation naturally leads to the assumption that households with a lower income should be less ambiguity averse in the stock market. As stated in the previous chapter, stocks that are more ambiguous tend to be non-dividend paying growth stocks. These stocks are relatively cheap, but with the potential of significant capital gains. Therefore, given the low wealth level of low income households, the potential losses from the ambiguous stocks are limited, while the potential gains are much more significant, compared to high income households. Therefore losing such an opportunity will cause much more regret in the low income households, compared to wealthier ones. This assumption is backed up by Kumar (2009)'s result, which will be discussed later. It is also interesting that this is exactly the opposite point of FNE's prediction (FNE predicts that the higher income household should be less ambiguity averse).

On the other hand, if individuals are more resistant to the self-anticipated regret, i.e. more optimistic, they are less ambiguity averse. Since optimism and pessimism are defined as positive and negative outcome expectations, people with a predisposition to expecting things turning out well might perceive an ambiguous situation differently from someone who tends to expect the worst outcome to happen. Alpha-max-min model is a good example of showing the link between optimism and ambiguity aversion. In this model, people use the formula below to adjust their expectations:

$$E' = \delta(\alpha \text{Max}(w) + (1 - \alpha)\text{Min}(w)) + (1 - \delta)E(w) \quad (4.15)$$

Here the Max (w) and Min (w) are the max and min utility of event W respectively, and $E(w)$ is the initial expected utility. The weight of estimations δ indicates the level of ambiguity, when δ is the increase in investors who become more reliant on their estimation than on their initial judgement $E(w)$. The weight of the worst outcome $(1 - \alpha)$ can be seen as their ambiguity aversion level, when $(1 - \alpha)$ is an increase in investors who have less utility from the ambiguous situation. As stated above, if the agent holds a prior belief of the worst case, which equals to $\text{Min}(w)$, they will minimize their regret. Thus the amount of regret which the agent overcomes is $\alpha \text{Max}(w)$, and it is directly related to the optimism level. Therefore it is assumed that those who revealed a higher level of optimism in the SCF should be less ambiguity averse.

In addition, various studies show that life experience could also change optimism and ambiguity aversion levels. For example, Pulford (2009) speculates that optimistic people consider themselves to be lucky and such feelings have come from their experiences. His experiments proved that a recently ‘lucky’ experience can make individuals seek more ambiguity than usual, but this effect does not last for long. On the other hand, Malmendier and Nagel (2009) argue that past experience in the market has a significant impact on an individual’s future investment decision as well as their risk attitude in the long-term. Hence, there could be a similar impact for the ambiguity attitude: those who have had a negative life-time experience in the stock market are more pessimistic about equity and become more ambiguity averse. And the second assumption is that those who have experienced a negative life-time stock market return are more ambiguity averse.

4.2.3 Empirical factors for ambiguity aversion

The previous section introduced some the cause of ambiguity aversion, and listed several assumptions. This section discusses a method which is used to test these assumptions of heterogeneity in ambiguity aversion empirically.

a. Level of optimism

The first hypothesis is that the level of optimism affects the ambiguity aversion level, as the alpha-max-min model gives an intuitive explanation. Rigotti, Ryan, and Vaithianathan (2004) developed a model according to which optimists are more likely to accept ambiguous returns. In asset pricing, optimism among a small number of investors may cause stock prices to exceed fundamental values in the presence of short-sale constraints (Miller, 1977). Puri and Robinson (2005) also argue that optimism is an important factor in investors’ economic decisions. To measure the level of optimism, this chapter relies on the studies of Puri and Robinson (2005), which used the SCF data and measured the level of optimism by the household’s self-reported life expectation and opinions towards future economic situations and income status. Their studies show that optimists are more likely to believe that future economic conditions will improve. In the SCF the households were asked about what they thought about the US economy

over the next five years in comparison to the current situation, and the optional answers were: better, the same, or worse. If the household believed the economic situation would improve, they were identified as being more optimistic than the average. Therefore, it is further assumed that those who had more optimistic answers in the SCF not only invest more in stocks, but are also less sensitive to ambiguity changes.

b. Income, race, age and education

As stated above, it has been assumed that income, age and education can also impact the ambiguity aversion level, because these parameters are related to the capacity of dealing with the negative feedback of others, as well as the level of regret from self-evaluation. Moreover, such assumptions are backed up by empirical studies based on investment behaviour in lottery-like stocks. Walker (1992) and Kumara (2009) argue that common biological, psychological, religious and socio-economic factors jointly determine the psychological portrait of a gambler or a lottery-lover. It is credible that individuals with a preference for lotteries might adopt a gambling mindset in other domains of their lives. In particular, people's investment choices may be affected by their attitudes toward lottery-playing and gambling. Kumar (2009) states that gambling, lottery-playing and speculation in the market are related albeit distinct activities. He defines speculation as an activity where a risk-seeking (or a less risk averse) individual takes relatively larger bets with large risks and gets rewarded appropriately. Under this definition, speculative stocks are those that have a higher variance and higher expected returns. In contrast, gambling refers to an activity where an individual takes large risks but the reward does not correspond to the level of the risk being taken, for example when the expected return is negative. Gamblers still undertake such bets because they gain a utility by simply being involved in a risky situation. In other words, gamblers are likely to trade with lower returns for the utility they derive from the thrill of gambling. Lastly, lotteries are distinct from speculative and gambling activities. When an individual buys a lottery ticket, he spends a small amount of money and expects to earn a low negative return with a high probability and a large positive return with a very small probability. Lottery players accept a negative expected return for the possibility of a large positive pay-off. Under this characterization of lotteries, stocks with lottery-type features would have lower prices, higher volatility, and large positive skewness. It is important to emphasise that these features of lottery-like stocks are very similar to the

‘aggressive growth stocks’ in the previous chapter, where stock volatility is extremely high as well as its potential returns. The ‘aggressive growth stocks’ are recognized as a more ambiguous asset class in the previous discussion, and thus it is reasonable to assume that lottery-like stock investors may be less ambiguity averse .

Kumar (2009) found that such lottery loving behaviour is correlated with status such as race, age, religion, income, and educational level. Those who are middle-aged, or live in urban, republican dominated regions and belong to minority groups (African-American and Hispanic) and religious groups (Catholic) invest more in lottery type stocks. Additionally, investors who exhibit a stronger preference for lottery-type stocks experience a greater mean under-performance. Meanwhile, Clotfelter and Cook (1989) argue that poorer individuals are more likely to purchase lottery or lottery-like assets, and such a demand for lottery increases with the decrement of their income, according to Blalock, Just, and Simon (2004). This result is also confirmed by Kumar (2009), who found that people with a lower income level (annual income<\$15000 in 1993) have a significantly higher demand for lottery-type stocks, given that other factors are controlled. These results are consistent with the predictions from the self-evaluation theory in the previous section, that lower income households have more regrets about missing the opportunity of lottery and ambiguous stocks.

Therefore, it is also reasonable to assume that people with different income levels may have different ambiguity aversion levels, and those who have lower income should be less ambiguity averse due to their relative preference for lottery-like stocks, and vice versa for the high income investors. Following Kumar (2009)’s study, a test needs to be performed on whether people who are older (over 65 years old), have a lower income (annual income<\$15000 in 1993) and are uneducated (without a college degree) are less ambiguity averse. It is also noteworthy that FNE and the self-evaluation theory make opposite predictions on the relationship between income and ambiguity aversion.

c. Social capital

As stated above, it has been assumed that individuals with a higher level of social capital are less ambiguity averse because they might be less influenced by negative evaluations from others. To measure such social capital empirically, the method of social capital is applied. Putnam (1996), who popularized the concept of social capital, defines it as ‘networks, norms, and trust that enable participants to act together more effectively to pursue shared objectives’. The World Bank is more expansive and it suggests that ‘social capital refers to the institutions, relationships, and norms that shape the quality and quantity of a society’s social interactions’. In general, a higher social capital level means more social interaction between individuals, or more developed social networks. Since social capital is such wide concept, the empirical measures can be defined by various approaches. According to the British National Office of Statistics, one way to measure social capital is through participation, social engagement, and commitment: involvement in local groups, voluntary organizations, clubs, taking action regarding a local issue. Previous studies show that such empirical measures of social capital make a significant impact on investor’s financial decisions. Guiso, Sapienza and Zingales (2004) use the frequency of interactions with communities, such as attending church and blood donation centres as indicators for the social capital level. They found that in Italian regions, where people attend church or donate blood more frequently, the stock market participation rate is significantly higher. Similarly, Hong, Kubik, and Stein (2001) use the interactions with neighbours as another indicator for social capital. They researched on how closely people are connected within the local community and they found that the social capital level significantly affects stock market participation. The SCF survey contains data pertaining to whether the household is involved in voluntary charity work, and the assumption is made here that people who participate in voluntary charity activities have a higher level of social capital, and, therefore, their ambiguity aversion is lower.

d. Life-time experienced return

Malmendier and Nagel (2009) argue that past experience in the market can have a significant impact on an individual's future investment decision, as well as their risk attitude in the long-term. They found that individuals who experienced periods of low stock-market returns and who expressed a lower willingness to take a financial risk are less likely to participate in the stock market, and invest less in stocks. Malmendier and Nagel (2009)'s study shows that the relatively low stock-market participation of young households in the early 1980s can be explained by the disappointing stock-market returns in the 1970s; and the relatively high participation of young investors in the late 1990s is the result of the boom years in the mid 1990s. This result is consistent with the self-evaluation theory' prediction, that a pessimistic level will increase ambiguity aversion. Therefore, based on these studies, this chapter will also test the hotheouses as to whether households which experienced a negative life-time market return are more ambiguity averse.

e. Marital status

As stated above, the research shows that when people are members of groups the influence of potential negative evaluations from others is higher: Blinder and Morgan (2005) found that groups make different decisions rather than individuals. The decisions made by groups are the most explainable solutions but are not always the optimal ones. Keller and Sarin (2007) state when people are put into a group of two, their decisions reveal more ambiguity aversion in comparison to the average person who is not connected. Consequently, the assumption is made that for those households which are reported as married in SCF, the ambiguity aversion level is higher.

f. Smoking status

Apart from the predictions from the FNE and self-evaluation theory, studies based on experiments have also suggested some factors that could possibly affect ambiguity aversion, for example, smoking. Previous studies suggested that smokers have a lower level of risk aversion compared to non-smokers, which indicates a possible difference in

their physiological portraits: Sutter, Rützler and Trautmann (2010) found that impatient people tend to consume more alcohol and tobacco, demonstrate less saving behaviour and have a lower ambiguity aversion level. Furthermore, Viscusi, Magat and Huber (1999)'s findings reveal that students who are more ambiguity averse smoke less in daily life. Therefore, it also needs to be tested whether households in which there are smokers are less sensitive to changes in ambiguity.

4.2.4 Factors for individual household participation

From previous studies, income and wealth level is regarded as one of the most important factors affecting individual stock market participation. It is known that high income groups are more likely to have financial assets and, therefore, they are more active participants in the stock market (Vissing-Jorgensen (2000), Kumar (2009), Hong, Kubik and Stein (2001)). The explanation is that for entering the stock market there is always fixed costs, such as the account fee, broker's charge, and time invested in seeking the market information. A high income group has a significant advantage in handling such fixed costs as this only takes a small percentage of their wealth. In addition, because this group has a relatively larger investment scale, such fixed costs make less impact on their stock's net returns. Also, normally, there is a required minimum amount of investment in the stock account, which represents another entry barrier for low-income individuals.

Secondly, stock-market participation is found to increase with the level of education in the household. This result is confirmed by Kumar (2009), Hong, Kubik and Stein (2001). One natural interpretation of this fact is that education reduces the fixed information costs of participation. It is easier for educated investors to understand the market's risk-reward tradeoffs, or to deal with the mechanics of setting up an account and executing trades. Thus, they spend less time and effort dealing with information costs. Hsu (2012) also argues that households with higher human capital have higher wealth level expectations, and will be more interested in financial assets such as stocks in order to achieve their desired wealth level.

Another factor that can affect individual market participation on an individual level is the psychological anchoring to the historical price. George and Hwang (2004) state that traders might use the 52-week high as an anchor when assessing stock value

information. George and Hwang (2004) argue that if a stock's price is at or near its 52-week high, then this is a stock that has recently had good news, and this can be the time when traders' under-reaction to good news is at its peak. Therefore, closeness to the 52-week high is positively associated with expected returns in the cross-section. Peng and Xiong (2006) show that limited investor attention can result in category-learning behaviour: investors tend to process more market-wide information rather than firm-specific information. Due to the fact that the Dow index is one of the most widely available market information tools, investors are likely to use the Dow index as a benchmark when evaluating new market-wide information. As a result, the nearness of current price to the Dow 52-week high captures the extent of under-reaction, and it should forecast aggregate market returns. Li and Yu (2010) confirmed this result by showing that to evaluate information people use the Dow 52-week average and historical high/low price rather than the current price. Li and Yu (2010) found that when the recent 52 week average is close to the historical level, the momentum is much stronger and, thus, brings about more herding behaviour with corresponding long/short positions. Therefore, the market participation increases when prices are closer to the anchoring point. However, in general, households cannot hold a short position in the stock market and hence, in this research, only the high price anchoring points are considered

Finally, there is a documented link between race and participation. According to Kumar (2009), white, non-Hispanic households have much higher stock market participation rates, controlling wealth and education. In addition, Kumar (2009) also finds that age, marital and retirement status can also affect the demand for financial assets. This can be explained by the varying demand of income at different stages a person's life cycle. It is known that demand for saving and financial assets increases in later stages of the life cycle, because investors need income streams for their children and retirement. Kumar (2009)'s results are consistent with this theory.

4.2.5 Research questions

To sum up, Easley and O'Hara (2010) provide a theoretical structure to how heterogeneity in ambiguity aversion can affect the stock market. Ellsberg (1961) and Krahermer and Stone (2010) further argue that self-evaluation is the reason for ambiguity aversion, while Curley, Yates, and Abrams (1986) believe that the reason is the fear of negative evaluation by others (FNE).

Meanwhile, Dimmock, Kouwenberg, Mitchell and Peijnenburg (2013) observed strong heterogeneity in ambiguity aversion, as they found both ambiguity averse and ambiguity seeking decisions. However, their results are survey-based and do not use real market data. Therefore, it is worth investigating whether this model is held in the real market, and what the driving reason for the behaviour is. **Thus, the research questions are:**

Does ambiguity affect the stock market participation at household level? If so, do different investors have a different level of ambiguity aversion? What are the reasons for it?

4.3 Data and methodology

To answer the research questions listed above, empirical evidence is required. The following section introduces the data and methodology used for the testing.

4.3.1 Survey of consumer finance

In this chapter the main data source is the Survey of Consumer Finance (SCF). SCF can be obtained from the Board of Governors of the Federal Reserve. SCFs were conducted annually from 1946 to 1971, and were again administered in 1977, 1983, and then every three years after 1983. To produce the SCF, interviewing is conducted between January and August to gather the demographic and financial information, such as consumer income, assets, debt, and major transactions, from a large number of different US households. The samples of SCF are made up of a national cross-section of dwelling units to represent the total population of the US. In this sample the data from 8 different surveys ending in 2010 is included, with a total number of 176,160

available observations. It is noteworthy that each SCF survey picks its respondents independently from the previous survey and, therefore, provides randomly-sampled cross-section data of US households rather than time-series data. As a result, the data of SCF is suitable for the probability-based regression model, such as the logistic regression model.

Although this survey has been used extensively in academic research (i.e., Malmendier and Nagel, 2011), it has a well-known technical issue for its multiple-imputation. Multiple-imputation is a technique commonly used to deal with missing information on individual items in survey data. This technique employs multivariate statistical methods to impute missing data resulting in multiple complete data sets. Since 1989, the Survey of Consumer Finances (SCF) data files have contained five complete data sets, referred to as implicates. The benefit to researchers of these multiply imputed data files is that they contain no missing values; however, an appropriate method of analysing multiply imputed data is also required. One major way of doing this is known as repeated imputation inferences (RII) from Rubin (1987), which combines the results obtained independently in each of the separate implicates using multiple imputation combining rules. In this research the RII approach is followed and the regression results of all five imputations of SCF data are combined to generate the final standard error and estimations of coefficients.

4.3.2 Household's participation variables from SCF

In the SCF, households were asked whether they have any publicly traded stocks and whether they have any stock-based mutual fund assets. The variable of the stock market participation $SMP_{i,t}$ is 0 if the household i has a negative answer to both questions at time t , and 1 otherwise. The gross income variable $income_{i,t}$ is the logged reported total gross income of the household i in the last year of time t . For example, when the survey was completed in 1989, the reported income variable would be the logged gross income received in 1988, measured by nominal dollars. Marriage variable $marriage_{i,t}$ is set to 1 if the household i is married at time t and 0 otherwise, and retirement $retirement_{i,t}$ is set to 1 if the respondent (or couple) is (are both) retired. The race variable $race_{i,t}$ is set to 1 if the household is white and 0 for the rest. The risk

attitude variable is based on the following questions in the survey (the following example is from the 2010 survey):

Which of the following statements come the closest to describe the amount of financial risk that you are willing to take when you save or make investments?

1. *Not willing to take any financial risks*
2. *Take average financial risks expecting to earn average returns*
3. *Take substantial financial risks expecting to earn substantial returns*
4. *Take above average financial risks expecting to earn above average returns*

Thus a higher score in the risk attitude means more risk tolerance and and, therefore, less risk aversion. To present such a risk attitude, the dummy variable is set according to the answers: variable $risk_{i,t}$ is set to i if household i 's answer is number i at time t , and 0 otherwise. Age variable $age_{i,t}$ is the reported age of respondent i at time t , and the education variable $edu_{i,t}$ is set to 1 if respondent i at time t has a college degree, and 0 otherwise. The social capital level $Soc_{i,t}$ is the dummy that is set to 1 if members of the family have ever worked as volunteers for charity organisations and 0 otherwise. In addition, the market price anchoring variable Ach_t is a control variable for anchoring, which is equal to the average Dow index price for the last year of the survey/the historical high price of the Dow index till the last year of the survey. Since the survey is carried out every three years, the ambiguity $amb_{t-1,t-3}$ is measured by the average ambiguity between year $t-1$ and $t-3$ (again amplified by 100 times). The appendix of this chapter provides an example of the questions that were used to construct these variables in the SCF 2010.

4.3.3 Ambiguity aversion variables

To examine whether the empirical factors stated above can affect the ambiguity aversion, this chapter tests the corresponding variables including age, education, income level, social capital level, optimism level, marriage, life-time experienced stock market return and smoking status. Based on Kumar (2009)'s study the base line of the high and low income groups is drawn: the high income group comprises households that have higher than \$125000 income per year in 1993, and the low income level households that

have below \$15000 in 1993. This is adjusted by the CPI for every year and converted into the corresponding income level in each survey. If the reported household income exceeded the high income level for the year, the high income dummy $high\ ic_{i,t}$ is set to 1, and, if otherwise, 0. And vice versa for the low income dummy $low\ ic_{i,t}$. For age, dummies of younger (below 40), middle aged (40-65) and older (above 65) are set according to Kumar (2009)'s study. The education dummy is set to 1 when a household claimed to have a college degree and 0 for the rest.

In addition, it is also necessary to test whether there is a significant impact of the life-time experienced market return on ambiguity aversion. Following Malmendier and Nagel (2009)'s study the weighted life-time experienced stock market return is constructed as a controlling variable. The life-time experienced stock market return is a weighted average of S&P returns within an individual's lifetime. According to Malmendier and Nagel (2009), individuals may not have sufficient experience towards the market from their early childhood, therefore in this study the experienced return is measured from age 10 to the survey time. For household i at year t , the life-time experienced return is:

$$A_{(i,t)}(\lambda) = \sum_{k=1}^{age_{i,t}-11} w_{i,t}(k, \lambda) R_{t-k} \quad (4.16)$$

$$w_{i,t}(k, \lambda) = \frac{(age_{i,t} - k)^\lambda}{\sum_{k=1}^{age_{i,t}-11} (age_{i,t} - k)^\lambda} \quad (4.17)$$

where R_{t-k} is the market return of S&P 500 at year $t-k$.

As Malmendier and Nagel (2009) suggest, the memory fades with time and returns in the earlier years have less weight, thus the weight parameter is set to 1.25 to give a reduced weight to the earlier experienced return. If such an experienced life-time market return is negative, then the dummy for the negative life-time return is set to 1, otherwise it is 0.

For the social capital dummy, the method from the British Office for National Statistics is used, which measures social capital through the degree of involvement in voluntary charity work. In the SCF, households report if they volunteered to work in charity organizations during the past year of the survey, and the social capital dummy is

set to 1 for those who participated, and 0 if otherwise. Similarly, the smoker dummy is set to 1 if the household claimed to be smokers, and 0 if otherwise.

4.4 Regression results and analysis

4.4.1 Testing ambiguity and individual stock market participation

Firstly, the test is performed on whether ambiguity makes a significant impact on the individual level of market participation. The model includes individual-specific control variables: $income_{i,t}$, which is the natural logarithm of income deflated into 2010 dollars; $risk\ attitude_{i,t}$ (three dummies to present whether the individual is willing to take average, above average or substantial financial risks), $race_{i,t}$ (dummy equal to 1 if the household is white and 0 if otherwise), $marriage_{i,t}$ (dummy equal to 1 if the individual is married), $retired_{i,t}$ (dummy equal to 1 if the individual is retired), $age_{i,t}$ (in years) and $education_{i,t}$ (dummy equal to 1 if the individual has a college degree). The probability that the household owns equities is modelled conditionally on the predictors in the vector X outlined above, $p = \Pr(Stock = 1|X)$, and the linear logistic model is in the form:

$$\log\left(\frac{p}{1-p}\right) = a + b'X \quad (4.18)$$

where a is the intercept parameter and b is a vector with slope parameters estimated with maximum likelihood. The main regression is, thus, set as follows:

$$\begin{aligned} [P(SMP_{i,t}) = 1] & \\ & = a_0 + a_1amb_{t-1,t-3} + a_2vol_{t-1,t-3} + a_3ach_{i,t} + a_4risk_2 + a_5risk_3 \\ & + a_6risk_4 + a_7income_{i,t} + a_8race_{i,t} + a_9marriage_{i,t} \\ & + a_{10}retirement_{i,t} + a_{11}age_{i,t} + a_{12}edu_{i,t} + \varepsilon_{i,t} \end{aligned} \quad (4.19)$$

Since the SCF uses the multiple imputation technique to impute missing values, standard errors for multiple imputations (RII) are adjusted following the method of Rubin (1987) in this study. Table 4.1 shows the regression result for the main regression, and it can be seen that the ambiguity has negative and significant estimates. This result

confirms the hypothesis that on a household level market ambiguity can reduce participation. In addition, in this model it is estimated that a one standard deviation increase of ambiguity reduces the probability of participation by 4.66%. If the market risk (measured by realised volatility) increases, it makes a significant impact as the participation rate decreases. In addition, since the variables have positive and significant estimations, households with higher income and education levels have more chance to participate in the stock market. Furthermore the result shows that white and retired people have a higher chance to become stockholders as well as married couples. The chance of participation also increases with age and risk tolerance. Moreover, the price anchoring level has positive and significant estimates, which are consistent with previous studies. To test the robustness of the main result, a similar process to Kumar (2009) is used, which drops households with reported incomes of less than 1000 dollars, and households with a head who is less (more) than 24 (75) years old. Table 4.2 presents the result and it can be seen that the main result of this section is robust.

4.4.2 Testing heterogeneity in ambiguity aversion

In this section the second research question is examined in order to test the empirical factors that could affect ambiguity aversion. The RII is also applied here to adjust the results of this section to net out the influence of multiple-imputations in SCF.

a. Income and ambiguity aversion

As Kumar (2009) states income status affects people's investment choice. In addition, FNE theory predicts that households with a higher level of income are less ambiguity averse, while the self-evaluation theory makes the opposite prediction. Following Kumar (2009), in this study the income is controlled by high and low income lines. The high income line is equal to \$125,000 in 1993 and the low income line equal to \$15000 in 1993. The baselines of incomes are adjusted by CPI and, therefore, the corresponding baseline is produced each year. The regression is:

$$\begin{aligned}
[P(SMP_{i,t}) = 1] & \\
&= a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
&+ a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
&+ a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} + a_{13} high\ ic_{i,t} \\
&+ a_{14} low\ ic_{i,t} + a_{15} amb_{t-1,t-3} * high\ ic_{i,t} + a_{16} amb_{t-1,t-3} \\
&* low\ ic_{i,t} + \varepsilon_{i,t}
\end{aligned}
\tag{4.20}$$

Table 4.3 shows the regression results. It can be seen that the ambiguity and the risk are still negative and significant (estimated coefficient= -228.27, $p < 0.01$, which means a one standard deviation increase of ambiguity reduces the probability of participation by 4.16%), and the signs and the significance level for most of the previous variables stay the same. On the other hand, interactions with the income level dummy produce significant results, as high income groups have significant and negative estimates and low income groups have significant and positive estimates. These results are consistent with the previous findings that low income groups show lottery-loving behaviour and, thus are less ambiguity averse. In addition, it also fits the predictions from the self-evaluation theory.

b. Age and ambiguity aversion

Since older people are mature and are supposed to have an advanced self-cognition level, it is assumed in this study that older people are more resistant to negative feedback from others and, thus, are less ambiguity averse. This result is supported by Chen, Kim, Nofsinger, and Rui (2003), who find that older investors are more confident with their decisions. Therefore, following Kumar (2009), in this chapter the households are divided by the age of the main respondents in the regression. The older group are the respondents with a reported age of above 65 years old, and the regression is presented in the formula below:

$$\begin{aligned} [P(SMP_{i,t}) = 1] \\ = a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\ + a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\ + a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} \\ + a_{13} old_{i,t} + a_{14} amb_{t-1,t-3} * old_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{4.21}$$

The result is shown in Table 4.4, and it can be seen that while the main result is still held (estimated coefficient= -287.56, $p < 0.01$, which means a one standard deviation increase of ambiguity reduces the probability of participation by 5.24%), the older group has significantly less ambiguity aversion, since the interaction action of its dummy and the ambiguity has positive and significant estimates. Therefore, this result supports the hypothesis that older people are less ambiguity averse.

c. Education and ambiguity aversion

Previous studies discussed above indicate that education makes an important impact on the investors' risk aversion level, even with income and other variables being controlled. According to the FNE theory it is assumed that those individuals with a higher education are affected by negative evaluation less.

On the other hand, as Kumar (2009) suggest, individuals with lower education tend to be more interested in the lottery-type of stocks, which is more ambiguous as has been previously explained in this thesis. Therefore, it is also the research interest as to

whether education level has a direct effect on ambiguity aversion in the stock market, and the regression is presented below:

$$\begin{aligned}
[P(SMP_{i,t}) = 1] \\
&= a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
&+ a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
&+ a_{10} retirement_{i,t} + a_{11} age_{i,t} \\
&+ a_{12} edu_{i,t} + a_{13} amb_{t-1,t-3} * edu_{i,t} + \varepsilon_{i,t}
\end{aligned}
\tag{4.22}$$

The regression result is shown in Table 4.5. It does not show any significant outcome between the education and the ambiguity. Therefore, it cannot be concluded whether education level can affect ambiguity aversion.

[Table 4.5 here]

d. Marriage and ambiguity aversion

The FNE theory also states that people can be more ambiguity averse when they act as a group. Keller and Sarin (2007) found that, if people are put into a group of two, their decisions reveal more ambiguity aversion compared to the average of the two people's vaule. Thus, here the argument from the FNE theory is that married households can be more ambiguity averse. The test regression is presented below:

$$\begin{aligned}
[P(SMP_{i,t}) = 1] \\
&= a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
&+ a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
&+ a_{10} retirement_{i,t} + a_{11} age_{i,t} \\
&+ a_{12} edu_{i,t} + a_{13} amb_{t-1,t-3} * marriage_{i,t} + \varepsilon_{i,t}
\end{aligned}
\tag{4.23}$$

Table 4.6 shows the result of the regression, and it can be seen that the interaction term of the ambiguity and marriage status has a significant and negative estimation

(estimated coefficient= -164.07, $p < 0.01$, which means a one standard deviation increase of ambiguity reduces the probability of participation by 3%). Therefore this evidence supports the assumption that married couples are more ambiguity averse.

[Table 4.6 here]

e. Social capital and ambiguity aversion

As stated above, the research based on the FNE theory (Bochner, 1965; Foster and Rosenzweig, 1995) states that more ambiguity averse individuals can be less sociable, and in this study it is assumed that those who have higher social capital are less ambiguity averse. To test this hypothesis the dummy variable for the high social capital level is added into the regression. According to the standard of British National office of statistics, the social capital level can be defined through many different approaches including the frequency of interaction with local communities, the intensity of working for a charity or a voluntary work experience, the level of trusting people and organisations. In the SCF a household is asked whether they have participated in any charity or voluntary work in the last year, and this is used as a dummy proxy for the social capital. The dummy for high social capital $sc_{i,t}$ is set to 1 if the household i had a charity or voluntary work experience at time t , and 0 for the rest. The regression is presented below:

$$\begin{aligned}
 [P(SMP_{i,t}) = 1] \\
 = a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
 + a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
 + a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} \\
 + a_{13} sc_{i,t} + a_{14} amb_{t-1,t-3} * sc_{i,t} + \varepsilon_{i,t}
 \end{aligned}
 \tag{4.24}$$

The result is shown in Table 4.7. However, since the regression for the social capital is not significant, it cannot be confirmed that social capital can significantly affect the household's ambiguity aversion.

[Table 4.7 here]

f. Optimism and ambiguity aversion

The self-evaluation theory indicates that optimism directly influences the behaviour of individuals under ambiguous conditions. The Alpha-Max-Min model states that optimistic individuals put more weight on the best possible outcome and, thus, have higher expectations of the ambiguous event. It is assumed, therefore, that for households which reveal a more optimistic attitude, the stock market participation is less sensitive to the changes on the ambiguity level. Since 1989, in the SCF households have been asked for their opinions about the economic situation in the US over the following five years, and how their family income will change in the next five years. Following Puri and Robinson (2005) these are both indicators of the household optimism level. Therefore in the section the dummy for optimistic $op_{i,t}$ is set to 1 if the household believes both the future economy US is going to improve, and their family income growth will beat inflation growth; and 0 if otherwise. The regression is presented below:

$$\begin{aligned}
 [P(SMP_{i,t}) = 1] \\
 &= a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
 &+ a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
 &+ a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} \\
 &+ a_{13} op_{i,t} + a_{14} amb_{t-1,t-3} * op_{i,t} + \varepsilon_{i,t}
 \end{aligned}
 \tag{4.25}$$

[Table 4.8 here]

The result is shown in Table 4.8. It can be seen that ambiguity is significantly and negatively correlated to the market participation (estimated coefficient= -164.07, $p < 0.01$, which means a one standard deviation increase of ambiguity reduces the probability of participation by 4.8%), therefore the main result remains. On the other hand, the interaction variable between the market ambiguity and the optimism level is positive and has p -value of 0.07. This result is consistent with the previous assumption that people with a higher optimism level are less affected by ambiguity.

g. Experienced life time return and ambiguity aversion

The empirical research states that people's level of optimism is related to their past experience: those who have recently had a winning experience consider themselves as lucky and become less ambiguity averse. Meanwhile, a life-experience with the asset market can significantly affect people's investment decision making: people who experienced the great depression would have become less likely to participate in the stock market and become more risk averse (Malmendier and Nagel, 2009). Hence the hypothesis is made that those who experienced the negative life-time stock market return are more pessimistic towards stock market and their ambiguity aversion should be higher. To test this hypothesis, the life-time experienced stock market return variable is constructed based on Malmendier and Nagel (2009), which is equal to the weighted life-time stock market return from the age of 10. The negative experience dummy $negative_{i,t}$ is set to 1 if household i had a negative life-time experienced return at time t , and 0 if otherwise. The regression is presented below:

$$\begin{aligned} [P(SMP_{i,t}) = 1] \\ = a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\ + a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\ + a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} \\ + a_{13} negative_{i,t} + a_{14} amb_{t-1,t-3} * negative_{i,t} + \varepsilon_{i,t} \end{aligned} \tag{4.26}$$

Table 4.9 shows the result of the regression and it can be seen that the negative life-time experience has a negative and significant estimation. Thus, this evidence supports the assumption that the life-time experience can impact the ambiguity aversion.

[Table 4.9 here]

h. Smoking and ambiguity aversion

Apart from the FNE theory and the self-evaluation theory, there is empirical study (Sutter, Rutzler and Trautmann, 2010) which states that impatient people tend to be less ambiguity averse and are much more likely to be smokers. Accordingly, there is research interest in this study as to whether smoking status can affect a household's

ambiguity aversion level in the stock market. In the SCF a household is reported on whether the occupants are smokers, and the smoker dummy $Smoker_{i,t}$ is set to 1 if there are reported smokers and 0 if otherwise. The regression is presented below:

$$\begin{aligned}
[P(SMP_{i,t}) = 1] \\
&= a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
&+ a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
&+ a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} \\
&+ a_{13} smoker_{i,t} + a_{14} amb_{t-1,t-3} * smoker_{i,t} + \varepsilon_{i,t}
\end{aligned}
\tag{4.27}$$

The result is shown in the Table 4.10. Unfortunately the estimations of the ambiguity and the smoking status are not significant, therefore it cannot be concluded that there is a significant relationship between smoking status and ambiguity aversion.

[Table 4.10 here]

4.4.3 Robustness

a. Joint test

So far all of the factors have been tested individually, for robustness this section tests them all together. The regression is presented below:

$$\begin{aligned}
[P(SMP_{i,t}) = 1] \\
&= a_0 + a_1 amb_{t-1,t-3} + a_2 vol_{t-1,t-3} + a_3 ach_{i,t} + a_4 risk_2 + a_5 risk_3 \\
&+ a_6 risk_4 + a_7 income_{i,t} + a_8 race_{i,t} + a_9 marriage_{i,t} \\
&+ a_{10} retirement_{i,t} + a_{11} age_{i,t} + a_{12} edu_{i,t} \\
&+ a_{13} old_{i,t} + a_{14} op_{i,t} + a_{15} sc_{i,t} + a_{16} smoker_{i,t} + a_{17} negative_{i,t} \\
&+ a_{18} highic_{i,t} + a_{19} amb_{t-1,t-3} * old_{i,t} + a_{20} amb_{t-1,t-3} * op_{i,t} \\
&+ a_{21} amb_{t-1,t-3} * ed_{i,t} + a_{22} amb_{t-1,t-3} * sc_{i,t} + a_{23} amb_{t-1,t-3} \\
&* smoker_{i,t} + a_{24} amb_{t-1,t-3} * marriage_{i,t} + a_{25} amb_{t-1,t-3} * edu_{i,t} \\
&+ a_{26} amb_{t-1,t-3} * highic_{i,t} + \varepsilon_{i,t}
\end{aligned}
\tag{4.28}$$

The result is presented in the Table 4.11. When pooled together, the result is that some of the variables become insignificant. However, the ambiguity itself is still significant with negative coefficients (estimated coefficient= -267.53, $p < 0.01$, which means a one standard deviation increase of ambiguity reduces the probability of participation by 4.9%), and the interactions of the ambiguity with the age and the higher income status are still significant. Therefore, the joint test provides empirical evidence that while older people are less ambiguity averse, people with a higher income are more ambiguity averse, which supports the FNE and self-evaluation theory respectively.

[Table 4.11 here]

b. Robustness

This section applies a similar robust-checking process to Kumar (2009), which removed the households with a reported income of less than 1000 dollars, and the households with a head that is less (more) than 24 (75) years old. The reason for removing these households is that they might lack the ability to make viable financial decisions. All of the results from Tables 4.2-4.8, and 4.10 are recalculated with the new data, and the results are reported in Table 4.12. However, the result for the life-time experienced returns cannot be processed in such a way, due to very the strong correlation between age and life-time experienced returns: because of the bull market in the 1980s and 1990s the average age of people who have experienced positive life-time returns is 32. Therefore, by removing all of the samples below 24 years old, it significantly reduces the number of households that have experienced positive life-time returns. As a result, for life-time returns the robustness check removes households with a reported income of less than 1000 dollars and with a head more than 75 years old.

[Table 4.12 here]

For brevity, Table 4.12 reports the estimation of the ambiguity variables and the interaction terms only. It can be seen that the ambiguity remains negative and significant in all of the regressions; therefore the conclusion is still held that ambiguity reduces the household's stock market participation. The interaction terms of the ambiguity with age, marriage, income and life-time experience remains significant, while social capital, smoking and education status remain insignificant. In addition, the interaction between ambiguity and optimism becomes more significant (p -value

changes from 0.07 to 0.05). Therefore all of the results from the previous sections are robust.

4.5 Limitations and future developments

The main limitation of this chapter is the sample size. Since from 1983 the SCF data is only available every three years, there are up to 10 surveys that can be used for this research. In addition, many questions that have been used in order to construct the variables for this chapter have only been available since 1989, and thus further reduce the available number of surveys to 8. This relatively short time period may result in some bias in the conclusion. However, since each available survey still provides a considerable number of observations (176,160 observations in total), the results are still convincing.

Another limitation is that, due to privacy protection, some of the survey results are not available for the public, for example, data on address, religion and political affiliation. Previous studies documented that these variables can affect a household's financial decision, i.e. Catholics and Jews participate in lotteries more than Protestants and Mormons (Grichting, 1986; Clotfelter and Cook. 1989), urban people are more likely to buy lotteries than people located in rural areas (Kumar, 2009). However, since there is no evidence that these variables make an extremely strong impact on a household's stock holding, their influence on the result should be limited as a decent number of control variables are already included in the regression.

The final limitation of this chapter is the cohort effect. Due to the Great Depression in the 1930s and the bull market in the 1990s, a household's life-time experienced stock market return is significantly correlated to its age. In addition, there was a rapid growth in the number of baby boomers in the 1940s and 1950s who, during the sampling period 1989-2010, retired and thus changed their investment goals. All of these facts may have an impact on the results of this chapter.

Regarding future studies, the first suggestion is to test the stated models by using some annual or quarterly household data, or data which has a longer time period. The ambiguity measure used in this chapter starts from the year of 1968 and, therefore, is long enough to collaborate with most of the data sources. Secondly, the SCF also

reports the categories and the weights of the assets that households hold. Thus there is research interest to see of how ambiguity could affect a household's portfolio choice. The results from the previous chapter show the stocks that pay no dividends are affected by ambiguity, and investors transfer their capital from the stock market into more liquid asset classes to avoid a high ambiguity period. Therefore, it might be worth testing whether a household chooses to hold more liquid asset during a high ambiguity period, and whether the interaction of ambiguity and the average market dividend yield has a significant impact on a household's stock holding. Finally, efforts are also needed in order to test the data outside the US. As previously documented, ambiguity aversion has psychological roots that are correlated to people's personality and social status, such as their age, income and level of optimism. Therefore, it is reasonable to assume that in countries which have a very different mentality and social structure the overall reaction to ambiguity might differ from that of the US. For example, China and Japan overall have an older population compared to the US as well as a very different culture. As a result, the Asian market may have a distinctive feature in terms of the household's ambiguity aversion.

4.6 Conclusion

From the Easley and O'Hara (2010) model it can be seen that different levels of ambiguity aversion can affect the individual's market participation, and this thesis has had the aim of testing whether a such model is held in real life. Therefore, first of all the influence of ambiguity on the household level of stock market participation is tested. For the test, the data from the Survey of Consumer Finance (SCF), which is issued by the Board of Governors of the Federal Reserve, is applied. This survey is conducted every three years, and gathers demographic and financial information from a large number of different households. Since 1989 the SCF has applied the multiple-imputations technique to complete the missing data, which results a data with five-imputations in the final set. Accordingly, the repeated imputation inferences (RII) technique from Rubin (1987) is applied, which combines the results obtained independently on each of the separate implicates using multiple imputation combining rules. This provides unbiased standard errors and estimates. The result supports Easley and O'Hara (2010), as ambiguity have negative and significant estimations. Furthermore, the robustness check is applied by deleting data with an income lower

than \$1000 and age less than 24 or more than 75, and the result is still held. Therefore it is suggested that ambiguity aversion can significantly reduce household stock market participation.

The research interest of this chapter also focuses on the factors that can influence the households' ambiguity aversion level. There are two main psychological explanations of ambiguity aversion, which are based on the negative evaluations from others and self-evaluation theories. The FNE theory believes that the fear of negative evaluation from others is the main cause of ambiguity aversion, and the research shows that people become more ambiguity averse in groups, and those who are more ambiguity averse are also less social. Thus it is assumed that people with a higher social status (higher income/education) or a lower social capital are less ambiguity averse. Also married couples should be more ambiguity averse than a single person. On the other hand, since older people are more mature with self-cognition, it is supposed that their ambiguity aversion is lower due to greater self-confidence and the ability of resisting FNE. The self-evaluation theory states the reason for the ambiguity aversion is that people prefer to minimise their regret. It points out that the level of optimism is the key for the ambiguity aversion. Thus, it is assumed that the household which is reported to be more optimistic in the SCF, should be less ambiguity averse. In addition, based on the self-evaluation theory, a low income household will regret more if it misses the reward from ambiguous stocks. Therefore, in contrast to FNE, self-evaluation predicts a high income household with high ambiguity aversion. Furthermore, previous studies show that an optimistic attitude can be related to life experience; therefore the hypothesis is that a negative life-time stock market return leads to a low optimism level and thus increases the ambiguity aversion. Moreover, the research shows that impatient people tend to be less ambiguity averse, and they are more likely to be smokers. As a result the hypothesis is tested whether smokers reveal less ambiguity aversion in the stock market.

The results confirmed the majority of the assumptions made in this chapter, i.e. that heterogeneity in ambiguity aversion exists in the real market. In particular, older people (more than 65 years old) and those who are more optimistic about the future economy are significantly less ambiguity averse, while married, high income (>\$125,000 income in 1993) and those who experienced a negative life-time stock market return are significantly more ambiguity averse. Therefore, both FNE (older people are less ambiguity averse; married couples are more ambiguity averse) and self-

evaluation (optimistic people are less ambiguity averse; high income people are more ambiguity averse) theories have successful predictions. Thus the results show that ambiguity aversion might not be the single result of fear of negative evaluations from others or self, but a combination of both.

Chapter5. Conclusion

5.1 Introduction

The study was set to explore the impact of ambiguity aversion in stock market participation, and its corresponding ambiguity premium empirically. The study has also sought to identify the psychological reasons for ambiguity aversion, and whether certain investor types are more ambiguity averse than others, particularly among US households. The general theoretical literature on this subject predicted that, similar to risk aversion, ambiguity aversion causes a price premium and thus should be another asset pricing factor. However, most of the theories have no empirical support or have only been tested in simulated environment. Therefore, this study has the purpose finding empirical evidence by using real market data, and the research questions are:

1. Does ambiguity aversion reduce stock market participation? And if it does, what is the scale of the impact? Is there any asset class affected by ambiguity aversion more than the others?
2. Does ambiguity affect stock market participation on an individual level? If so, do investors have a different level of ambiguity aversion? What are the reasons for it?

5.2 Empirical findings

The main empirical findings are chapter-specific, and have been summarised within the particular empirical chapters of this thesis. This section synthesises the empirical findings in order to answer the research questions stated above.

1. Does ambiguity aversion affect stock market participation at all? And if it does, what is the scale of the impact? Is there any asset class which is affected by the ambiguity aversion more than others?

a. Ambiguity aversion does reduce stock market participation. By using the American mutual fund flow data as a measure of the stock market participation, and the divergence of analysts' opinions as a measure of stock market ambiguity, it has been found that the increment in the stock market ambiguity is positively and negatively

correlated to the mutual fund flows. Therefore, an increase in stock market ambiguity will result in the outflow of capital from the mutual funds, which means a reduction in the stock market participation level.

b. Apart from the simple outflow, the results of this thesis also show that during a period of high ambiguity, part of the outflows from the stock market, which are caused by ambiguity aversion, are transferred into the money market. Based on the ICI definition, money market funds are invested in low risk, high-grade assets that receive the full principal and interest within 90 days on average, therefore investors can switch to the short-term investments with a higher liquidity in order to avoid stock market uncertainty.

c. Stocks without dividend payments are more sensitive to changes in the ambiguity level. The result shows that in comparison to the rest of the equity funds, the ‘growth’ and ‘aggressive growth’ style funds have greater and more significant correlation coefficients with the ambiguity. According to the ICI definition, ‘aggressive growth’ and ‘growth’ funds invest in non-dividend-paying stocks with a focus on the capital gains, and thus are considered as more ambiguous. Consequently, their flows are more sensitive to the changes in the ambiguity level.

2. Does ambiguity affect the stock market participation at household level? If so, do different investors have a different level of ambiguity aversion? What are the reasons for it?

a. Ambiguity aversion does reduce stock market participation at household level. By using reported stock holding from the Survey of Consumer Finance as a measure of stock market participation, and the convergence of analyst’s opinions as a measure of the ambiguity (Anderson et al., 2009), it has been found that the probability of a household holding equity assets is negatively and significantly correlated to the ambiguity level. Therefore, an increment in the ambiguity level reduces the overall participation rate of US households.

b. There is heterogeneity in the households’ ambiguity aversion level. The results show that the following households have a higher ambiguity aversion level: married households, households with higher income (>125,000 annual incomes in dollars, deflated into the 1993 level), and households which experienced the negative life-time

stock market returns. On the other hand, older (age>65) households and households which reveal more optimistic views towards the future US economy are less ambiguity averse.

c. Ambiguity aversion is the result of a combination of the effect of the fear of negative evaluation from others (FNE) and the fear of negative self-evaluation. FNE theory predicts that investors, who make decisions as a group, are more ambiguity averse, while those who are more mature are less ambiguity averse. These predictions are confirmed by results from this thesis. Also the results prove that some of the predictions from the self-evaluation theory exist in the real market: investors who are more optimistic are less ambiguity averse, and investors who have had a negative past experience with the stock market, or who have a higher income, are more ambiguity averse. Therefore, ambiguity aversion is not driven solely by FNE or self-evaluation, but is the combined result of the two motivations.

5.3 Implications

5.3.1 Theoretical implications

Knight (1921) suggested that when a situation is ambiguous an investor's decision making process is different to when they are in a risky situation. Based on Knight's study, many theoretical works predicted a distinctive price premium for ambiguity aversion, which is independent from the risk premium (Dow and Werlang, 1992; Maenhout, 2004; Cao, Wang and Zhang, 2005; Garlappi et al., 2007; Epstein and Schneider, 2010). Easley and O'Hara (2009) stated further that the price premium for ambiguity can be a result of limited market participation caused by ambiguity aversion, and that the individual ambiguity aversion level can be heterogeneous. Such heterogeneity can be the result of either a different level of fear of negative evaluations from others (FNE) (Curley, Yates, and Abrams. 1986), or a different level of self-evaluation (Ellsberg, 1961; Hamm & Bursztajn, 1979).

This study measures stock market participation in two ways: using US equity mutual funds capital flows from the Investment Company Institute (ICI) data, and using the proportion of the households that invest in equities, estimated in the Survey of Consumer Finances (SCF). The measure of the market-wide ambiguity is based on a

recent study by Anderson et al. (2009) and reflects the dispersion in analysts' implied forecasts about market returns. This measure corresponds closely to the definition of the ambiguity provided by Ellsberg (1961).

The results show that ambiguity is significantly and negatively correlated to equity fund flows. As a result, an increment of the market ambiguity leads to a capital outflow from mutual funds, which means a reduced level of market participation. Therefore, this thesis empirically supports the theory that predicts limited market participation due to ambiguity aversion (Easley and O'Hara, 2009). The thesis also provides strong empirical evidence for the existence of the ambiguity premium and, therefore, supports all of the corresponding literature (Knight, 1921; Ellsberg, 1961; Dow and Werlang, 1992; Maenhout, 2004; Cao, Wang and Zhang, 2005). In addition, the results show that for funds investing in non-dividend paying assets, their flows are more significantly affected by ambiguity. This is consistent with the results of Zhang (2006) who argues that non-dividend paying stocks have low information quality, or, in other words, high ambiguity and, therefore, carry the ambiguity premium. Moreover, it is found here that, during periods of high ambiguity, capital flows from equity funds into the money market funds, suggesting that investors flee to safer and more liquid asset classes when faced with a high ambiguity about stock returns. This is supporting evidence for Ozsoylev and Werner (2009)'s and Eichberger and Spanjers (2007)'s studies, which argue that illiquidity could be associated with high ambiguity and causes the price premium.

The results from the SCF analysis data are corroborating. By using the data from multiple surveys, a logistic model has been estimated, where the dependent variable is a binary indicator of whether the individual owns equities. Controlling for various household and market level factors, it has been found that higher ambiguity leads to a reduction in the probability that the average household owns equities. This provides reinforcement of the findings from the mutual fund flow data, and further supports the existence of the ambiguity premium. In addition, it has also been found that there is heterogeneity in households' ambiguity aversion. The results show that, as the FNE theory predicts (Curley, Yates, and Abrams, 1986; Ellsberg, 1961; Knight, 1921), married households are more ambiguity averse, while older (age>65) households are less ambiguity averse. Meanwhile, although the self-evaluation theory (Ellsberg, 1961; Hamm & Bursztajn, 1979) was not significant in Curley, Yates, and Abrams (1986)'s test, some of its predictions are significant here: households who have experienced

negative life-time stock market returns are more ambiguity averse, and households who reveal a more optimistic view towards the future are less ambiguity averse. In addition, as the self-evaluation theory suggests, households with a higher income have higher ambiguity aversion. Therefore, the result not only supports the Easley and O'Hara (2009)'s model of heterogeneity in ambiguity aversion, but also provides empirical evidence for both the FNE and the self-evaluation theory.

5.3.2 Policy and practical implications

In addition, this research also provides several policy implications concerning ambiguity, in order to improve the stock market's performance and the participation rate. Higher stock market participation not only helps stabilise asset prices, but also helps to improve risk diversification which will benefit the whole economy. Therefore, firstly, improving information disclosure should be an integral theme of policy. Adequate resources of information, and better communication between companies and investors, can reduce ambiguity and therefore increase overall stock market participation. Secondly, governments should consider policies that can reduce investors' losses in extreme events. Ambiguity averse investors' decisions are highly affected by the worst-case results. If regulations can provide back up for losses in extreme events, the ambiguity premium could be reduced and, therefore, increase market participation and asset prices.

The results of this study also suggest that mutual funds' managers could improve their marketing efficiency by targeting the correct ambiguity averse group. The results of Chapter 3 show that the mutual funds which invest in non-dividend paying stocks ('aggressive growth' and 'growth') are more affected by ambiguity. Meanwhile, Chapter 4 shows that older households, or households with lower income or are not married, have a relatively lower ambiguity aversion. Therefore, for the 'aggressive growth' and 'growth' style of mutual funds, the marketing strategy of mutual fund managers should be more focused on these households.

5.4 Limitations of study

This study has the motivation to empirically investigate the impact of ambiguity aversion on the stock market. As a direct consequence of the methodology, the study has encountered a number of limitations, which need to be considered. The main limitation of this study is that due to limited data access, the ambiguity measure is based on US market information only. As a result, all research in this study has been based on the US market. This could lead to a potential bias in the result, since ambiguity aversion relies on social status and mentality. American investors may reveal some distinct preferences towards ambiguity, and the result of this study can be market-specific. However, the American stock market is one of the most developed markets in the world and it is studied widely; therefore this research will be useful for further studies in the field.

The second limitation of this study is the limited number of control variables. Again, due to limited data access, some control variables cannot be constructed and presented in the model. For example, the rapid growth of substitutes to mutual funds such as exchange-traded-funds (ETF), are suggested to impact changes of fund flow. But due to a lack of ETF data, this is not included in the models in Chapter 3. Similarly, religion and political preference are also suggested to have an influence on households' stockholding. Due to privacy policy, such information is not available for the SCF's public users, and therefore it is not presented in the models of Chapter 4. However, based on previous studies, there is no evidence to show that these missing variables have any overwhelming impact. Since the models in this thesis have already considered most of the control variables, missing a small number of control variables is not likely to affect the conclusions significantly.

The third limitation of the study is the measure of ambiguity. The study majorly relies on Anderson et al. (2009)'s measure of ambiguity, which is the divergence of analyst's forecasts about future stock market returns. Such a measure has been chosen as the measure of ambiguity because it is in line with Ellsberg (1961)'s definition of ambiguity, and it corresponds with many applications of ambiguity in finance (Hansen and Sargent, 2001; Ulrich, 2013; Drechsler, 2012; Shi, 2013). However, there are other empirical measures of ambiguity such as age of a firm and analyst coverage (Zhang, 2006), which could be also significant. Therefore future studies could investigate

whether any other ambiguity measure is significantly correlated to stock market participation.

The final limitation of this research is the coverage. This thesis mainly investigates the influence of ambiguity on stock market participation. However, investors can be ambiguous about other aspects of the stocks, such as volatility and liquidity. There are theoretical models which discuss the impact of ambiguity on stock market volatility (Epstein and Schneider, 2010), and the result of this study also shows that investors choose more liquid assets in order to avoid a period of high ambiguity in the stock market. Therefore, it is worthwhile investigating the influence of ambiguity on the stock market volatility, or on market liquidity. Due to limited time it has not been possible to construct a reliable empirical measure of ambiguity on market volatility/liquidity, and thus the corresponding research cannot be accomplished. However, since there is a lack of prior empirical studies in these fields, this could be an important motivation for further research.

5.5 Suggestions for further research

The first suggestion presented here is to extend the research into markets outside the US, especially into non-western countries. As the result indicates, ambiguity aversion has very subtle psychological roots which are related to the social status and mentality of households. Therefore it is very important to investigate ambiguity aversion in markets with a different culture and social structure. For example, Japan is known for its conservative culture and also it has a much older population compared to the US, therefore the scale of the ambiguity aversion could be very different in Japanese market.

Secondly, the impact of ambiguity on volatility and liquidity also needs to be empirically investigated. The market return is not the only parameter that can be ambiguous to investors; various studies conclude that ambiguity in volatility (Epstein and Schneider, 2010) and liquidity (Ozsoylev and Werner, 2009; Eichberger and Spanjers, 2007) can also result in a price premium. Since these studies are still at a theoretical level, empirical evidence is of great importance.

An effort is also required to perform a more detailed analysis of ambiguity on different asset classes. The results of this thesis show that flows of the funds that invest

in non-dividend paying stocks are more sensitive to ambiguous changes. Therefore, it would be beneficial to produce more detailed ambiguity indicators for individual stocks, and to examine the long-term returns of such high-ambiguity portfolios in order to see whether they produce abnormal returns. In addition, this study also finds that during a high ambiguity period in the stock market, investors transfer their capital into much more liquid asset classes such as the money market. As a result, an effort to find the correlation of ambiguity between different asset classes will help to produce hedging strategies, or, at least, will help to avoid the ambiguity.

5.6 Conclusion

In spite of being neglected by mainstream asset pricing models, ambiguity is an important factor in investors' decision making. The behaviour of ambiguity averse investors is not only observed at an aggregate level, but also in individual household's investment decisions. As much theoretical literature has predicted, overall ambiguity reduces the stock market participation and, therefore, results in a price premium. Similar to risk, heterogeneity of ambiguity aversion is observed in different asset classes and different types of investors. Investors in the US market have stronger reactions towards the stocks that pay little or no dividends; and when the stock market experiences a high-ambiguity period, they transfer their capital into more liquid markets. In the US, the married, wealthier investors, and those who have had a negative experience in the stock market, are more ambiguity averse. Meanwhile, older and optimistic people are less ambiguity averse. Although some studies believe that education level can reduce the ambiguity aversion, this study did not find any correlation between education level and ambiguity aversion.

This study fills a gap in knowledge in ambiguity studies by providing empirical evidence that ambiguity reduces the market participation. It also supports the literature that predicts an ambiguity premium in the stock market and helps to identify high/low ambiguity asset classes, and high/low ambiguity averse investor groups. Due to lack of data, the study is limited to the American market and has a limited sample time period. Further research is suggested in order to expand this study into other markets or different asset classes. In addition, empirical studies of ambiguity in market volatility, or market liquidity, are also urgently needed.

References

- Allen, F. and Gale, D., 1994, 'Limited market participation and volatility of asset prices', *American Economic Review*, 84, 933-955.
- Amihud, Y., 2002, 'Illiquidity and stock returns: cross-section and time-series effects', *Journal of Financial Markets*, 7, 31-56.
- Anderson, E., Ghysels, E., and Juergens, J., 2009, 'The impact of risk and uncertainty on expected returns', *Journal of Financial Economics*, 94, 233-263.
- Arin, K., Molchanov, A., and Reich, O., 2007, 'Politics, Stock Markets, and Model Uncertainty', working paper, Massey University.
- Aumann, R., 1997, 'Rationality and Bounded Rationality', *Games and Economic Behaviour*, 21, 2-14.
- Antoniou, C., Galariotis, C., and Read, D., 2014, 'Ambiguity aversion, company size and the pricing of earnings forecasts', *European Financial Management*, 20, 633-651.
- Antoniou, C., Doukas, A., and Subrahmanyam, A., 2014, 'Investor sentiment, beta, and the cost of equity capital', *Management Science*, forthcoming.
- Aydin, S., 2008. An investigation on the language anxiety and fear of negative evaluation among Turkish EFL learners, *Asian EFL Journal: Teaching Articles*, 31, 19-35.
- Baker, M., and Wurgler, J., 2006, 'Investor sentiment and the cross-section of stock returns', *Journal of Finance*, 61, 1645-1680.
- Baker, M., and Wurgler, J., 2007, 'Investor sentiment in the stock market', *Journal of Economic Perspectives*, 21, 129-152.
- Backus, D., Routledge, B., and Zin, S., 2004, 'Exotic preferences for macroeconomists', *NBER Macroeconomics Annual*, 19, 319-390.
- Baele, L., Pungulescu, C., and Horst, J., 2007, 'Model uncertainty, financial market integration and the home bias puzzle', *Journal of International Money and Finance*, 26, 144-158

- Baillie, T., and DeGennaro, P., 1990, 'Stock returns and volatility', *Journal of Financial and Quantitative Analysis*, 25, 203–214.
- Ben-Rephael, A., Kandel, S., and Wohl, A., 2011, 'The price pressure of aggregate mutual fund flows', *Journal of Financial and Quantitative Analysis*, 46, 585-603.
- Ben-Rephael, A., Kandel, S., and Wohl, A., 2012, 'Measuring investor sentiment with mutual fund flows', *Journal of Financial Economics*, 104, 363-382.
- Blalock, G., Just, D., and Simon, D., 2004, 'Hitting the jackpot or hitting the skids: Entertainment, poverty, and the demand for state lotteries', Working Paper, Cornell University.
- Blinder, S., and Morgan, J., 2005, 'Are Two Heads Better than One? Monetary Policy by Committee', *Journal of Money, Credit, and Banking*, 17, 789-812.
- Bochner, S., 1965, 'Defining intolerance of ambiguity', *Psychological Record*, 15, 393-400.
- Bossaerts, P., Ghirardato, P., Guarnaschelli, S., and Zame, W., 2010, 'Ambiguity in Asset Markets: Theory and Experiment', *Review of Financial Studies*, 23, 1325-1359.
- Brenner, M., and Izhakian, Y., 2011, 'Asset Prices and Ambiguity: Empirical Evidence', working paper, stern School of Business,
- Camerer, C. and Weber, M., 1992, 'Recent developments in modelling preferences: Uncertainty and ambiguity', *Journal of Risk and Uncertainty*, 5, 325-370
- Campbell, J., 2006, 'Household Finance', *Journal of Finance*, 4, 1553-1603.
- Cao, H., Wang, T., and Zhang, H., 2005, 'Model uncertainty, limited market participation and asset prices', *Review of Financial Studies*, 18, 1219–1251.
- Caskey, J., 2009, 'Information in equity markets with ambiguity-averse investors', *Review of Financial Studies*, 22, 3595-3627.
- Chan, C., Jegadeesh, N., and Lakonishok, J., 1996, 'Momentum strategies', *Journal of Finance*, 51, 1681-1713.
- Chen, G., Kim, K., Nofsinger, J., and Rui, O., 2003, 'Does investor sophistication influence investing behaviour and trading performance? Evidence from China', working paper, Hong Kong Polytechnic University

- Chen, N., Roll, R., Ross, S., 1986, 'Economic forces and the stock market', *The Journal of Business*, 59, 383-403.
- Chen, Z., and Epstein, L., 2002, 'Ambiguity, Risk, and Asset Returns in Continuous Time', *Econometrica*, 70, 1403-1443.
- Chow, C. and Sarin, K., 2001, 'Comparative ignorance and the Ellsberg paradox', *Journal of Risk and Uncertainty*, 22, 129–139.
- Clotfelter, T. and Cook, P., 1989, 'Implicit taxation in lottery finance', *National Tax Journal*, 40, 533-546.
- Collard, F., S. Mukerji, K. Sheppard, and J. Tallon., 2011, 'Ambiguity and the historical equity premium', working paper, University of Oxford.
- Curley, S., Yates, P., and Abrams, A., 1986, 'Psychological sources of ambiguity avoidance', *Organizational Behaviour and Human Decision Processes*, 38, 230-256.
- Daniel, K., Hirshleifer, D., and Subrahmanyam, A., 1998, 'Investor Psychology and Security Market Under- and Overreactions', *The Journal of Finance*, 53, 1839–1885.
- De Long, B., Shleifer, A., Summers, L., and Waldmann, R., 1991, 'Noise Trader Risk in Financial Markets', working paper, University of Chicago.
- Dimmock, S., Kouwenberg, R., Mitchell, O., and Peijnenburg, K., 2013, 'Ambiguity Aversion and Household Portfolio Choice: Empirical Evidence', working paper, the National Bureau of Economic Research.
- Dimson, E., Marsh, P., and Saunton, M., 2006, 'The Worldwide Equity Premium: A Smaller Puzzle', working paper, London Business School.
- Dow, J., and Werlang, S., 1992, 'Ambiguity aversion, risk aversion, and the optimal choice of portfolio', *Econometrica*, 60, 197–204.
- Drechsler, I., 2012, 'Uncertainty, time-varying fear, and asset prices', *Journal of Finance*, forthcoming.
- Easley, D. and M. O'Hara, 2009, 'Ambiguity and Non-participation: The Role of Regulation', *Review of Financial Studies*, 22, 1817-1844.
- Easley, D., and O'Hara, M., 2010, 'Liquidity and valuation in an uncertain world', *Journal of Financial Economics*, 97, 1-11.

- Edwards, F., and Zhang, X., 1998, 'Mutual Funds and Stock and Bond Market Stability', *Journal of Financial Services Research*, 13, 257-282.
- Eichberger, J., Spanjers, W., 2007, 'Liquidity and Ambiguity: Banks or Asset Markets?', working paper, University of Heidelberg
- Ellsberg, D., 1961, 'Risk, ambiguity, and the Savage axioms', *Quarterly Journal of Economics*, 75, 643– 669.
- Epstein, L., and Schneider, M., 2010, 'Ambiguity and asset market', *Annual Review of Financial Economics*, 2, 315-346.
- Epstein, L., and Wang, T., 1994, 'Intertemporal asset pricing under Knightian ambiguity', *Econometrica*, 62, 283–322.
- Fama, E, and French, K., 1989, 'Business Conditions and Expected Returns on Stocks and Bonds', *Journal of Financial Economics*, 25, 23–49.
- Fama, E, and French, K., 1993, 'Common risk factors in the returns on stocks and bonds', *Journal of Financial Economics*, 33, 3-56.
- Foster, A., Rosenzweig, M., 1995, 'Learning by doing and learning from others: human capital and technical change in agriculture', *The Journal of Political Economy*, 103, 1176-1209.
- Fox, R., and Tversky, A., 1995, 'Ambiguity aversion and comparative ignorance', *The Quarterly Journal of Economics*, 110, 585–603.
- Fox, R., and Weber, M., 2002, 'Ambiguity Aversion, Comparative Ignorance, and Decision Context', *Organizational Behaviour and Human Decision Processes*, 88, 476–498.
- French, K., and Poterba, J., 1991, 'Investor diversification and international equity markets', *American Economic Review*, 81, 222–226.
- French, R., Schwert, W., and Stambaugh, F., 1987, 'Expected stock returns and volatility', *Journal of Financial Economics*, 19, 3–29.
- Frisch, D. and Baron, J., 1988, 'Ambiguity and rationality', *Journal of Behavioural Decision Making*, 1, 149-157.
- Gallaher, S., Kaniel, R., and Starks, L., 2006, 'Mutual funds and advertising, mimeo', working paper, University of Texas at Austin.

- George, T., and Hwang, C., 2004, 'The 52-Week High and Momentum Investing', *The Journal of Finance*, 59, 2145-2176.
- Ghysels, E., Santa-Clara, P., Valkanov, R., 2005, 'There is a risk-return tradeoff after all', *Journal of Financial Economics*, 76, 509–548.
- Gilbert, P., and McGuire, M., 1998, Shame, social roles and status: The psychobiological continuum from monkey to human. In *Shame: Interpersonal Behavior, Psychopathology and Culture*, Gilbert P, Andrews B (eds). Oxford University Press, New York.
- Gleason, A., and Lee, C., 2003, 'Analyst forecast revisions and market price discovery', *The Accounting Review*, 78, 193-225.
- Glosten, R., Jagannathan, R., and Runkle, E., 1993, 'On the relation between the expected value and the volatility of the nominal excess return on stocks', *Journal of Finance*, 48, 1779–1801.
- Grichting, L., 1986, 'The impact of religion on gambling in Australia', *Australian Journal of Psychology* 38, 45–58.
- Guerdjikova, A., 2006, 'Portfolio Choice and Asset Prices in an Economy Populated by Case-Based Decision Makers', working paper, Cornell University.
- Guiso, L., Haliassos, M., and Jappelli, T., 2003, 'Household stockholding in Europe: where do we stand, and where do we go?', working paper, Centre for Economic Policy Research.
- Guiso, L., Sapienza, P., and Zingales, L., 2004, 'The Role of Social Capital in Financial Development', *The American Economic Review*, 94, 526-556.
- Haliassos, M. and Bertaut, C., 1995, 'Why do so few hold stocks?', *The Economic Journal*, 105, 1110-1129.
- Hamm, M., and Bursztajn, H., 1979, 'A medical version of a decision theory paradox: Is it still a paradox?', Paper presented at the First Annual Meeting of the Society for Medical Decision Making, Cincinnati, OH.
- Hansen, P., and Sargent, J., 2001, 'Robust Control and Model Uncertainty', *American Economic Review*, 91, 60-66.
- Harvey, R., 2001, 'The specification of conditional expectations', *Journal of Empirical Finance*, 8, 573–638.

- Heath, C. and Tversky, A., 1991, 'Preference and belief: Ambiguity and competence in choice under uncertainty,' *Journal of Risk and Uncertainty*, 4, 5–28.
- Hirshleifer, D., and Teoh, S., 2003, 'Limited attention, information disclosure, and financial reporting', *Journal of Accounting and Economics*, 36, 337–386.
- Hong, H, and Stein, J., 1999, 'A theory of under-reaction, momentum trading, and overreaction in asset markets', *Journal of Finance*, 54, 2143-84.
- Hong, H., Kubik, J., and Stein, J., 2004, 'Social interaction and stock-market participation', *Journal of Finance*, 59, 137-163.
- Hong, H., Scheinkman, J., and Xiong, W., 2008, 'Advisors and asset prices: A model of the origins of bubbles', *Journal of Financial Economics*, 268-287
- Hsu, J., 2012, 'What drives equity market non-participation?', *North American Journal of Economics and Finance*, 23, 86-114.
- Hvidkjaer, S., 2006, 'Small trades and the cross-section of stock returns', working paper, University of Maryland.
- Investment Company Institute, 2003, *Mutual Fund Fact Book: A Guide to Trends and Statistics in the Mutual Fund Industry*, 43rd Edition.
- Investment Company Institute, 2012, *Investment Company Fact Book: A Review of Trends and Activity in the U.S. Investment Company Industry*, 52nd Edition.
- Ippolito, A., 1992, 'Consumer reaction to measures of poor quality: evidence from the mutual fund industry', *Journal of Law and Economics*, 35, 45-70.
- Jorion, P., 2009, 'Risk Management Lessons from the Credit Crisis', *European Financial Management*, 15, 923–933.
- Kamstra, J., Kramer, A., Levi, D., and Wermers, R., 2011, 'Seasonal asset allocation: evidence from mutual fund flows', working paper, University of Toronto.
- Kamstra, J., Kramer, A., Levi, D., and Wermers, R., 2014, 'Seasonal asset allocation: evidence from mutual fund flows', working paper, University of Toronto.

- Karceski, J., 2002, 'Returns-Chasing Behaviour, Mutual Funds, and Beta's Death', *Journal of Financial and Quantitative Analysis*, 37, 559-594.
- Keller, L. R., Sarin, K., and Souderpandian, J., 2007, 'An examination of ambiguity aversion: Are two heads better than one?', *Judgment and Decision Making*, 2, 390-397.
- Kelsey, D., Ford, J., and Pang, W., 2013, 'Ambiguity in Financial Markets: Herding and Contrarian Behaviour', *Theory and Decision*, 75, 1-15.
- Kelsey, D., Kozhan, R., and Pang, W., 2008, 'Asymmetric momentum effects under uncertainty', *Review of Finance*, 15, 603-63
- Keynes, J., 1921, *A Treatise on Probability*, Cornell University Library, New York.
- Knight, F., 1921, *Risk, Uncertainty and Profit*, Houghton Mifflin Company, Boston, MA.
- Kumar, A., 2009, 'Who gambles in the stock market?', *The Journal of Finance*, 64, 1889-1933.
- Lamont, O., and Jones, C., 2002, 'Short-sale constraints and stock returns', *Journal of Financial Economics*, 66, 207–239.
- Lerner, S., and Tetlock, E., 1999, 'Accounting for the effects of Accountability', *Psychological Bulletin*, 125, 255-275.
- Levy I., Snell J., Nelson J., Rustichini A., and Glimcher P., 2010, 'The neural representation of subjective value under risk and ambiguity', *Journal of Neurophysiology*, 103, 1036-1047.
- Malmendier, D., and Nagel S., 2009, 'Depression Babies: Do Macroeconomic Experiences Affect Risk-Taking?', Working paper, UC Berkeley and NBER
- Mehra, R., 2003, 'The equity premium: Why is it a puzzle?', *Financial Analysts Journal*, 59, 54-69.
- Mehra, R., and Prescott, E., 1985, 'The equity premium: a puzzle', *Journal of Monetary Economics*, 15, 145–161.
- Merton, C., 1973, 'Theory of rational option pricing'. *Bell Journal of Economics and Management Science*, 4, 141–183.

- Michaely, R., Thaler, R., and Womack, K., 1995, 'Price reactions to dividend initiations and omissions: over-reacting or drift?', *Journal of Finance*, 50, 573-608.
- Miller, E., 1977, 'Risk, Uncertainty, and Divergence of Opinion'. *The Journal of Finance*, 32, 1151-1168.
- Mukerji, S., and Tallon, J., 2001, 'Ambiguity aversion and incompleteness of financial markets,' *Review of Economic Studies*, 68, 883-904.
- Nelson, B., 1991, 'Conditional heteroskedasticity in asset returns: a new approach', *Econometrica*, 59, 347-370.
- Newey, W., and West, K., 1987, 'A simple positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix', *Econometrica*, 55, 703-708.
- Newey, W., and West, K., 1994, 'Automatic lag selection in covariance matrix estimation', *Review of Economic Studies*, 61, 631-653.
- Oh, N., and Parwada, J., 2007, 'Relations between Mutual Fund Flows and Stock Market Returns in Korea', *Journal of International Financial Markets, Institutions & Money*, 17 140-151.
- Olsen R., and Troughton G., 2000, 'Are risk Premium anomalies caused by ambiguity?', *Financial Analysts Journal*, 56, 24-31.
- Ozsoylev, N., and Werner, J., 2009, 'Liquidity and Asset Prices in Rational Expectations Equilibrium with Ambiguous Information', working paper, University of Oxford.
- Peng, L., and Xiong, W., 2006, 'Investor attention, over-confidence and category learning', *Journal of Financial Economics*, 80, 563-602.
- Pulford, B., 2009, 'Is luck on my side? Optimism, pessimism, and ambiguity aversion', *Quarterly Journal of Experimental Psychology*, 62, 1079 - 1087.
- Puri, M., and Robinson, D., 2005, 'Optimism and economic choice', working paper, Duke University
- Putnam, D., 1996, 'The Strange Disappearance of Civic America', *The American Prospect*, winter, 34-48.

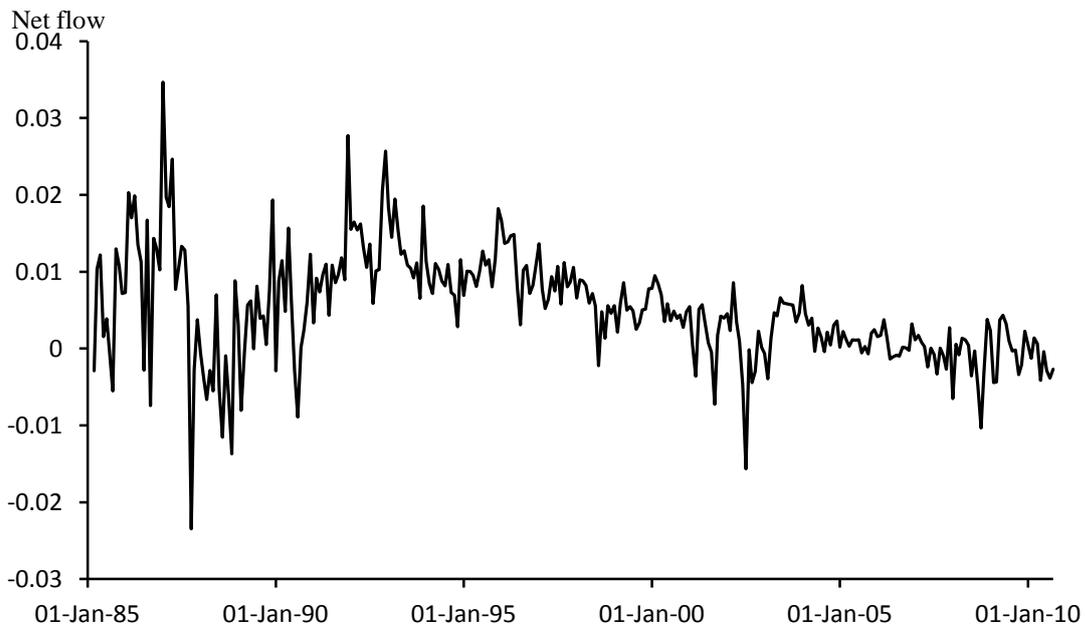
- Rigotti, L., Ryan, M. and Vaithianathan, R., 2004, 'Entrepreneurial innovation', Working paper, Duke University.
- Rubin, D., 1987, *Multiple Imputation for Non-response in Surveys*, New York: John Wiley & Sons.
- Savage, J. 1954. *The Foundations of Statistics*. New York, Wiley.
- Savage, J, and Friedman, M., 1948, 'The Utility Analysis of Choices Involving Risk', *The Journal of Political Economy*, 56, 279-304.
- Shefrin, H., 2002, 'Behavioral decision making, forecasting, game theory, and role-play', *International Journal of Forecasting*, 18, 375–382
- Shi, Z., 2013, 'Time-varying ambiguity and asset pricing puzzles', Working paper, Penn State University.
- Sibley, S., Xing, Y., Zhang, X., 2013, 'Is sentiment sentimental?', working paper, Purdue University.
- Simonson, I., 1989, 'Choice Based on Reasons: The Case of Attraction and Compromise Effects,' *Journal of Consumer Research*, 16, 158-174.
- Sirri, E., and Tufano, P., 1998, 'Costly search and mutual fund flows', *Journal of Finance*, 53, 1589–1622.
- Stickel, S., 1991, 'Common stock returns surrounding earnings forecast revisions: More puzzling evidence', *The Accounting Review*, 66, 402-416.
- Stone, R., and Krähmer, D., 2010, 'Anticipated Regret as an Explanation of Uncertainty aversion', working paper, University of Bonn.
- Sutter, M., Rutzler, M., and Trautmann, S., 2010, 'Impatient and uncertainty: experimental decision predict adolescents field behavior', IZA discussion paper No 5404.
- Takashi, U., 2011, 'The ambiguity premium vs. the risk premium under limited market participation', *Review of Finance*, 15, 245-275.
- Ulrich, M., 2013, 'Inflation ambiguity and the term structure of U.S. government bonds', *Journal of Monetary Economics*, 60, 295-309.

- Viscusi K., and Chesson H., 1999, 'Hopes and fears: the conflicting effects of risk ambiguity', *Theory and Decision*, 47, 153-178.
- Viscusi, W., Magat, W., and Huber, J., 1999, 'Smoking Status and Responses to Ambiguous Scientific Risk Evidence', *Southern Economic Journal*, 66, 250-270.
- Vissing-Jorgenson, A., 2002, 'The returns to entrepreneurial investment: a private equity premium puzzle?', *American Economic Review*, 92, 745-778.
- Walker, D., and Michael B., 1992, *The Psychology of Gambling*, Pergamon Press, Oxford, UK.
- Warther, A., 1995, 'Aggregate Mutual Fund Flows and Security Returns', *Journal of Financial Economics*, 39, 209-235.
- Werner, J., 2001, 'Participation in risk-sharing under ambiguity', working paper, University of Minnesota.
- Williamson. S., 1994, 'Liquidity and Market Participation', *Journal of Economic Dynamics and Control*, 18, 629-670.
- Yaron, A., and Zhang, H., 2000, 'Fixed Costs and Asset Market Participation', *Economic Analysis Review*, 15, 89-109.
- Zhang, F., 2006, 'Information Uncertainty and Stock Returns', *Journal of Finance*, 55, 105-136.

Figure 3.1: Net flows and net exchanges for the equity asset class.

The figure reports the monthly net flows and net exchanges for the equity asset class, which comprises funds within the 'aggressive growth', 'growth', 'sector', 'growth and income', and 'income equity' investment objective categories. The data is from the ICI and covers the period from March 1985 to December 2010. Net flows (Panel A) and net exchanges (Panel B) are calculated according to Equations 3.19 and 3.20.

Panel A. Net flows



Panel B. Net exchanges

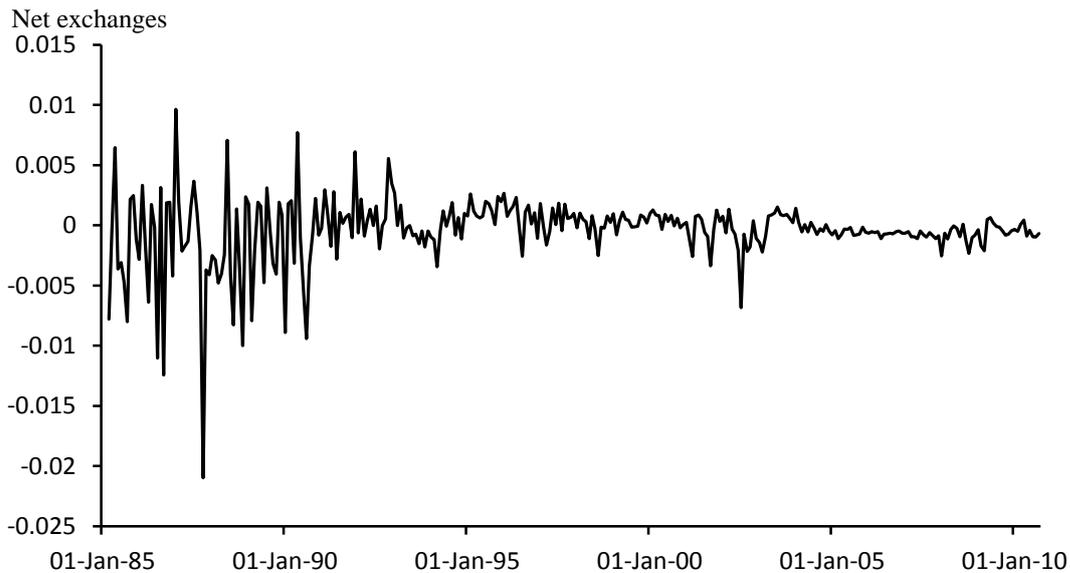
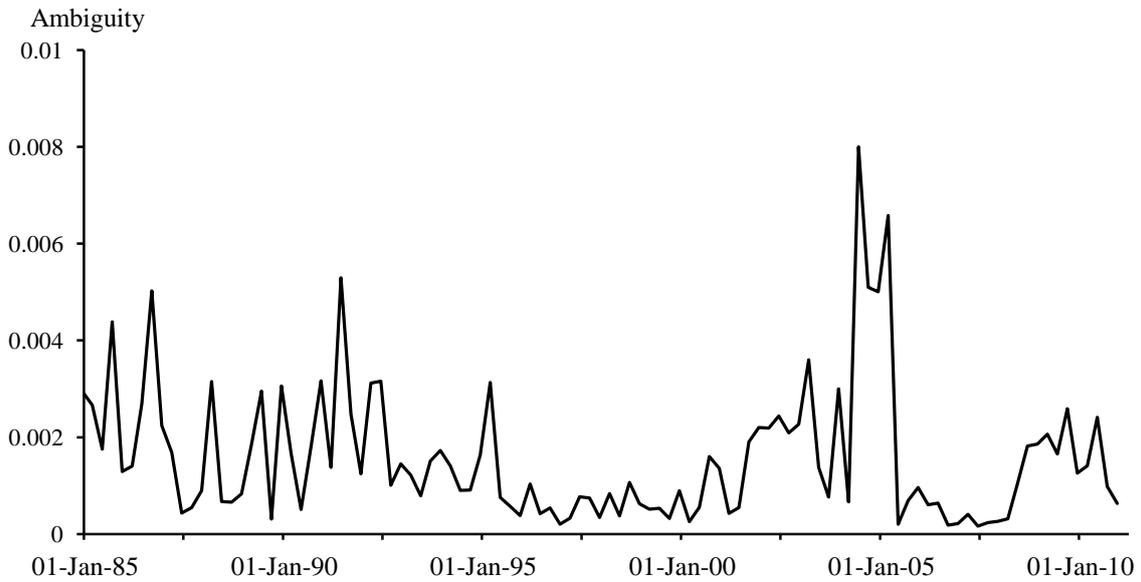


Figure 3.2: Ambiguity

The figure reports the quarterly ambiguity (Panel A) and the changes in the monthly ambiguity (Panel B), from 1985 to 2010. The ambiguity measure reflects the dispersion in forecasts for market returns, obtained using the data from the Survey of Professional Forecasters. To calculate this measure Anderson et al. (2009) is followed, and the forecasts of aggregate output, the output deflator, and corporate profits after taxes are used and combined as per Equation 3.19 with $\nu=15.346$. Monthly ambiguity is computed from the quarterly measure by linear interpolation. Both series are scaled by 100.

Panel A. Quarterly ambiguity



Panel B. Changes in monthly interpolated ambiguity

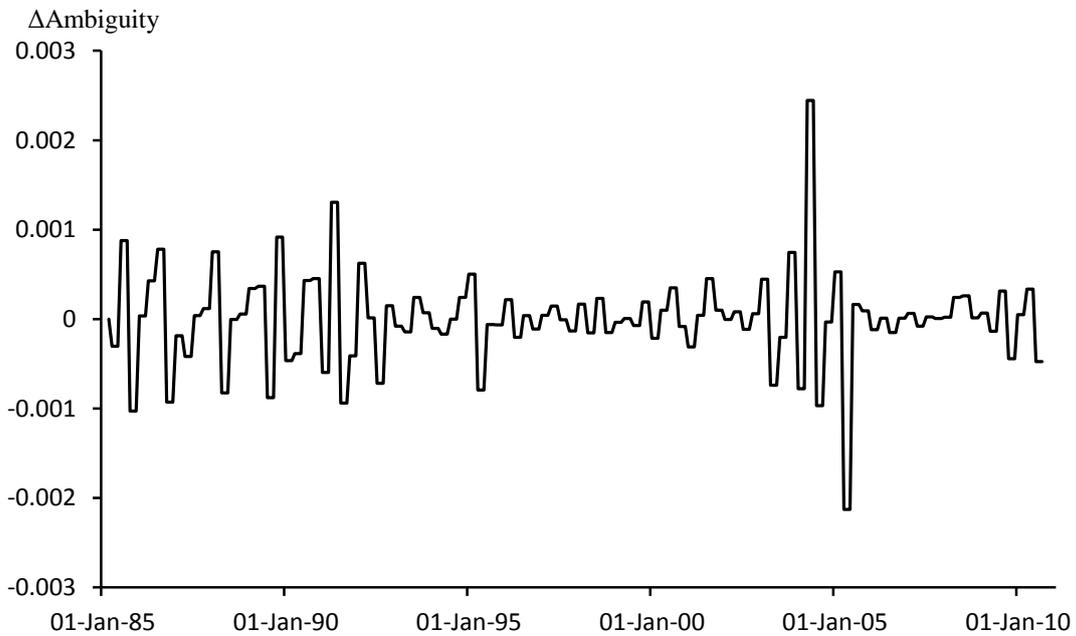


Figure 3.3: Ambiguity and Risk over Time.

The figure reports the monthly ambiguity (dashed line) and the conditional variance of market returns (solid line) from March 1985 to December 2010. The conditional variance is calculated following Anderson et al. (2009) using data from the CRSP. In panel B a scatter plot of the ambiguity and risk is produced.

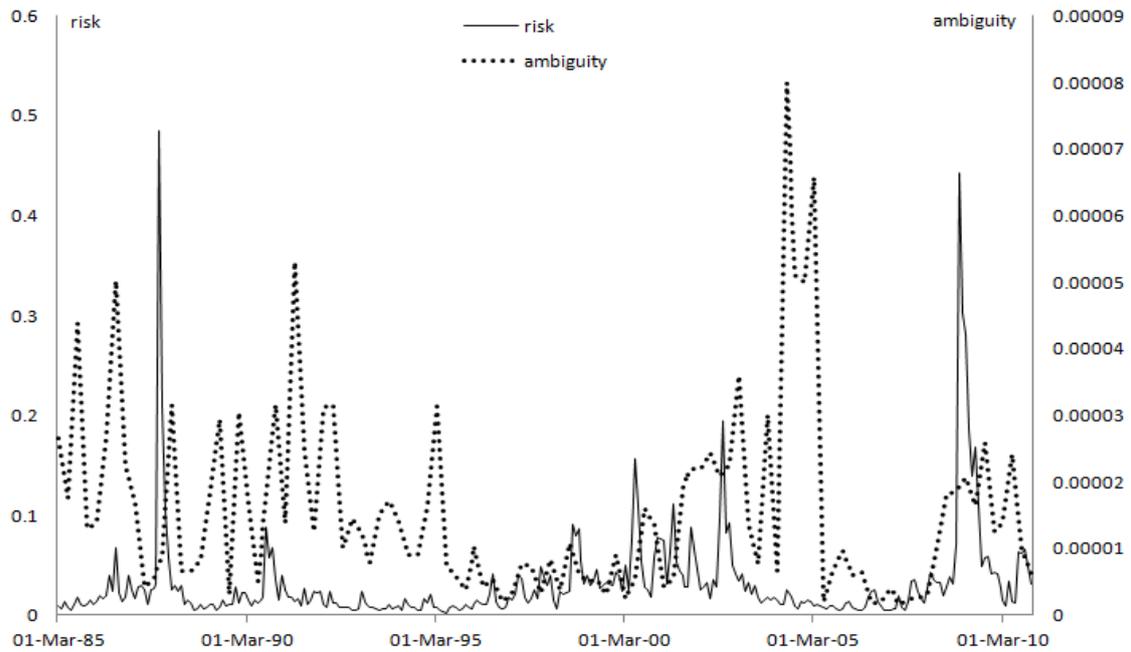


Figure 3.3: Continued.

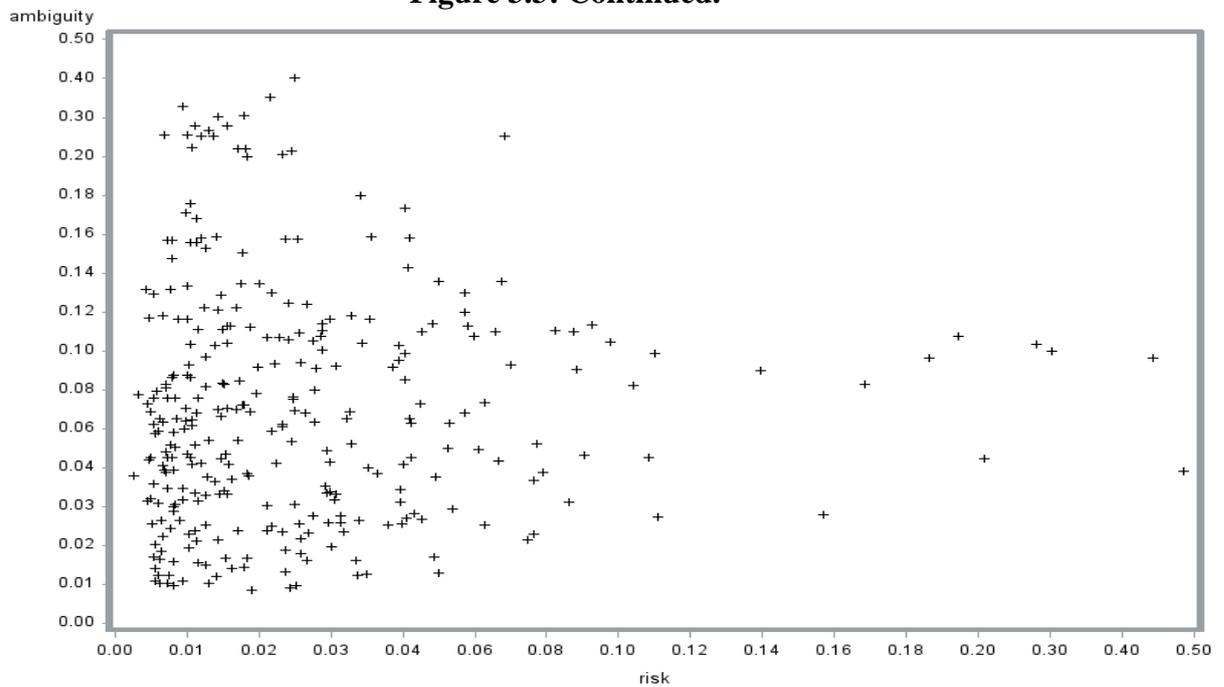


Table 3.1: Classification of mutual funds

The table reports the categorisation of the ICI fund investment objective categories by the asset class based on Kamstra et al. (2011).

Fund Investment Objective	Fund Asset Class
Aggressive Growth	Equity
Growth	Equity
Sector	Equity
Growth and Income	Equity
Income Equity	Equity
Asset Allocation	Hybrid
Balanced	Hybrid
Flexible Portfolio	Hybrid
Income Mixed	Hybrid
Corporate - General	Corporate Fixed Income
Corporate - Intermediate	Corporate Fixed Income
Corporate - Short Term	Corporate Fixed Income
High Yield	Corporate Fixed Income
Strategic Income	Corporate Fixed Income
Government Bond - General	Government Fixed Income
Government Bond - Intermediate	Government Fixed Income
Government Bond - Short Term	Government Fixed Income
Mortgage Backed	Government Fixed Income
State Municipal Bond - General	Government Fixed Income
State Municipal Bond - Short Term	Government Fixed Income
National Municipal Bond - General	Government Fixed Income
National Municipal Bond - Short Term	Government Fixed Income
Taxable Money Market - Government	Money Market

Table 3.2: Summary of Statistics

The table reports the summary of statistics for the various variables for the period from March 1985 to December 2010. The ambiguity measure, amb_t reflects the dispersion in forecasts for the market returns, calculated using the data from the Survey of Professional Forecasters. To calculate this measure Anderson, et al. (2009) is followed, and the forecasts of aggregate output, the output deflator, and corporate profits after taxes are used and combined as per Equation 3.19 with $\nu=15.346$ and scale by 100. $cvar_t$ is the conditional variance calculated according to Anderson et al. (2009). Δamb_t is the change in ambiguity, adv_t is the aggregate cost of print advertising across all the funds, divided by the previous year's total advertising costs, cap_t is the capital gains in month t , from Kamstra et al. (2011, Table 1). sav_t is the personal savings rate taken from the Bureau of Economic Analysis (series PSAVERT), $rfund_{t-1, t-12}$ is the aggregate return of equity funds over the previous 12 months, $rmkt_{t-3, t-1}$ is the return on the CRSP value-weighted index (series VWRETD) over the last 3 months and $llq_{t-3, t-1}$ is the Amihud liquidity measure aggregated on the value-weighted basis and averaged for the previous three months.

	Mean	Std	Skew	Kurt	Max	Min
<i>Net exchanges</i>	0.000	0.003	-2.149	13.960	0.010	-0.021
<i>Net flow</i>	0.005	0.007	0.242	1.692	0.035	-0.023
amb_t	0.002	0.001	1.800	4.206	0.008	0.000
Δamb_t	0.000	0.001	0.296	5.993	0.002	-0.002
$cvar_t$	0.032	0.051	5.496	38.664	0.485	0.003
$\Delta cvar_t$	0.000	0.043	4.827	61.813	0.454	-0.276
$rfund_{t-1, t-12}$	0.167	0.218	-0.966	0.868	0.583	-0.580
adv_t	0.086	0.012	0.811	5.933	0.144	0.038
cap_t	8.408	19.446	2.938	6.827	72.000	0.900
sav_t	0.049	0.019	0.007	-0.801	0.103	0.009
$rmkt_{t-3, t-1}$	0.028	0.085	-1.076	2.855	0.264	-0.367
$llq_{t-3, t-1}$	0.000	0.000	3.015	14.236	0.000	0.000

Table 3.3: Correlations

The table reports the correlation matrix for the various variables for the period from March 1985 to December 2010, defined as in Table 3.2.

	<i>Net flows</i>	<i>Net exchanges</i>	Δamb_t	$\Delta cvar_t$	adv_t	cap_t	sav_t	$rfund_{t-1,t-12}$	$rmkt_{t-3,t-1}$	$llq_{t-3,t-1}$
<i>Net flows</i>	1	0.655	-0.098	-0.033	0.048	0.158	0.316	0.482	0.210	0.291
<i>Net exchanges</i>		1	-0.159	-0.035	-0.021	0.058	-0.049	0.059	0.031	0.001
Δamb_t			1	-0.017	0.028	-0.003	0.046	0.010	-0.017	-0.018
$\Delta cvar_t$				1	-0.028	-0.026	0.003	0.001	-0.137	-0.012
adv_t					1	-0.007	0.025	0.159	0.074	-0.127
cap_t						1	0.006	0.001	-0.072	-0.026
sav_t							1	0.206	0.075	0.484
$rfund_{t-1,t-12}$								1	0.442	0.203
$rmkt_{t-3,t-1}$									1	0.050
$llq_{t-3,t-1}$										1

Table 3.4: Ambiguity and Equity Fund Flows and Exchanges

The table reports the results of estimation of the net flows model (Panel A) and the net exchanges model (Panel B) for the equity asset class, for the period from March 1985 to December 2010. The variables are defined as in Table 3.2. P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test firstly the models shown in equations 3.27 and 3.28 are jointly estimated for all 5 fund families using GMM, and then the residuals are used in order to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator. The classification of mutual funds is based on Kamstra et al (2011), and it is shown in Table 1. All coefficients and standard errors are multiplied by 1000.

Panel A. Net flows				
	Estimate	Std Err	t statistic	prob> t
<i>Intercept</i>	-0.75	1.08	-0.69	0.49
Δamb_t	-1359.33	323.70	-4.21	0.00
$\Delta cvar_t$	-2.21	2.69	-0.82	0.41
adv_t	2.03	10.80	0.19	0.85
cap_t	-0.46	0.21	-2.23	0.03
$rfund_{t-12,t-1}$	0.57	0.53	1.09	0.28
$rmkt_{t-3,t-1}$	-3.50	1.72	-2.04	0.04
<i>Net flow</i> $t-1$	357.39	40.70	8.78	0.00
<i>Net flow</i> $t-2$	145.02	73.20	1.98	0.05
<i>Net flow</i> $t-3$	303.99	32.10	9.46	0.00
sav_t	30.48	10.50	2.89	0.00
$llq_{t-3,t-1}$	14884.77	20389.20	0.73	0.47
Jan_t	1.20	0.63	1.89	0.06
Feb_t	0.00	0.56	0.00	1.00
Nov_t	5.03	1.70	2.96	0.00
Dec_t	37.96	14.60	2.60	0.01
N	309	Adj. R ²	0.52	
P_A	0.09	P_O	0.99	
Panel B. Net exchanges				
<i>Intercept</i>	-0.23	0.47	-0.49	0.62
Δamb_t	-502.43	204.10	-2.46	0.01
$\Delta cvar_t$	-2.12	0.94	-2.27	0.02
adv_t	-1.90	5.37	-0.35	0.72
cap_t	0.01	0.00	3.05	0.00
$rfund_{t-12,t-1}$	0.28	0.22	1.28	0.20
$rmkt_{t-3,t-1}$	-1.02	0.81	-1.25	0.21
<i>Net exchange</i> $t-1$	71.06	24.00	2.96	0.00
<i>Net exchange</i> $t-2$	10.36	21.80	0.47	0.64
<i>Net exchange</i> $t-3$	213.28	23.30	9.14	0.00
$llq_{t-3,t-1}$	-0.23	0.47	0.50	0.62
N	309	Adj. R ²	0.04	
P_A	0.17	P_O	0.99	

Table 3.5: Ambiguity and Different Style Equity Funds

The table reports the estimated coefficients on the change in the ambiguity for each category when estimate the models net flows (Panel A) and net exchanges models (Panel B) for the different style equity funds separately for the period from March 1985 to December 2010. The variables are defined as in Table 3.2. P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test, the models shown in equations 3.27 and 3.28 are jointly estimated for all 5 equity categories using GMM, and then the residuals are used in order to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator. The classification of mutual funds is based on Kamstra (2011), and it is shown in Table 3.1. All coefficients and standard errors are multiplied by 1000.

Panel A: Net flow							
Fund Style	Estimate	Std Err	t statistic	prob> t	N	Adj. R^2	P_A
Aggressive Growth	-2939.61	931.40	-3.16	0.00	309	0.41	0.75
Growth	-1446.01	482.50	-3.00	0.00	309	0.36	0.20
Growth and Income	-558.95	198.40	-2.82	0.01	309	0.66	0.92
Income Equity	-502.37	309.00	-1.63	0.11	309	0.75	0.86
Sector	1465.36	1264.40	1.16	0.25	309	0.17	0.31
P_O	0.99						
Panel B: Net exchanges							
Aggressive Growth	-1973.21	689.00	-2.86	0.00	309	0.16	0.17
Growth	-831.71	285.70	-2.91	0.00	309	0.13	0.13
Growth and Income	-162.82	89.00	-1.83	0.07	309	0.25	0.03
Income Equity	-333.22	141.00	-2.36	0.02	309	0.50	0.79
Sector	1143.76	629.90	1.82	0.07	309	0.07	0.00
P_O	0.99						

Table 3.6: Ambiguity and non-Equity Mutual funds

The table reports the estimated coefficients on the changes in the ambiguity when the net flows (Panel A) and net exchanges models (Panel B) for the non-equity mutual funds are estimated separately for the period from March 1985 to December 2010. The variables are defined as in Table 3.2. P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test the models shown in equations 3.27 and 3.28 are jointly tested for all 5 fund families using GMM, and then the residuals are used in order to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator. The classification of mutual funds is based on Kamstra (2011), and it is shown in Table 1. All coefficients and standard errors are multiplied by 1000.

Panel A: Net flow							
Fund Family	Estimate	Std Err	t statistic	prob> t	N	Adj. R^2	P_A
Hybrid	-582.00	309.20	-1.88	0.06	309	0.73	0.13
Government Fixed Income	-922.56	364.30	-2.53	0.01	309	0.87	0.23
Corporate Fixed Income	-1029.34	406.20	-2.53	0.01	309	0.60	0.73
Money Market	1000.78	636.50	1.57	0.12	309	0.14	0.30
P_O	0.99						
Panel B: Net exchanges							
Hybrid	-56.63	41.80	-1.35	0.18	309	0.61	0.13
Government Fixed Income	-358.85	139.30	-2.58	0.01	309	0.30	0.03
Corporate Fixed Income	-385.64	161.40	-2.39	0.02	309	0.09	0.79
Money Market	223.50	75.70	2.95	0.00	309	0.04	0.00
P_O	0.99						

Table 3.7: Ambiguity and macro economic variables

This table reports the regression results of the one-month-ahead quarterly ambiguity (calculated by using equation 3.19) on quarterly average value of the U.S. recession dummy (NBER), U.S. unemployment rate (Unemp), change in consumption (dCons), default spread (Def), term spread (Term), the value-weighted market return including dividends (VWRETD), from Dec 1968 to Dec 2010. The NBER recession dummy indicates the US recession period, which is a period between a peak and a trough based on the National Bureau of Economic Research (NBER)'s business cycle data. Unemployment (Unemp) rate is from the Bureau of Labour Statistics Data, series ID: LNS14000000. Changes in consumption (Cons) are the changes in the consumption rate of real personal consumption expenditures per capita, which is constructed by dividing the CITIBASE series of seasonally adjusted real consumption (excluding durables) by the Bureau of Census's monthly population, following Chen, Roll and Ross (1986). The default spread (Def) is defined as the difference in yields between Baa-rated corporate bonds and AAA-rated corporate bonds as stated in Fama and French (1989, 1996). The term spread (Term) is defined as the difference in yields between the 10-year Treasury bond and the 3-month T-bill as stated in Chen, Roll, and Ross (1986). Vwretd is the Centre for Research in Security Prices (CRSP)'s value-weighted stock market return index. To adjust for serial correlation, one and two quarter lags of the ambiguity are added. To make result more clear, the estimation of the parameters and the standard errors are multiplied by 1,000,000.

	Estimate	Std Err	t statistic	prob>t
intercept	-1.77	5.92	-0.30	0.77
nber	11.00	4.43	2.55	0.01
unemp	1.70	1.55	1.10	0.27
dcons	3.63	3.93	0.92	0.36
def	-1.83	3.15	-0.58	0.56
term	0.05	1.38	0.04	0.97
vwretd	-0.98	0.54	-1.82	0.07
Amb_t	164010.00	88400.00	1.86	0.07
Amb_{t-1}	252844.00	70900.00	3.57	0.00
N	166	Adj. R ²	0.24	
P_A	0.27			

Table 3.8: Quarterly Regressions

The table reports the results of estimation of the net flows model (Panel A) and the net exchanges model (Panel B) for the combined equity asset class, for the period from March 1985 to December 2010 using non-interpolated quarterly data. Net flows and exchanges are calculated on a quarterly basis. The changes in the ambiguity and conditional variance are equal to $\Delta Qcvar_t = cvar_t - cvar_{t-3}$ and $\Delta Qamb_t = amb_t - amb_{t-3}$, respectively. Quarterly capital gains, savings and advertising costs are equal to the sum of the monthly values over each quarter. Lagged market return, illiquidity premium and fund return are defined as in Table 3.2. In this table lags of one and four quarters are included for the dependent variables. P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test the models shown in equations 3.31 and 3.32 are jointly estimated for all 5 fund families using GMM, and then the residuals are used in order to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator. The classification of mutual funds is based on Kamstra (2011), and it is shown in Table 1. All coefficients and standard errors are multiplied by 1000.

Panel A. Quarterly Net Flows				
	Estimate	Std Err	t -statistic	prob>t
<i>Intercept</i>	-1.77	4.72	-0.37	0.71
Δamb_t	-1483.41	355.50	-4.17	0.00
$\Delta cvar_t$	-38.63	14.60	-2.64	0.01
adv_t	-21.95	19.40	-1.13	0.26
cap_t	-0.02	0.02	-1.08	0.28
$rfund_{t-12,t-1}$	6.62	1.92	3.45	0.00
$rmkt_{t-3,t-1}$	52.60	5.99	8.78	0.00
<i>Net flow</i> $t-3$	630.61	29.50	21.38	0.00
<i>Net flow</i> $t-12$	109.35	28.20	3.87	0.00
$Sav_{t-3,t}$	42.43	9.05	4.69	0.00
$Ilq_{t-3,t-1}$	52067.12	31622.60	1.65	0.10
Dec_t	9.81	1.36	7.21	0.00
N	102	Adj. R ²	0.60	
P_A	0.45	P_O	0.99	
Panel B. Quarterly Net Exchanges				
<i>Intercept</i>	0.80	1.52	0.53	0.60
Δamb_t	-684.93	195.70	-3.50	0.00
$\Delta cvar_t$	-28.95	4.89	-5.93	0.00
adv_t	-9.65	5.91	-1.63	0.11
cap_t	0.02	0.01	2.77	0.01
$rfund_{t-12,t-1}$	1.63	0.83	1.97	0.05
$rmkt_{t-3,t-1}$	9.26	2.17	4.27	0.00
<i>Net exchange</i> $t-3$	230.87	29.60	7.79	0.00
<i>Net exchange</i> $t-12$	49.61	38.80	1.28	0.20
$Ilq_{t-3,t-1}$	-195.61	15292.70	-0.01	0.99
N	102	Adj. R ²	0.15	
P_A	0.36	P_O	0.99	

Table 3.9: Additional Control Variables

The table reports the results of estimation of the net flows model (Panel A) and the net exchanges model (Panel B) for the equity asset class, for the period from 1986 to 2010. All the variables are defined as in Table 3.4. *Sent* is the sentiment index of Baker and Wurgler (2007) at time t , orthogonalized to macroeconomic variables, and *Medforecast* is the median SPF forecast at time t . The remaining variables are defined as in Table 3.2. P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test the models shown in equations 3.27 and 3.28 are jointly estimated for all 5 fund families using GMM, and then the residuals are used to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator. The classification of mutual funds is based on Kamstra (2011), and it is shown in Table 1. All coefficients and standard errors are multiplied by 1000.

Panel A. Net flows								
	Estimate	Std Err	t - statistic	prob> t	P_A	N	Adj. R^2	P_O
Δamb_t	-1500.67	333.6	-4.50	0.00	0.08	309	0.53	0.99
Δrv_t	-28.47	14.3	-1.99	0.05				
Δamb_t	-1246.75	281.2	-4.43	0.00	0.09	309	0.53	0.99
$\Delta GARCH_{1,t}$	-276.71	134.5	-2.06	0.04				
Δamb_t	-1308.78	285.6	-4.58	0.00	0.11	309	0.53	0.99
$\Delta GARCH_{2,t}$	-132.62	36	-3.68	0.00				
Δamb_t	-986.21	244.4	-4.03	0.00	0.61	298	0.65	0.99
Δvix_t	-0.53	0.04	-12.72	0.00				
Panel B. Net exchanges								
Δamb_t	-498.83	204.3	-2.44	0.02	0.24	309	0.04	0.99
Δrv_t	-3.1	7.29	-0.43	0.67				
Δamb_t	-512.01	207.4	-2.47	0.01	0.12	309	0.04	0.99
$\Delta GARCH_{1,t}$	-209.07	72.7	-2.88	0.00				
Δamb_t	-502.7	209.1	-2.40	0.02	0.11	309	0.04	0.99
$\Delta GARCH_{2,t}$	-42.91	15.6	-2.75	0.01				
Δamb_t	-393.28	128.6	-3.06	0.00	0.07	298	0.23	1.00
Δvix_t	-0.26	0.03	-9.73	0.00				

Table 3.10: Different Risk Measures

The table reports the results of estimation of the net flows model (Panel A) and the net exchanges model (Panel B) for the equity asset class, for the period from March 1985 to December 2010. The variables are defined as in Table 3.2. The models are estimated using all controls as in Table 3.4, and only report results on the coefficients on the changes in the ambiguity and the changes in the risk. In each specification a different specification for the risk variable is used: rv_t is realised volatility, which is the sum of daily squared market returns using the CRSP value weighted index. $GARCH_1$ is an in-sample forecast of volatility in month $t+1$ using a simple GARCH model specification, and $GARCH_2$ is an in-sample forecast of market return volatility in month $t+1$ using the GJR GARCH model specification. The specifications of the GARCH models are taken from Glosten, Jagannathan and Runkle (1993). vix_t is the VXO index. For each risk measure the change from month $t-1$ and t is used in the models. The fund flow regressions in Panel A where the risk is measured using the GARCH models, in addition to the controls in Table 4, include a fourth lag of the dependent variables. The fund exchange regression in Panel B where the risk is measured with VIX, in addition to the controls in Table 4, include a sixth lag of the dependent variables. P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test the models shown in equations 3.27 and 3.28 are jointly estimated for all 5 fund families using GMM separately for each risk specification, and then the residuals are used to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator. The classification of mutual funds is based on Kamstra (2011), and it is shown in Table 3.1. All coefficients and standard errors are multiplied by 1000.

Panel A. Net flows								
	Estimate	Std Err	t - statistic	prob> t	P_A	N	Adj. R^2	P_O
Δamb_t	-1500.67	333.6	-4.50	0.00	0.08	309	0.53	0.99
Δrv_t	-28.47	14.3	-1.99	0.05				
Δamb_t	-1246.75	281.2	-4.43	0.00	0.09	309	0.53	0.99
$\Delta GARCH_{1,t}$	-276.71	134.5	-2.06	0.04				
Δamb_t	-1308.78	285.6	-4.58	0.00	0.11	309	0.53	0.99
$\Delta GARCH_{2,t}$	-132.62	36	-3.68	0.00				
Δamb_t	-986.21	244.4	-4.03	0.00	0.61	298	0.65	0.99
Δvix_t	-0.53	0.04	-12.72	0.00				
Panel B. Net exchanges								
Δamb_t	-498.83	204.3	-2.44	0.02	0.24	309	0.04	0.99
Δrv_t	-3.1	7.29	-0.43	0.67				
Δamb_t	-512.01	207.4	-2.47	0.01	0.12	309	0.04	0.99
$\Delta GARCH_{1,t}$	-209.07	72.7	-2.88	0.00				
Δamb_t	-502.7	209.1	-2.40	0.02	0.11	309	0.04	0.99
$\Delta GARCH_{2,t}$	-42.91	15.6	-2.75	0.01				
Δamb_t	-393.28	128.6	-3.06	0.00	0.07	298	0.23	1.00
Δvix_t	-0.26	0.03	-9.73	0.00				

Table 3.11: Different ν values for the ambiguity measure

In this table different values of the weighting parameter ν are used in equation 3.17 to construct four different time series of the ambiguity measure, and then the same regression is run as in table 3.4. Briefly, this table only reports the estimation of the changes of the ambiguity variable. Δamb_t , P_A is the p -value from the Breusch-Godfrey autocorrelation test. To perform the test the models shown in equations 3.27 and 3.28 are jointly estimated for all 5 fund families using GMM separately for each risk specification, and then the residuals are used to test whether they exhibit autocorrelation up to the sixth lag. P_O is the p -value from the test of over-identifying restrictions. Standard errors are adjusted for autocorrelation and heteroscedasticity using the Newey-West (1994) estimator.

Panel A. Net flows								
ν value	Estimate of Δamb_t	Std Err	t -statistic	prob> t	P_A	N	Adj. R ²	P_O
1	0.01	0.02	0.45	0.65	0.14	309	0.52	0.99
5	-0.35	0.10	-3.47	0.00	0.18	309	0.52	0.99
10	-0.91	0.23	-3.96	0.00	0.13	309	0.53	0.99
20	-1.95	0.41	-4.75	0.00	0.08	309	0.53	0.99
Panel B. Net exchanges								
1	0.01	0.01	0.85	0.40	0.12	309	0.02	0.99
5	-0.10	0.05	-1.96	0.05	0.12	309	0.02	0.99
10	-0.28	0.13	-2.15	0.03	0.16	309	0.03	0.99
20	-0.67	0.25	-2.67	0.01	0.14	309	0.04	0.99

Table 4.1: Ambiguity and household stock market participation

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010), to estimate a logistic model where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}, marriage_{i,t}, edu_{i,t}, race_{i,t}$ are dummy variables that flag whether the individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $risk_j(j = 2,3,4)$ is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	Prob>t
<i>amb</i>	-255.479	32.136	<.01
<i>vol</i>	-7.184	1.486	<.01
<i>ach</i>	0.782	0.170	<.01
<i>risk2</i>	1.221	0.069	<.01
<i>risk3</i>	1.142	0.038	<.01
<i>risk4</i>	1.519	0.046	<.01
<i>edu</i>	0.729	0.031	<.01
<i>income</i>	0.621	0.013	<.01
<i>race</i>	0.708	0.043	<.01
<i>marriage</i>	0.163	0.036	<.01
<i>retirement</i>	0.221	0.060	<.01
<i>age</i>	0.025	0.001	<.01
N	176160	R ²	0.34

Table 4.2: Robustness for ambiguity and household stock market participation

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The data of those who have income less than 1000 dollars and have an age below 25 and above 75 is removed. The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether the individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $riskj_{i,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	Prob>t
<i>amb</i>	-275.552	33.789	<.01
<i>vol</i>	-7.997	1.562	<.01
<i>ach</i>	0.782	0.180	<.01
<i>risk2</i>	1.233	0.072	<.01
<i>risk3</i>	1.162	0.042	<.01
<i>risk4</i>	1.536	0.049	<.01
<i>edu</i>	0.735	0.032	<.01
<i>income</i>	0.638	0.014	<.01
<i>race</i>	0.685	0.044	<.01
<i>marriage</i>	0.164	0.038	<.01
<i>retirement</i>	0.278	0.069	<.01
<i>age</i>	0.023	0.001	<.01
N	156939	R ²	0.34

Table 4.3: Ambiguity and household stock market participation: The effect of income

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns, calculated using $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether the individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $riskj_{i,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable *highic* (*lowic*) is set to 1 if a household i in cohort t decaded income of more than \$125,000 (\$15,000) per year in 1993, and 0 otherwise. These figures are deflated using the CPI of each year and converted into the corresponding income level for each survey. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-228.269	36.194	<.01
<i>amb * highic</i>	-170.473	59.474	<.01
<i>amb * lowic</i>	210.214	110.945	0.06
<i>vol</i>	-7.140	1.483	<.01
<i>ach</i>	0.783	0.170	<.01
<i>income</i>	1.217	0.069	<.01
<i>risk2</i>	1.126	0.038	<.01
<i>risk3</i>	1.506	0.046	<.01
<i>risk4</i>	0.719	0.031	<.01
<i>race</i>	0.505	0.023	<.01
<i>marriage</i>	0.702	0.043	<.01
<i>retirement</i>	0.154	0.036	<.01
<i>age</i>	0.201	0.060	<.01
<i>edu</i>	0.026	0.001	<.01
<i>highic</i>	0.538	0.111	<.01
<i>lowic</i>	-0.809	0.185	<.01
N	176160	R ²	0.34

Table 4.4: Ambiguity and household stock market participation: The effect of age

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of the Dow index up until the last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}, marriage_{i,t}, edu_{i,t}, race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $risk_j_{i,t} (j = 2,3,4)$ is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable old is set to 1 if the main respondent of household i in time t is above 65 years old. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-287.560	34.601	<.01
<i>amb * old</i>	151.463	59.883	0.01
<i>vol</i>	-7.239	1.487	<.01
<i>ach</i>	0.757	0.171	<.01
<i>income</i>	1.222	0.069	<.01
<i>risk2</i>	1.148	0.038	<.01
<i>risk3</i>	1.525	0.046	<.01
<i>risk4</i>	0.627	0.013	<.01
<i>race</i>	0.705	0.043	<.01
<i>marriage</i>	0.162	0.036	<.01
<i>retirement</i>	0.193	0.061	<.01
<i>age</i>	0.022	0.001	<.01
<i>edu</i>	0.733	0.031	<.01
<i>old</i>	-0.075	0.107	0.48
N	176160	R ²	0.34

Table 4.5: Ambiguity and household stock market participation: The effect of education

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model where the dependent variable $SMP_{i,t}$ is a binary indicator of stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $risk_{j,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-231.957	41.940	<.01
<i>amb * edu</i>	-43.145	49.714	0.39
<i>vol</i>	-7.178	1.485	<.01
<i>ach</i>	0.787	0.170	<.01
<i>income</i>	1.221	0.069	<.01
<i>risk2</i>	1.143	0.038	<.01
<i>risk3</i>	1.520	0.046	<.01
<i>risk4</i>	0.621	0.013	<.01
<i>race</i>	0.708	0.043	<.01
<i>marriage</i>	0.163	0.036	<.01
<i>retirement</i>	0.221	0.060	<.01
<i>age</i>	0.025	0.001	<.01
<i>edu</i>	0.794	0.081	<.01
N	176160	R ²	0.34

Table 4.6: Ambiguity and household stock market participation: The effect of marriage

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $risk_{j,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-164.074	53.618	<.01
<i>amb * marriage</i>	-122.306	57.733	0.03
<i>vol</i>	-7.368	1.489	<.01
<i>ach</i>	0.764	0.171	<.01
<i>income</i>	1.221	0.069	<.01
<i>risk2</i>	1.143	0.038	<.01
<i>risk3</i>	1.520	0.046	<.01
<i>risk4</i>	0.620	0.013	<.01
<i>race</i>	0.707	0.043	<.01
<i>marriage</i>	0.353	0.097	<.01
<i>retirement</i>	0.218	0.060	<.01
<i>age</i>	0.025	0.001	<.01
<i>edu</i>	0.730	0.031	<.01
N	176160	R ²	0.34

Table 4.7: Ambiguity and household stock market participation: The effect of social capital

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price of the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $risk_j_{i,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable for social capital sc is set equal to 1 if a household is reported to have participated in voluntary charity work in the last year, and 0 for the rest. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-261.841	37.628	<.01
<i>amb * sc</i>	-19.229	50.876	0.71
<i>vol</i>	-8.417	1.504	<.01
<i>ach</i>	0.618	0.173	<.01
<i>income</i>	1.209	0.069	<.01
<i>risk2</i>	1.132	0.038	<.01
<i>risk3</i>	1.506	0.046	<.01
<i>risk4</i>	0.612	0.013	<.01
<i>race</i>	0.709	0.043	<.01
<i>marriage</i>	0.149	0.036	<.01
<i>retirement</i>	0.219	0.060	<.01
<i>age</i>	0.025	0.001	<.01
<i>edu</i>	0.711	0.031	<.01
<i>sc</i>	0.223	0.083	0.01
N	176160	R ²	0.34

Table 4.8: Ambiguity and household stock market participation: The effect of optimism level

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price for the last year/the historical high price of the Dow index up until the last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}, marriage_{i,t}, edu_{i,t}, race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $riskj_{i,t} (j = 2,3,4)$ is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable $optim$ is set to 1 if a household i in cohort t believe that both the US economy improves in the next 5 years compared to the last 5 years, and their family income will grow faster than the inflation. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-263.575	32.538	<.01
<i>amb * optim</i>	165.868	92.352	0.07
<i>vol</i>	-6.894	1.493	<.01
<i>ach</i>	0.822	0.171	<.01
<i>income</i>	1.218	0.069	<.01
<i>risk2</i>	1.141	0.038	<.01
<i>risk3</i>	1.515	0.046	<.01
<i>risk4</i>	0.727	0.031	<.01
<i>race</i>	0.620	0.013	<.01
<i>marriage</i>	0.712	0.043	<.01
<i>retirement</i>	0.161	0.036	<.01
<i>age</i>	0.222	0.060	<.01
<i>edu</i>	0.025	0.001	<.01
<i>optim</i>	-0.154	0.145	0.29
N	176160	R ²	0.34

Table 4.9: Ambiguity and household stock market participation: The effect of life-time experienced return

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price for the last year/the historical high price of the Dow index up until the last year. The variable $income_{i,t}$ is the logged reported income of a household i for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $riskj_{i,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable $negative$ is set equal to 1 if the household has experienced a negative life-time stock market return, which is calculated from formula (4.16), and 0 for the rest. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	689.101	381.062	0.07
<i>amb * negative</i>	-943.226	379.435	0.01
<i>vol</i>	-6.761	1.517	<.01
<i>ach</i>	0.758	0.174	<.01
<i>income</i>	1.222	0.069	<.01
<i>risk2</i>	1.142	0.038	<.01
<i>risk3</i>	1.519	0.046	<.01
<i>risk4</i>	0.621	0.013	<.01
<i>race</i>	0.706	0.043	<.01
<i>marriage</i>	0.163	0.036	<.01
<i>retirement</i>	0.221	0.060	<.01
<i>age</i>	0.025	0.001	<.01
<i>edu</i>	0.730	0.031	<.01
<i>negative</i>	1.173	0.458	0.01
N	176160	R ²	0.34

Table 4.10: Ambiguity and household stock market participation: The effect of smoking

In this table the data is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price for the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}$, $marriage_{i,t}$, $edu_{i,t}$, $race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of a respondent, and $risk_{j,t}$ ($j = 2,3,4$) is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable $smoking$ is set equal to 1 if the household is reported to be a smoking household, and 0 for the rest. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t
<i>amb</i>	-276.414	33.239	<.01
<i>amb * smoking</i>	-41.081	74.618	0.58
<i>vol</i>	-7.618	1.483	<.01
<i>ach</i>	0.782	0.171	<.01
<i>income</i>	1.217	0.069	<.01
<i>risk2</i>	1.138	0.038	<.01
<i>risk3</i>	1.513	0.046	<.01
<i>risk4</i>	0.614	0.013	<.01
<i>race</i>	0.714	0.043	<.01
<i>marriage</i>	0.155	0.036	<.01
<i>retirement</i>	0.215	0.060	<.01
<i>age</i>	0.024	0.001	<.01
<i>edu</i>	0.698	0.031	<.01
<i>smoking</i>	-0.306	0.112	0.01
N	176160	R ²	0.34

Table 4.11: Ambiguity and household stock market participation: the joint test

In this table the data used is from the Survey of Consumer Finances, which is conducted every three years (1989 to 2010) to estimate a logistic model, where the dependent variable $SMP_{i,t}$ is a binary indicator of the stock market participation of a household i . The ambiguity measure $amb_{t-1,t-3}$ reflects the average ambiguity during the three years before year t when the survey is conducted. The ambiguity measure is based on the dispersion in the forecasts for market returns. $vol_{t-1,t-3}$ is the corresponding realised volatility based on the return of the CRSP value-weighted index (series VWRETD). Ach_t is the average Dow index price for the last year/the historical high price of Dow index until last year. The variable $income_{i,t}$ is the logged reported income of a household for the previous year, deflated to 2010 dollars. $retirement_{i,t}, marriage_{i,t}, edu_{i,t}, race_{i,t}$ are dummy variables that flag whether an individual is retired, married, holds a college degree and white, respectively. $age_{i,t}$ is the reported age of respondent, and $risk_{j,t} (j = 2,3,4)$ is set to 1 if the respondent has chosen the answers 2, 3 or 4 when asked about his/her risk preferences. The dummy variable $smoking$ is set equal to 1 if the household is reported to a smoking household, and 0 for the rest. The dummy variable $negative$ is set equal to 1 if the household has experienced a negative life-time stock market return, which is calculated from formula (4.16), and 0 for the rest. The dummy variable $optim$ is set to 1 if the household i in cohort t believes that the US economy improves in the next 5 years compare to the last 5 years. The dummy variable for social capital sc is set equal to 1 if the household is reported to have participated in voluntary charity work in the last year, and 0 for the rest. The dummy variable old is set to 1 if the main respondent of the household i in time t decaded is above 65 years old. The dummy variable $highINC$ is set to 1 if the household i in cohort t reported income of more than \$125,000 per year in 1993, and 0 otherwise. The regression model is 4.28. Standard errors are adjusted for multiple imputations.

	Estimate	Std Err	prob>t		Estimate	Std Err	prob>t
<i>amb</i>	-267.531	67.970	<.01	<i>op</i>	-0.088	0.087	0.32
<i>vol</i>	-8.564	1.541	<.01	<i>sc</i>	0.166	0.085	0.05
<i>ach</i>	0.535	0.178	<.01	<i>smoking</i>	-0.287	0.115	0.01
<i>income</i>	1.206	0.070	<.01	<i>negative</i>	-1.180	0.461	0.01
<i>risk2</i>	1.135	0.039	<.01	<i>high inc</i>	0.391	0.112	<.01
<i>risk3</i>	1.505	0.046	<.01	<i>amb * older</i>	157.554	61.525	0.01
<i>risk4</i>	0.578	0.021	<.01	<i>amb * op</i>	70.593	52.911	0.18
<i>race</i>	0.712	0.043	<.01	<i>amb * ed</i>	28.738	52.757	0.59
<i>marriage</i>	0.263	0.099	<.01	<i>amb * sc</i>	15.038	52.133	0.77
<i>retirement</i>	0.191	0.061	<.01	<i>amb * smoking</i>	-46.879	75.879	0.54
<i>age</i>	0.021	0.002	<.01	<i>amb * marriage</i>	-77.608	59.475	0.19
<i>edu</i>	0.640	0.086	<.01	<i>amb * negative</i>	926.275	380.721	0.02
<i>smoking</i>	-0.100	0.109	0.36	<i>amb * high inc</i>	-171.801	61.553	0.01
<i>older</i>	-267.531	67.970	<.01				
N	176160	R ²	0.35				

Table 4.12: Robustness for Table 4.3-4.10

This table reports the results of robustness for tables 4.3 - 4.10. For tables 4.3 - 4.8 and 4.10, the model is again tested with the sample that removes the households who have an income lower than \$1000, and the households who are older (younger) than 75 (25) years old. For table 4.9, due to a strong cohort effect, most of the samples that experienced positive life time stock market returns are young. Therefore, for table 4.10 the new sample removes the households who have an income lower than \$1000, and the households who are older than 75 years old. For brevity, this table reports an estimation on the ambiguity and interactions only.

	Estimate	Std Err	prob>t	N	R ²
<i>amb</i>	-243.003	37.727	<.01	156939	0.34
<i>amb*highic</i>	-184.199	61.371	<.01		
<i>amb</i>	-301.167	35.501	<.01	156939	0.34
<i>amb*old</i>	172.304	73.168	0.02		
<i>amb</i>	-243.698	44.901	<.01	156939	0.34
<i>amb*edu</i>	-56.087	52.506	0.29		
<i>amb</i>	-177.548	58.704	<.01	156939	0.34
<i>amb*marriage</i>	-127.612	62.786	0.04		
<i>amb</i>	-285.240	39.808	<.01	156939	0.34
<i>amb*sc</i>	-11.235	53.448	0.83		
<i>amb</i>	-285.438	34.250	<.01	156939	0.34
<i>amb*optim</i>	183.853	95.748	0.05		
<i>amb</i>	-271.932	33.581	<.01	156939	0.34
<i>amb*negative</i>	819.818	390.290	0.04		
<i>amb</i>	-295.696	35.006	<.01	156939	0.34
<i>amb*smoke</i>	-56.821	77.508	0.46		

Appendix

Here there are examples of the questions and variables taken from the survey to construct the regression variables. This example is based on the 2010 SCF survey. Please note that the content of the questions and the variable numbers can be changed in different years.

Stock market participation:

X3913: Do you (or anyone in your family living here) own any stock which is publicly traded?

X3821: I need to know what types of funds you have. Do you have stock mutual funds?

Race:

X6814: Which of these categories do you feel best describe you:

White, black or African-American, Hispanic or Latino, Asian, American Indian or Alaska Native, Hawaiian Native or other Pacific Islander, or another race?

Marriage:

X7372(#1): What is your current legal marital status? Are you married, separated, divorced, widowed, or have you ever been married?

Education:

X5905(#1): What is the highest degree (you/he/she/he or she) have(s) ?

Age:

X13: Respondent: Age computed from date of birth

Retirement:

X4100(#1): What is your current work status?

Income:

X5729: Roughly how much was the total income you (and your family living here), received in 2009 from all sources, before taxes and other deductions were made?

Risk:

X3014: Which of the following statements comes closest to describing the amount of financial risk that you and your husband/wife/partner are willing to take when you save or make investments?

1. Take substantial financial risks expecting to earn substantial returns
2. Take above average financial risks expecting to earn above average returns
3. Take average financial risks expecting to earn average returns
4. Not willing to take any financial risks

Optimism

X301: How do you think the future of US economy in the next 5 years compared to the last 5 years?

1. Better
2. The same
3. Worse

X303 : Over the next five years, do you expect your total (family) income to go up more than prices, or about the same as prices?

1. Up more
2. Up less
3. About the same