LIMITED STOCK MARKET PARTICIPATION AND AMBIGUITY AVERSION

ADVISOR: FABIO A. MACCHERONI
DISCUSSANT: MASSIMO MARINACCI

MASTER OF SCIENCE THESIS OF:
CHU, YINXIAO
STUDENT ID: 1491117

ACADEMIC YEAR: 2012/2013

March 2013
Milan, Italy
At times our own light goes out and
is rekindled by a spark from another person.
Each of us has cause to think with deep gratitude of
those who have lighted the flame within us.
—Albert Schweitzer

ACKNOWLEDGMENTS

It seems that this tiny space and my flat words may not be enough
to express my gratitude to all the people along the way to obtain
my achievements in the past two and half years in Milan. However,
I deeply realize that I can hardly reach my attainment without their
care, support, guidance, assistance and, even, arguments with them.
These people are:

My family members, not only including my parents but also my
grandparents, aunts, uncles, and cousins are the very first ones I
would like to thank to. I thank them not only for the distant sup-
porting but also their care from the moment when I was born. They
nurture me, encourage me and even punish me when I was wrong in
my early age.

And I owe much to the friends whom I have met in Milan. They
have been constantly assisting my life in Milan, and their support
eases my life far away from my hometown. I owe as much to my old
friends who are apart from me everywhere in the world.

Last but most important and responsible for my experience of study-
ing at Bocconi, I am greatly indebted to those who teach me, advise
me and enlighten me. They are the professors in the Bocconi Uni-
versity, among whom I, especially, appreciate Prof. Maccheroni’s dedica-
tion for tutoring me and other support. In addition, many thanks to
Prof. Battigalli, Prof Botticini and Prof. Marinacci for the instructive
talks with them.
To my dearest parents and grandparents.
The limited stock market participation is a well-documented phenomenon despite a high equity premium in both the US and other developed economies. However, many attempts cannot successfully explain it. A simple version of the puzzle is formulated in a two-asset portfolio choice problem under mean-variance preference with a participation cost. Although it can generate non-participation as an optimum, we cannot get a reasonable estimation of relative risk aversion coefficient from data. Robust mean-variance preference is introduced and it captures the feature of ambiguity aversion. The high risk aversion can be reconciled under robust mean-variance preference by adjusting the unobserved ambiguity with reasonable parameter values. Also, the results reveal the fact that ambiguity aversion affect one’s portfolio decision substantially. Furthermore, some sources of ambiguity are studied to capture some ideas about the ambiguity and ambiguity aversion in turn. Finally, extended theoretical discussions illustrate some qualitative effects of learning on momentum trading, and also a generalized multiple-asset framework is proposed.
LIST OF TABLES

Table 1  Estimates in 2004 & 2007  15
Table 2  Estimates of robust mean-variance model  21
Table 3  Estimates with simple linear models of return  24
Table 4  Estimates with ambiguous idiosyncratic information  25
INTRODUCTION

The financial market puzzles economists because there exist some discrepancies between well-formed theories and empirical observations. One of them being noticed is the limited household participation in stock market despite of a very high equity premium. After synthesize the data from Panel Study of Income Dynamics (PSID) in 1984, Mankiw and Zeldes (1991) summarize the fact that, of the 2998 sampled families, only 27.6 percent hold some stocks, and among those who hold other liquid assets in excess of 100,000, only 47.7 percent hold stock. In addition, Haliassos and Bertaut (1995) document the same phenomenon, that is, roughly 75% of the US households are kept themselves away from stock market. More importantly, limited stock market participation is not a unique phenomenon in the United States; in fact, it exists in most the developed markets (See Household Portfolios edited by Guiso, Haliassos, and Jappelli, 1999).

The importance of understanding limited stock market participation is that it helps to account for other related puzzles in financial economics, prominently, equity premium puzzle (Mehra and Prescott, 1985). Mankiw and Zeldes (1991) shows that the correlation between consumption growth and equity return of stockholder and non-stockholder provides a potential explanation of the empirical failure of the consumption-CAPM model. However, understanding why majority of the households choose not to hold stock is indeed another difficult task. Haliassos and Bertaut (1995) also conclude that risk aversion per se, heterogeneity in beliefs, habit persistence, time non-separability and borrowing constraints cannot be used to account for the phenomenon.

However, limited participation in stock market itself does not constitute the puzzle; the indeed puzzle is that the theoretical framework with reasonable parameters cannot be reconciled with empirical data, or the empirical data implies an unreasonable parameter under the theoretical framework. The very first puzzle is that, in the presence of equity premium, zero stock-holding will never be the case for a well-informed expected utility maximizer without market friction and under the assumption of rational expectation.

But, if considering some market frictions, especially a participation cost, it is possible to model the situation where agent holds zero stock as an optimum (Haliassos and Michaelides, 2001). Some empirical evidences also support the argument of participation cost (Vissing-Jorgensen, 2002). But, calibrating some observed parameters from empirical data, under mean-variance model, will generate an in-
consistently high parameter values, like risk aversion coefficient and participation cost, which renders the puzzle not fully resolved.

The motivation of this paper is to show that it offers promising solutions of puzzles in financial decision to introduce new features of agent’s preferences. Ellsberg’s famous thought experiment suggests that the decision maker (DM) behave substantially differently under ambiguous situation when the objective probability is not told than in the risky situation where the only objective probability is known. This feature has great implication on people’s financial decision since usually the distribution of asset return is not known. There are several utility representations are developed to capture the DM’s aversion to ambiguity. Among them, the smooth model (Kilbanoff et al. 2005) enjoys some advantages in tractability and similarity with the expected utility theory. The advantageous potion of introducing ambiguity lies in two facts: 1) it is well-axiomatic, which means that the analysis based on incorporating ambiguity has a well-formed behavioral foundation; 2) ambiguity is proved to account for equity premium puzzles (Ju and Miao, 2012).

In this paper, I apply the robust mean-variance preference (Maccheroni et al., 2010), which is further derived from the smooth model in the similar way of applying Arrow-Pratt approximation on the expected utility model, to study a simple household portfolio choice problem. Before that, in an analytical exercise, it is possible to obtain a optimum where the representative household with a mean-variance preference (under the expected utility framework) does not hold stock. However, when introducing the heterogeneity (in wealth and beliefs) to the model and matching it to the empirical data (Survey of Consumer Finance, 2004 & 2007), it turns out that the estimated relative risk aversion coefficient is beyond the reasonable range, say less than 2, which lies in the heart of the puzzle, while the estimated participation cost is moderate and in agreement with other studies.

To reconcile the high risk aversion, robust mean-variance preference is employed to show that a lower relative risk aversion coefficient suffices to match the empirical distribution of stock-holding analytically. Again, the calibration and estimation of the empirical data under this framework reveals that, as a matter of fact, the ambiguity aversion (the product of ambiguity and attitudes towards ambiguity) in one’s decision is a truly substantial component to determine the portfolio weight. However, our knowledge about DM’s ambiguity towards financial market and their attitudes towards ambiguity is very limited. But admitting a robust mean-variance functional has the advantage to allow a relative rich comparative statics to study how changes in DM’s ambiguity and attitude to ambiguity affect her decision.

To further explore the agent’s attitude towards ambiguity, I propose some scenario to construct proxies of agent’s ambiguity and,
thus, recover the agent’s ambiguity aversion attitude in some sense. Furthermore, interesting questions of that how the DM reacts to the changes in the observables are also discussed qualitatively. If some conditions hold for a learning structure, we will observe DM’s behavior as a momentum trader. Finally, I extend the framework into a multiple-asset setting, which allows us to study more problems concerning comparison between several assets.

1.1 RELATED LITERATURE

In term of related literature, besides the empirical investigations of limited stock market participation and decision-theoretical papers capture the ambiguity aversion, there are some papers discussing model uncertainty and ambiguity aversion and portfolio decisions, and more papers concerning ambiguity and financial market in general. Dow and Werlang (1992) pioneer the application of ambiguous preference to portfolio choice problem; using capacity instead of measure, they show the condition where no trades happen. Epstein and Schneider (2007) give a simple example of non-participation of a multiple-priors agent with one ambiguous asset; the key intuition is that the participation cost arises endogenously from learning under ambiguity. Also, Cao et al. (2005) and Easley and O’Hara (2009) demonstrate that non-participation is endogenously arise when taking into account of ambiguity-averse investor modeled by multiple-prior utility. Considering the application of smooth ambiguity model in studying finance problems, Ju and Miao (2012) provides a nice example. In the calibrated model, their results match the empirical data. Maccheroni et al. (2010) develop the robust mean-variance preference, which is the basic building block for this paper.

1.2 ORGANIZATION

The rest of the paper is organized as following. Chapter 2 starts with a simple review of Arrow-Pratt approximation and mean-variance preference, and then proceeds to an approximation of smooth model under ambiguity and robust mean-variance model, which captures the agent’s ambiguity aversion. In Chapter 3, a simple household portfolio problem is formulated, where it shows that, with a participation cost, zero stock-holding is optimal for some agents, however, puzzles emerge when calibrating and estimating the risk aversion coefficient from empirical data. Following the previous steps, I re-evaluated the household portfolio problem with robust mean-variance using same data in Chapter 4. In the Chapter 5, there are some discussions on decomposing ambiguity and attitude towards ambiguity. Chapter 6 is an extension of relevant topics; it contains a qualitative analysis of
how the DM react to changes in observed stock prices and a general framework for multiple assets. Finally Chapter 7 concludes.
This chapter is a review of mean-variance preference and robust mean-variance preference. For a more rigorous treatment, please refer to Maccheroni, Marinacci and Ruffino (2010).

2.1 THE MEAN-VARIANCE PREFERENCE

The Arrow-Pratt approximation (1964) provides the mean-variance preference a theoretical foundation, hence a careful examination of Arrow-Pratt approximation is going to reveal some meaningful aspects of mean-variance preference itself. Consider an expected utility maximizer decision maker (DM), facing a risky investment \( h \), whose Bernoulli utility function is \( u \) and wealth is \( w \), the Arrow-Pratt approximation of his certainty equivalent for the uncertain result of investment \( w + h \) is

\[
c(w + h, P) = u^{-1}(E_P(u(w + h))) \approx w + E_P(h) - \frac{1}{2} \lambda_u(w) \text{Var}_P(h)
\]

(1)

where \( P \) is the probabilistic model that describes the uncertainty of \( h \), and \( \lambda_u(x) = -\frac{u''(x)}{u'(x)} \) is the risk aversion coefficient at wealth level \( w \) by definition. Connecting this to the mean-variance preference, where \( f \) is the uncertain prospect and \( \lambda \) is the parameter attached to agent’s dislike to risk, the following form

\[
U(f) = E_P(f) - \frac{1}{2} \lambda \text{Var}_P(f)
\]

(2)

is straightforward, simply by setting \( f = w + h \) and \( \lambda = \lambda_u(w + E_P(h)) \). Note that if normalizing the expectation of the risk prospect, i.e. \( w = 1 \), then \( \lambda = \lambda_u(1) = -\frac{u''(1)}{u'(1)} \) is not only the absolute risk aversion coefficient, but also equal to the relative risk aversion coefficient. Under the assumption of constant relative risk aversion (CRRA)\(^{1}\), \( \lambda \) is the relative risk aversion coefficient at any wealth level.

2.2 AMBIGUITY AND THE SMOOTH MODEL

However, the DM may not only be faced with a risky investment where there is a uniquely defined probabilistic model to capture the uncertainty, but indeed she would also be confronted with an ambiguous situation in which the agent does not have only one well-defined

\(^{1}\) Indeed, we can assume such specification of utility function that \( u(x) = \frac{x^{1+\lambda}}{1+\lambda} \).
probabilistic model to describe the uncertain prospect. The famous Ellsberg thought experiment reveals the fact that agent’s choice under ambiguity is not consistent with the subjective expected utility theory a la Savage.

To capture the feature of the ambiguity and ambiguity aversion, Klibanoff, Marinacci and Mukerji (2005, KMM hereafter) introduced and axiomized the smooth model, where it assumes that the DM considers alternative models $Q$ instead of the single model $P$, and $\Delta$ is the space of possible models $Q$ and $\mu$ is the DM’s prior probability measure on space $\Delta$. The probability space is given by $(\Omega, \mathcal{F}, P)$ and the Radon-Nikodym derivative $q = dQ/dP$ is the square integrable density of probability measures $Q$ on $\mathcal{F}$. The DM ranks the prospect through the following functional:

$$V(f) = \int_\Delta \phi \left( \int_\Omega u(f(\omega)) q(\omega) dP(\omega) \right) d\mu(q)$$

where $f$ is defined on a Hilbert space of square integrable random variables (denoted by $L^2 = L^2(\Omega, \mathcal{F}, P)$) consisting of its almost surely bounded elements (denoted by $L^\infty = L^\infty(\Omega, \mathcal{F}, P)$) with the essential supremum and infimum of $f$ contained in $I$, which is denoted by

$$L^\infty(I) = \{ f \in L^\infty : \text{essinf } f, \text{essup } f \in I \}$$

Set $v = \phi \circ u$, then the above functional can be also rewritten (2.3) in the form

$$V(f) = \int_\Delta (v \circ u^{-1}) \left( \int_\Omega u(f(\omega)) q(\omega) dP(\omega) \right) d\mu(q)$$

Consequently, the certainty equivalent is given by

$$C(f) = u^{-1} \left( \phi^{-1} \left( \int_\Omega u(f(\omega)) q(\omega) dP(\omega) \right) d\mu(q) \right), \forall f \in L^\infty(I).$$

Using the functional (2.4), equivalently,

$$C(f) = v^{-1} \left( \int_\Delta v \left( u^{-1} \left( \int_\Omega u(f(\omega)) q(\omega) dP(\omega) \right) \right) d\mu(q) \right), \forall f \in L^\infty(I)$$

It shows that the certainty equivalent $C(f)$ is decomposed into two monetary certainty equivalents,

$$c(f, Q) = u^{-1} \left( \int_\Omega u(f(\omega)) q(\omega) dP(\omega) \right)$$

and

$$C(f) = v^{-1} \left( \int_\Delta v(c(f, Q)) d\mu(q) \right)$$

KMM, in conclusion, shows that the attitude towards ambiguity, i.e. the model uncertainty, is characterized by the evaluation function $\phi$. 

2.3 ROBUST MEAN-VARIANCE PREFERENCE

Particularly, the concavity of $\phi$ implies the ambiguity aversion and a positive Arrow-Pratt coefficient $\lambda_\phi = -\phi''/\phi'$. Notice that

$$\lambda_\phi(u(w)) = \frac{1}{u'(w)} (\lambda_\nu(w) - \lambda_u(w))$$ \hspace{1cm} (10)

due to $\lambda_\nu - \lambda_u > 0$ is a key condition to ensure ambiguity aversion, i.e. the concavity of $\phi$.

2.3 ROBUST MEAN-VARIANCE PREFERENCE

Thanks to the smooth nature of the evaluation function $\phi$, the smooth model has the tractability comparable to expected utility model. Thus, the extended the Arrow-Pratt approximation can be applied to the certainty equivalent function. Let $w \in \text{int } I$ be a scalar representing current wealth level, and there is an uncertain investment $h \in L^\infty$ such that $w + h \in L^\infty(I)$. The extended Arrow-Pratt approximation yields

$$C(w + h) \approx w + E_Q(h) - \frac{1}{2} \lambda_u(w) \text{Var}_\mu(h) - \frac{1}{2} (\lambda_\nu(w) - \lambda_u(w)) \text{Var}_\mu(E_Q(h))$$ \hspace{1cm} (11)

where

$$E_Q(h) = \int_\Omega h dQ$$ \hspace{1cm} (12)

$$\text{Var}_\mu(E_Q(h)) = \int_\Delta (E_Q(h))^2 d\mu - \left( \int_\Delta (E_Q(h)) d\mu \right)^2$$ \hspace{1cm} (13)

$$\text{Var}_\mu(h) = \int_\Omega h^2 d\bar{Q} - \left( \int_\Omega hd\bar{Q} \right)^2$$ \hspace{1cm} (14)

and $\bar{Q}$ is the probability measure on $\mathcal{F}$ defined as

$$\bar{Q}(A) = \int_\Delta Q(A) d\mu(Q), \forall A \in \mathcal{F}$$ \hspace{1cm} (15)

$\bar{Q}$ is called reduction of $\mu$ on $\Omega$, which can be interpreted as reduction of compound lotteries.

Parallel to the mean-variance preference, the robust mean-variance functional, $C : L^2 \rightarrow \mathbb{R}$, which the DM uses to rank the prospect $h$ in $L^2$ is given by

$$C(h) = E_p(h) - \frac{\lambda}{2} \text{Var}_p(h) - \frac{\theta}{2} \text{Var}_\mu(h), \forall h \in L^2$$ \hspace{1cm} (16)

under the assumption that $\bar{Q} = P$, which allows us to compare it with the classic case. The robust mean-variance functional can be viewed as a local quadratic approximation of smooth ambiguity preference at some $w$ such that $\lambda = \lambda_u(w)$ and $\theta = \lambda_\nu(w) - \lambda_u(w)$. The parameters $\lambda$ and $\theta$ describe the DM’s negative attitudes attached to the risk and ambiguity, respectively. Note, similarly, if normalizing the
baseline wealth level \( w = 1 \), the \( \lambda = \lambda_u(1) \) is, again, both absolute risk aversion coefficient at 1 and the constant relative risk aversion coefficient if assuming the DM is CRRA. On top of that, if we further assume that \( -\frac{v''(w)}{v'(w)} w \) is constant, we can show that the agent is also constant relative ambiguity averse (CRAA)\(^2\) defined by KMM; furthermore, at normalized unit wealth level, one can also show that, in the robust mean-variance utility, \( \theta = \lambda_v(1) - \lambda_u(1) \) is the constant relative ambiguity coefficient.

\(^2\) Similar to the specification of \( u(\cdot) \), we can assume that \( v(x) = \frac{x^{1-\nu}}{1-\nu} \).
Even though there is an economically significant equity premium, holding stock is not a common investment decision made by majority of the households. Non-participation in stock market is a well-documented fact observed not only in the United States, but also in other developed economies. Theoretically, introduction of a participation cost is enough to generate non-participation for the agents with mean-variance preference. However, as mentioned in the introduction, a simple calibration exercise will result in a tension between parameter values; furthermore, estimates of relative risk aversion parameter from micro-data is beyond the reasonable region.

3.1 Portfolio Allocation with a Participation Cost

3.1.1 A representative agent model of non-participation

To simplify the problem, we consider a simple one-period portfolio allocation problem with only two assets, a risk-free asset with return $r_f$ and a risky asset (say stock) with a return $r_e$. The DM use probabilistic model $P$ to describe the stochastic nature of $r_e$. The initial wealth level $w$ is normalized to unit, and also there is a participation cost $c$ realized if the agent decide to hold stock, so $\kappa = c/w$ is defined as the cost-to-wealth ratio, or simply the normalized cost. So the end-of-period wealth is equal to $r_f$ if the agent only holds risk-free asset, and, if holding stock, it would be the total return from the investment deducted by the cost of holding stock:

$$r_q = \begin{cases} r_f + q \cdot (r_e - r_f) - \kappa & q > 0 \\ r_f & q = 0 \end{cases} \quad (17)$$

where $q$ is the wealth (weight) invested in stock, which needs to be determined.

A discussion on participation cost

Intuitively, a participation cost in stock market is widely-observed market friction and composed of some types of transaction costs, such

---

1. Such a simple model can be justified by assuming the DM is myopic. Indeed, this assumption is more realistic than assumptions in a dynamic model with infinite horizon. More importantly, such simple model will be enough tailored to account for the problem.

2. The probabilistic model $P$ can be actually the true distribution of data, or just a construction in DM’s belief.
as the entry cost, which is used to set up the account and search for information, and commission fee for each transaction; else, for those who participate stock market by investing in mutual fund, they have to pay a management fee for the fund managers. However, the participation cost is not only restricted to literal payment, but also consisted of the time and efforts invested to determine the portfolio and the psychological burden for doing risky investment, which is associated with opportunity cost. Indeed, there is a line of empirical research showing that many factors have significant impact on the decision of whether to hold stocks; but from my perspective, such factors, e.g. social trust (Georgrakos and Pasini, 2011), financial literacy (van Rooij et al., 2011), information technology (Bogan, 2008), IQ (Grinblatt et al. 2011) etc., are part of the participation cost. If part of the participation cost affects the portfolio decision substantially, it is, then, necessary to be incorporated into the models to account for relevant empirical puzzles.

Another line of research is aiming at estimating the participation cost. Vissing-Jorgensen (2002) estimates that a $50 participation cost in price of 2000 per period is enough to generate half non-participation; and in 1989 and 1994 the 75% non-participation can be generated by a participation cost of 150 per year (about $260 in 2000 price). Alan (2005) estimates a one-time entry cost which is roughly 2% of the (annual) permanent income. This result is much lower than the estimation of Haliassos and Michaelides (2003).

Else, the participation cost should vary from individual to individual. However, in the rest of the chapter, it is convenient and almost innocuous to stick to the assumption of a fixed participation cost. Taking into account of the heterogeneity in wealth levels, the fixed participation cost generates a pattern between the cost and wealth, that is, a wealthier household has a smaller normalized participation cost. This offers a simple but good approximation aligned with the empirical studies. For example, high wealth level is positively correlated to higher education, higher IQ, and easier access to information technology, which as discussed before lower the participation.

Finally, a further analysis later will review that variety and uncertainty in participation cost do affect the portfolio decision, but empirically the variety in participation cost is too small to alter the decisions significantly. Moreover, for the purpose of estimation, a fixed participation cost results in a linear model that can be easily estimated.

To find out the optimal allocation of wealth, the DM solves the maximization problem (18) subject to the no-short-selling condition

\[
\max_q U(r_q) = E_p(r_q) - \frac{1}{2} \lambda Var_p(r_q)
\]

(18)

\[^3\text{It is usual that household investors do not short sell, and most of them have no access to hedge fund investment either.}\]
The solution is simply given by the standard solution of mean-variance portfolio allocation,
\[
\hat{q} = \frac{E_p(r_e) - r_f}{\lambda \text{Var}_P(r_e)}
\]
if \( E(r_e) - r_f \geq 0 \), i.e. in the presence of equity premium; otherwise, \( \hat{q} = 0 \). This implies that, under the restriction of no-short-selling, non-participation is generated only when there is a negative equity premium, and this fact is not affected by whether there is a participation cost.

However, in the presence of equity premium, the agent still need to compare the utility that can be attained by only holding risk-free asset, which is \( U(r_0) = r_f \), with the maximum utility of holding the costly stock, which is
\[
U(r_{\hat{q}}) = r_f + s - \kappa
\]
where \( s = \frac{[E_p(r_e) - r_f]^2}{2 \lambda \text{Var}_P(r_e)} \) is the excess utility for investing in stock. This leads to the decision of whether to hold stock in the presence of equity premium solely depends on the sign of \( (s - \kappa) \). In sum, the agent decides not to hold any stock if the excess utility of holding stock is less than the (normalized) cost, i.e.
\[
s - \kappa < 0
\]

### 3.2 The Empirical Puzzle

#### 3.2.1 The tension between parameter values in the calibration

Substitute \( \kappa = c/w \) and rearrange the inequality, equation (22) is equivalent to
\[
\lambda c > \frac{[E_p(r_e) - r_f]^2}{2 \text{Var}_P(r_e)}w.
\]

A simple calibration exercise could reveal the tension between two parameters, the relative risk aversion coefficient, \( \lambda \), and the fixed participation cost, \( c \). To see this, assuming the evaluation period is one-year long\(^5\), \( E_p(r_e) - r_f \) is nothing but equity premium, which is

---

4 The fact that \( \lambda \) here is approximately equal to the relative risk aversion coefficient need to be justified. Two ways to think about it. Suppose that the agent is CRRA indeed, and Taylor expansion to derive mean-variance utility function is taken place centering on \( E_p(r_q) \), which by construction is very close to 1. So it is possible to do the approximation around 1; therefore \( \lambda = \lambda_u(1) \), which as discussed before \( \lambda_u(1) \) equates the constant relative risk coefficient. This argument will be held also for the following estimation part.

5 Since that we assume that the DM is myopic, they will re-balance the portfolio after certain period. Benartzi and Thaler (1995) argue that myopic loss aversion could offer some explanation for equity premium. They also calibrate that the myopic DM evaluate portfolios annually to match the equity premium.
commonly recognized as 6% annually; the standard deviation of annual stock return is roughly 20%. Considering the at least 60% of the households do not hold any stock, and from Table A1 (in Appendix), we can see the threshold for 50% participation is in the 7th decile of financial assets, which is ranging from 47,000 $ to 137,800 $, and hence, let $w = 50,000$. we obtain the inequality that $\lambda c > 2,250$. For example, the participation cost is estimated to be around 300 $, then the risk aversion coefficient is over 7.5. But what do we know about risk aversion? Macro-economists and finance researchers believe that the relative risk aversion coefficient should not exceed 2 in most cases, and it leads to the lower bound of participation cost over 1,125 $, which is far above the estimation in other studies (see Vissing-Jorgensen, 2002).

3.2.2 Estimation from micro-data

By introducing heterogeneity in wealth and idiosyncratic information, it is possible to estimate a uniform participation cost assuming everyone has the same relative risk aversion coefficient as well. The idea is that the variation in participation cost and relative risk aversion coefficient is neglected since its small magnitude and little impact, while wealth distribution and a random idiosyncratic information constitute a substantial component of one’s decision.

the return of stock is added by a piece of idiosyncratic information $\epsilon_i$ that is orthogonal to market uncertainty, i.e. $r_{e,i} = r_e + \epsilon_i$; and for the purpose of estimation, let us assume that $\epsilon_i$ is normally distributed among households with zero mean and variance $\sigma^2_{\epsilon}$, i.e. $\epsilon_i \sim N(0, \sigma^2_{\epsilon}) \forall i$. The idiosyncratic information can be easily recognized as either insider information, or different interpretation of corporation’s financial statements, or anythings else. However, it is unobserved. The zero mean condition, intuitively, is aligned with non-arbitrage condition, which says that, at the end of the day, the trading activities homogenize the heterogeneous belief into one price. Then the expected return conditional on $i$’s information is just the market expectation corrected by the idiosyncratic information, i.e. $E_p(r_{e,i}|\epsilon_i) = E_p(r_e) + \epsilon_i$.

Furthermore, for each individual household $i$, its wealth is denoted by $w_i$, and consequently, its cost-to-wealth ratio is denoted by $\kappa_i = c/w_i$. The heterogeneous wealth levels do not have an effect on asset returns, but heterogenize the normalized participation cost. And wealthier households have a relatively lower participation cost.

Finally, the end-of-period wealth of holding stock is

$$r_{q,i} = r_f + q_i \cdot (r_{e,i} - r_f) - \kappa_i$$

where $q_i$ is the weight of wealth invested into stock market and is chosen by the DM of household $i$. 

Models for estimation

The optimization problem and its solution are similar to what are formulated before, but taking into account the heterogeneity, which is
\[
\max_{q_i} U(q_{q,i}) = E_p(r_{q,i}) - \frac{1}{2} \lambda Var_p(r_{q,i}).
\] (25)

\[
s.t. \ q_i \geq 0
\] (26)

The solution is
\[
\hat{q}_i = \frac{E_p(r_{e,i}|\epsilon_i) - r_f}{\lambda Var_p(r_{e,i}|\epsilon_i)} = \frac{E_p(r_e) + \epsilon_i - r_f}{\lambda Var_p(r_e)}
\] (27)

if \( E(r_e) + \epsilon_i - r_f > 0 \) (positive equity premium condition); otherwise \( \hat{q}_i = 0 \), again. Parallel to the analysis in representative model, the maximum utility can be attained by holding stock is
\[
U(q_{q,i}) = r_f + s_i - \kappa_i
\] (28)

where \( s_i = \frac{[E_p(r_{e,i}|\epsilon_i) - r_f]^2}{2\lambda Var(r_{e,i})} = \frac{[E_p(r_e) + \epsilon_i - r_f]^2}{2\lambda Var(r_e)} \). Thus, individual participation decision depends only the sign of \((s_i - \kappa_i)\).

On top of positive equity premium condition, simple algebra leads to
\[
s_i - \kappa_i > 0 \Leftrightarrow \frac{E_p(r_e) - r_f}{\sigma_e} - \sqrt{2Var_p(r_e)\frac{\lambda c}{\sigma_e^2}} \sqrt{\frac{1}{w_i} + \frac{\epsilon_i}{\sigma_e^2}} > 0.
\] (29)

It shows that the right hand side of the inequality (3.12) is linear in \( \sqrt{1/w_i} \), and it is possible to estimate a standard probit model with a linear potential outcome equation with following specification:
\[
y_i = \beta_0 + \beta_1 \sqrt{1/w_i} + \epsilon_i / \sigma_e
\] (30)

where \( \beta_0 = \frac{E_p(r_e) - r_f}{\sigma_e}, \beta_1 = -\sqrt{2Var_p(r_e)\frac{\lambda c}{\sigma_e^2}} \). \( w_i \) is the wealth of household \( i \), and \( \epsilon_i / \sigma_e \) has a standard normal distribution. Consequently, we can also recover the estimation of \( \sigma_e \) and \( \lambda c \) thanks to the nice property of MLE in the following way:
\[
\hat{\sigma_e} = \frac{E_p(r_e) - r_f}{\hat{\beta}_0}
\] (31)
\[
\hat{\lambda c} = \frac{\hat{\beta}_1^2 \sigma_e^2}{2Var_p(r_e)}
\] (32)

where \( \hat{\beta}_0 \) and \( \hat{\beta}_1 \) are the MLEs of \( \beta_0 \) and \( \beta_1 \), respectively.

A further step to estimate is to exploit the observed weights that each household invest into stock. Since we only observe the stock-holding of those who decide to buy stock, the estimation has to evoke

\[\text{Assume this condition holds throughout the paper unless specified.}\]
Heckman selection model with minor modification, whereas the main equation is

\[ \hat{q}_i = \frac{E_p(r_e) + \epsilon_i - r_f}{\lambda Var_p(r_e)} = \frac{E_p(r_e) - r_f}{\lambda Var_p(r_e)} + \frac{\epsilon_i}{\lambda Var_p(r_e)} \] (33)

and the selection equation is

\[ \beta_0 + \beta_1 \sqrt{\frac{1}{w_i}} + \frac{\epsilon_i}{\sigma_{\epsilon}} > 0 \] (34)

which is the same as (3.13). Some more calculations result in

\[ E(\hat{q}_i|\epsilon_i/\sigma_{\epsilon} > -N_i) = 1 \cdot \frac{E_p(r_e) - r_f + \sigma_{\epsilon} \cdot \gamma(N_i)}{\lambda Var_p(r_e)} \] (35)

where \( N_i = \beta_0 + \beta_1 \sqrt{\frac{1}{w_i}} \), inverse mills ratio \( \gamma(N_i) = \frac{\varphi(N_i)}{\Phi(N_i)} \), and \( \varphi(\cdot) \) and \( \Phi(\cdot) \) are density function and cumulative probability function of standard normal distribution respectively. \( N_i \) and \( \frac{E_p(r_e) - r_f + \sigma_{\epsilon} \cdot \gamma(N_i)}{\lambda Var_p(r_e)} \) are possible to be constructed by using the estimates in the previous steps. Then, we can identify \( 1/\lambda \) by regressing \( q_i \) on \( \frac{E_p(r_e) - r_f + \sigma_{\epsilon} \cdot \gamma(N_i)}{\lambda Var_p(r_e)} \) without constant term, and thus uncover \( \lambda \).

**SUMMARY OF ESTIMATION STEPS** is therefore followed:

1. Run a probit model, where the dependent variable is whether respondent participates stock market \( D_i \), the independent variable is the square root of the reciprocal of liquid wealth, \( \sqrt{1/w_i} \).

\[ D_i = \begin{cases} 1 & \text{if } \beta_0 + \beta_1 \sqrt{1/w_i} + \epsilon_i/\sigma_{\epsilon} > 0 \\ 0 & \text{otherwise} \end{cases} \] (36)

2. Recover the estimation of \( \sigma \) and \( \lambda c \) followed by the estimated \( \hat{\beta}_0 \) and \( \beta_1 \):

\[ \hat{\sigma}_{\epsilon} = \frac{E_p(r_e) - r_f}{\hat{\beta}_0} \] (37)

\[ \hat{\lambda} \hat{c} = \frac{\hat{\beta}_1^2 \hat{\sigma}_{\epsilon}^2}{2 \lambda Var_p(r_e)} \] (38)

3. Compute the inverse Mills ratio \( \gamma(N_i) \) for each observation using \( \hat{\sigma} \) and \( \hat{\lambda} \hat{c} \) estimated before, where \( N_i = \hat{\beta}_0 + \hat{\beta}_1 \sqrt{1/w_i} \).

4. regress \( q_i \) observed on \( y_i = \frac{E_p(r_e) - r_f + \epsilon_i \gamma(N_i)}{\lambda Var_p(r_e)} \) (without constant term), i.e. regression equation

\[ q_i = \gamma y_i \] (39)

Then the estimation of \( \lambda \) is \( \hat{\lambda} = \frac{1}{\hat{\gamma}} \), and consequently \( \hat{c} = \hat{\lambda} \hat{c} / \hat{\lambda} = \hat{\lambda} \hat{c} \).
Table 1: Estimates in 2004 & 2007

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\lambda}$</th>
<th>$\hat{c}$</th>
<th>$\hat{\sigma}_\epsilon$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>4.0</td>
<td>275</td>
<td>0.106</td>
</tr>
<tr>
<td>2007</td>
<td>5.2</td>
<td>180</td>
<td>0.107</td>
</tr>
</tbody>
</table>

$\hat{\lambda}$: the estimates of risk aversion
$\hat{c}$: the estimates of the cost

Data

Survey of Consumer Finance (SCF) is a normally triennial cross-sectional survey of U.S. families, but offers panel data occasionally. The survey data includes information on assets composition of U.S. households. Since we focus on how the liquid wealth is allocated in the U.S. households, household $i$’s liquid (or investable) wealth $w_i$ is defined by the sum of bonds, mutual funds, stocks, and other kinds of savings. Not only direct stock holding is considered, but also indirect stock holding in investment of mutual funds. The consideration is that the household decide to hold stock, but he decide to transfer their efforts, a form of participation cost, into management fee, another type of participation cost. Nevertheless, as long as the household decide to invest in mutual funds, they know that they are going to assume the uncertainty in stock market. This is different from pension account even though it is common that pension funds usually hold stock; however, the scheme of pension payment is more or less fixed when the contract is signed, so the uncertainty in stock market should has little influence in the decision of signing the pension scheme. Thus, the proportion of stock in total liquid wealth $q_i$ is calculated by the sum of direct and indirect stock holding divided by total liquid wealth.

3.2.2.1 The estimation results

Using the data in 2004 and 2007, the parameters are estimated with samples in different years, respectively, listed in Table 1. As illustrated in the representative agent model, the participation cost is more or less in the reasonable region (less than $300), however, the risk aversion coefficient is still a bit higher than reasonable value (both exceeding 2), after introducing heterogeneity.

And notice that, in different year, the estimations vary from each other, the variation in participation cost is easy to be accounted for since a lot of facts could affect participation cost physically and psychologically. However, the instability of risk aversion is indeed not well-accepted. Considering in an experiment setting, if the risk situation is given, although it is possible to observe the variation of risk attitude among individuals, the overall distribution should stay stable for different representative samples. Both the magnitude and
instability of risk aversion coefficient suggest that the puzzles of portfolio formation may be rooted in the way that we model uncertainty that faced by the agent. The introduction of ambiguity aversion is just aiming at understanding DM’s behavior when being faced with not only the risky events, but also model uncertainty about describing such risk.
In the light of the fact that the households are never well-informed of the nature of uncertainty in stock market, it should be more appropriate to model the stock return as an ambiguous prospect that the DM holds a non-singleton set of probabilistic models to describe its uncertainty instead of a unique model. In this view, modeling agent with ambiguity aversion can be closer to the reality and provide more insights; and the ambiguity aversion is captured by the robust mean-variance utility.

4.1 Model

Once again, recall the end-of-period wealth is the same as that in the previous chapter,
\[ r_{q,i} = r_f + q_i (r_{e,i} - r_f) - \tilde{\kappa}_i \]
where \( \tilde{\kappa}_i = \tilde{c}/w_i \)\(^1\) and \( r_{e,i} = r_e + \epsilon_i \). However, in this case, the stock return, \( r_e \), is ambiguous rather than risky in the sense that the DM has a set, \( \Delta \), consisting of all the probabilistic models, \( Q \), and she also has a probability measure, \( \mu \), over the set \( \Delta \)\(^2\). Then the way that DM of household \( i \) ranks the uncertain prospect switches to robust mean-variance functional and maintain the assumption \( \bar{Q} = P \), and finally the portfolio choice problem is as follow,
\[
\max_{\tilde{q}_i} V(r_{q,i}) = E_P(r_{q,i}|\epsilon_i) - \frac{\tilde{\lambda}}{2} Var_P(r_{w,i}|\epsilon_i) - \frac{\theta}{2} Var_\mu[E_Q(r_{w,i}|\epsilon_i)] \\
\text{s.t. } q_i \geq 0
\]
(41)
(42)

The solution of optimal stock holding is
\[
\tilde{q}_i = \frac{E_P(r_e) + \epsilon_i - r_f}{\tilde{\lambda} Var_P(r_e) + \theta Var_\mu[E_Q(r_e)]}
\]
(43)

if \( E_P(r_e) + \epsilon_i - r_f > 0 \); and \( \tilde{q}_i = 0 \) otherwise.

The robust mean-variance preference has a great advantage to admit a richer comparative statics. Notice that, clearly, assuming the same risk attitude (i.e. \( \tilde{\lambda} = \lambda \)), if the agent is ambiguity averse (\( \theta > 0 \)) and the information about return is ambiguous (\( Var_\mu[E_Q(r_e)] > 0 \)),

---

1 Here, using \( \tilde{\cdot} \) is to distinguish the parameters in robust mean-variance model and mean-variance model.

2 Assume the \( \sigma \)-algebra endowing the probability measure is the power set \( 2^\Delta \).
the agent must allocate less proportion of wealth in stock than mean-variance utility maximizer. Indeed, it is a special case of comparative statics for changing $\theta$ while fixing $\lambda$, $\text{Var}_p(r_e)$, and positive $\text{Var}_\mu[\text{E}_Q(r_e)]$.

**Proposition 1.** Consider the optimal asset allocation equation (43) derived from the robust mean-variance portfolio choice problem with the fixed participation cost in this chapter, the followings hold

- Fix $\lambda$, $\text{Var}_p(r_e)$ and positive $\text{Var}_\mu[\text{E}_Q(r_e)]$. If the DM is more ambiguity averse, i.e. larger $\theta$, then the optimal stock holding $\tilde{q}_i$ will be lower;
- Similarly, fix $\lambda$, $\text{Var}_p(r_e)$ and positive $\theta$, i.e. the DM is ambiguity averse. If the DM has a poorer information on the expected stock return, i.e. greater $\text{Var}_\mu[\text{E}_Q(r_e)]$, then the optimal stock holding $\tilde{q}_i$ will be lower.

In addition, with the same participation cost, i.e. setting $\tilde{c} = c$ and thus $\bar{\kappa}_i = \kappa_i$, ambiguity averse agent is more likely to keep herself away from stock market. With the solution above, we can easily calculate the maximum utility can be attained with robust mean-variance utility is

$$V(r_{q_i}, i) = r_f + \bar{s}_i - \bar{k}_i$$

where $\bar{s}_i = \frac{[\text{E}(r_e) + c - r_f]^2}{2(\lambda \text{Var}_p(r_e) + \theta \text{Var}_\mu[\text{E}_Q(r_e)])}$. Again, the decision is made not to hold stock when $\bar{s}_i - \bar{k}_i < 0$, which, under no short-selling condition, is equivalent to

$$\epsilon_i < -[\text{E}(r_e) - r_f] + \sqrt{2\bar{k}_i (\lambda \text{Var}_p(r_e) + \theta \text{Var}_\mu[\text{E}_Q(r_e)])}$$

That is to say, for a generic household $i$, the probability of NOT holding stock is $\Phi(N_i^A / \sigma_e)$ where $\Phi(\cdot)$ is the standard normal distribution function and with

$$N_i^A = -[\text{E}(r_e) - r_f] + \sqrt{2\bar{k}_i (\lambda \text{Var}_p(r_e) + \theta \text{Var}_\mu[\text{E}_Q(r_e)])}$$

Consider that, in the mean-variance framework, the condition for not holding stock is

$$\epsilon_i < -[\text{E}(r_e) - r_f] + \sqrt{2\bar{k}_i \lambda \text{Var}_p(r_e)}$$

and therefore the probability of NOT holding stock for a mean-variance DM is $\Phi(N_i^R / \sigma_e)$ where

$$N_i^R = -[\text{E}(r_e) - r_f] + \sqrt{2\bar{k}_i \lambda \text{Var}_p(r_e)}$$

By assumptions, $N_i^A$ is greater than $N_i^R$, and hence the inequality

$$\Phi\left(\frac{N_i^A}{\sigma_e}\right) > \Phi\left(\frac{N_i^R}{\sigma_e}\right)$$

3 Similar argument as before applied for that $\lambda$ and $\theta$ are approximately equal to the constant relative risk aversion coefficient and relative ambiguity aversion coefficient, respectively, if assuming the agent is CRAA and has a CRRA $u(\cdot)$.  

holds which means that ambiguity averse agent is more likely to hold zero stock compared to mean-variance agent with same risk attitude and participation cost.

Indeed, to carry out the comparative static analysis exercises, simply look at the changes in $N_A^i$: greater $N_A^i$ leads to higher probability of not holding stock, and vice versa. Hence, the following results can be easily seen:

**Proposition 2.** Consider the probability of holding stock for household $i$, $\Phi(N_A^i / \sigma_e)$, the followings hold

- Fix $\bar{\lambda}$, $\text{Var}_P(\epsilon)$, positive $\bar{\kappa}_i$ and postie $\theta$. If the DM is more ambiguity averse, i.e. ambiguity aversion coefficient $\bar{\lambda}$ is greater, then she is less likely to hold the ambiguous asset, i.e. stock in this case;
- Fix $\bar{\lambda}$, $\text{Var}_P(\epsilon)$, positive $\bar{\kappa}_i$ and postie $\theta$. If the DM has poorer information, i.e. greater $\text{Var}_\mu[E_Q(\epsilon)]$, then she is less likely to hold stock.

Recall the normalized participation cost $\bar{\kappa}_i = \bar{\kappa} / w_i$ is also interpreted as the cost-to-wealth ratio, we can further easily see the following results.

**Proposition 3.** Again consider the probability of holding stock for household $i$, $\Phi(N_A^i / \sigma_e)$, the followings hold

- Fix $\bar{\lambda}$, $\text{Var}_P(\epsilon)$, $\theta$, $\text{Var}_\mu[E_Q(\epsilon)]$ and positive household wealth $w_i$. A higher participation cost, i.e. a greater $c$, will be more likely to keep the household away from holding stock;
- Fix $\bar{\lambda}$, $\text{Var}_P(\epsilon)$, $\theta$, $\text{Var}_\mu[E_Q(\epsilon)]$ and positive participation cost $c$. A wealthier household, i.e. household with greater $w_i$, is more likely to hold stock.

### 4.2 Estimation

The estimation of this model is in line with the discussion in the previous chapter, but in this case, we have

$$\bar{s}_i - \bar{\kappa}_i > 0 \iff$$

$$\frac{E_P(\epsilon) - r_f}{\sigma_e} - \sqrt{2(\bar{\lambda}\text{Var}_P(\epsilon) + \theta\text{Var}_\mu[E_Q(\epsilon)])} \left[ \frac{\bar{c}}{\sigma_e^2} \sqrt{\frac{1}{w_i} + \frac{\epsilon_i}{\sigma_e}} \right] > 0. \quad (50)$$

Employing the probit model in (37), we can still recover the $\sigma_e$ with the same formula in (37), but it is impossible to recover either $\bar{\lambda}$ or $\theta$ alone without sufficient knowledge. A trick proposed here is introducing the ambiguity-to-risk ratio, $m = \frac{\theta \text{Var}_\mu[E_Q(\epsilon)]}{\bar{\lambda}\text{Var}_P(\epsilon)}$. Then, similar to (3.20), we can recover

$$\hat{\lambda} \hat{c} = \frac{\hat{\beta}^2 \sigma_e^2}{2(1 + m)\text{Var}_P(\epsilon)}.$$  \quad (51)
as long as we specify some value of $m$.

Moreover, we can, again, use Heckman selection model to estimate relative risk aversion coefficient $\hat{\lambda}$ and recover the participation cost $c$ as well. However, fortunately, some calculations shows that we can further exploit the estimates in chapter 3, which makes the recovery of parameters simply as follow (See Appendix A.2):

$$\hat{\lambda} = \frac{1}{1 + m} \hat{\lambda}$$

$$\hat{c} = \hat{c}$$

(52)

(53)

where $\hat{\lambda}$ and $\hat{c}$ are the estimates under mean-variance model. We can clearly discern that if $m = 0$ it is equivalent to say there is no ambiguity or ambiguity aversion, and the model is reduced to mean-variance utility agent model and the estimates coincide.

With exploitation previous studies in risk aversion, we can specify some value of $m$ to calibrate estimates of relative risk aversion coefficient within the range consistent with other studies. Then we can only identify the product of ambiguity and ambiguity aversion, i.e. $\theta Var_{\mu}[E_{\mu}(r_e)]$.

However, only a few research are talking about $\theta$. One example is that Ju and Miao (2012) calibrate the relative ambiguity aversion coefficient to match the mean equity premium, which give a $\theta$ roughly equal to 7. $^4$ Then we can further reveal the value of $Var_{\mu}[E_{\mu}(r_e)]$.

The Table 2 reports the estimates in year 2004 and 2007, respectively. Two observations need to be noticed: to get a reasonable relative risk aversion coefficient $\lambda$ around 2, we should have the ambiguity-to-risk ratio $m = 1.6$ in year 2007 and $m = 1$ in 2004, which means than the ambiguity part is substantial, equal to or greater than the risk part (the product of risk aversion and risk); in addition, since we know that the economic situation becomes turbulent in 2007, given the same parameter of preferences, the measurement of uncertainty becomes bigger. For example, both set $\lambda = 2$ and $\theta = 7$, the estimates of $Var_{\mu} [E_{\mu}(r_e)]$ is 1.14% and 1.83%, respectively in 2004 and 2007, where the latter is over 60% higher than the former.

$^4$ Recall that we set $\phi = \nu \circ u^{-1}$. Indeed, in Ju and Miao’s paper, they assume that $\nu = \frac{x^{1-\lambda}}{1-\lambda}$ and $u = \frac{x^{1-\lambda}}{1-\lambda}$, which is aligned with the assumption of CRAA and CRRA in this paper, and calibrate that $\lambda = 2$ and $\eta = 8.864$. Notice that in this case, as discussed before, $\theta = \eta - \lambda = 6.864$. 

Table 2: Estimates of robust mean-variance model

<table>
<thead>
<tr>
<th></th>
<th>2004</th>
<th>2007</th>
</tr>
</thead>
<tbody>
<tr>
<td>(m)</td>
<td>(\hat{\lambda})</td>
<td>(\hat{c})</td>
</tr>
<tr>
<td>0.5</td>
<td>2.67</td>
<td>275</td>
</tr>
<tr>
<td>1</td>
<td>2</td>
<td>275</td>
</tr>
<tr>
<td>2</td>
<td>1.33</td>
<td>275</td>
</tr>
</tbody>
</table>

the ambiguity aversion \(\theta\) is set equal to 7;

\(m\): the ambiguity-to-risk ratio;

\(\hat{\lambda}\): the estimates of risk aversion;

\(\hat{c}\): the estimates of the cost.
At the moment, it is very difficult to go further to decompose the quantity of ambiguity and ambiguity aversion since we have limited knowledge of either one. In the previous chapter, I fix the ambiguity aversion coefficient to reveal the ambiguity, but let us consider another way around in this chapter. Let us first study some significant part of the ambiguity and thus construct some proxy and then recover the corresponding ambiguity aversion attitude. Also, there will be a discussion on how the ambiguity in participation cost itself alter the investment decision since the uncertainty of participation cost is not resolved when the agent makes the decision whether to hold stock or not.

5.1 Simple Linear Model of Time Varying Stock Return

The smooth model works as if the agent has a set of models for the risky prospect, so a simple way to construct such scenario is just assuming the agent has a linear model for the expected return, which depends on some time varying observed variables (e.g. price-earning ratio); and this leads to a time varying expected return as well.

Let us assume that the agent has the following simple regression model on the annual stock return:

\[ r_{e,t} = \delta_0 + \delta_1 (PE \ ratio)_t + \iota_t \]  

where \( \iota_t \) is i.i.d with \( E(\iota_t) = 0 \), and \( Var(\iota_t) = \sigma^2_\iota \). The expected return

\[ E_p(r_{e,t}) = \hat{r}_{e,t} = \hat{\delta}_0 + \hat{\delta}_1 (PE \ ratio)_t \]  

where \( \hat{\delta}_0 \) and \( \hat{\delta}_1 \) are OLS regression coefficients. In this case, the ambiguity is characterized by the DM’s uncertainty on the estimated regression coefficients.

The equity data is extracted from Robert Shiller’s data on historical equity return, which also contains the PE ratios, and 20 observations before the year when portfolio decision is made is used to do the estimation and regression. First, \( Var_p(r_e) \) is estimated by using historical data. Then we can reproduce the procedure of estimation with robust mean-variance preferences. Moreover, the let us specify a ambiguity-to-risk ratio, \( m \), such that it fixes \( \lambda = 2 \). Since \( Var_p[E_Q(r_e)] \) can be estimated by using the prediction of the linear model, that is, \( Var_p[E_Q(r_e)] = Var(\hat{r}_e) \), then the ambiguity aversion coefficient, \( \theta \), is consequently recovered.
Table 3: Estimates with simple linear models of return

<table>
<thead>
<tr>
<th></th>
<th>$\hat{\delta}_0$</th>
<th>$\hat{\delta}_1$</th>
<th>$E_P(r_{e,t}) - r_{f,t}$</th>
<th>$\text{Var}_P(r_e)$</th>
<th>$\text{Var}_\mu[E_Q(r_e)]$</th>
<th>$m$</th>
<th>$\hat{\theta}$</th>
<th>$\hat{c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>0.29</td>
<td>-0.0088</td>
<td>0.08</td>
<td>0.0256</td>
<td>0.0064</td>
<td>3.15</td>
<td>25</td>
<td>230</td>
</tr>
<tr>
<td>2007</td>
<td>0.31</td>
<td>-0.0077</td>
<td>0.087</td>
<td>0.0225</td>
<td>0.0066</td>
<td>5.7</td>
<td>39</td>
<td>260</td>
</tr>
</tbody>
</table>

The relative risk aversion coefficient $\hat{\lambda}$ is set equal to 2;
$\hat{\delta}_0, \hat{\delta}_1$: regression coefficients of the expected return equation;
$m$: the ambiguity-to-risk ratio;
$\hat{\theta}$: the estimates of ambiguity aversion;
$\hat{c}$: the estimates of the cost.

The Table 3 reports the results in 2004 and 2007, where it shows that there is a very high relative ambiguity aversion coefficient. We may take the stand arguing that ambiguity might be small and ambiguity aversion is great. However, it may the other case. The ambiguity aversion could be moderate, but the ambiguity constructed in this way is too small and could not capture the significant part of it.

5.2 AMBIGUOUS IDIOSYNCRATIC INFORMATION

To further explore the possible value of ambiguity, one possible source of ambiguity may lie in the idiosyncratic information. To introduce the ambiguity of idiosyncratic information, set

$$\tilde{\epsilon}_i = \epsilon_i + u_i$$

such that $E_\mu(u_i|\epsilon_i) = 0$ and $u_i \in M^\perp = \{h \in L^2 : Var_\mu(E(h)) = 0\}^\perp$. That is to say, the idiosyncratic information is noisy and can be decomposed to a certain part $\epsilon_i$ (it is known for household $i$) and an ambiguous part $u_i$.

In this case, $r_{e,i} = r_e + \tilde{\epsilon}_i$

$$\text{Var}_\mu[E(r_{e,i})] = \text{Var}_\mu[E_Q(r_e)] + \text{Var}_\mu[E_Q(\tilde{\epsilon}_i)\epsilon_i]$$

(57)

Furthermore, let us assume that the DM makes no mistake of estimating the variance of the idiosyncratic information $\sigma_\epsilon^2$, which is just equal to $\text{Var}_\mu[u_i]$. Exploiting the estimates aforementioned,

$$\hat{\theta} = \frac{m\hat{\lambda}\text{Var}_P(r_e)}{\text{Var}_\mu[E_Q(r_{e,i})]} = \frac{m\hat{\lambda}\text{Var}_P(r_e)}{\text{Var}_\mu[E_Q(r_e)] + \text{Var}_\mu[E_Q(\tilde{\epsilon}_i|I_i)]} = \frac{m\hat{\lambda}\text{Var}_P(r_e)}{\text{Var}_\mu[E_Q(r_e)] + \hat{\sigma}_\epsilon^2}$$

(58)

Again, assuming the agent uses the linear models as last section, and the idiosyncratic information is also purely ambiguous, using the estimates in the previous section, the results of re-estimation on $\hat{\theta}$ is reported in the following table. The estimates of $\theta$ are improved...
The relative risk aversion coefficient $\hat{\lambda}$ is set equal to 2; 
$\hat{\delta}_0, \hat{\delta}_1$: regression coefficients of the expected return equation;
$m$: the ambiguity-to-risk ratio;
$\hat{\theta}$: the estimates of ambiguity aversion;
$\hat{c}$: the estimates of the cost.

greatly in the sense that they become a moderate value and much more close to the calibrated one in Ju and Miao’s paper. However, by my conjecture, even under this construction, it could not cover the sources of ambiguity exhaustively, but a significant part of it.

### 5.3 Ambiguity in Participation Cost

Now let us consider a setting where the participation cost $\tilde{c}$ is ambiguous. In fact, before deciding purchase stock or not, the uncertainty of participation cost is not resolved. Again each DM $i$ has a set of probabilistic model, $\Delta$, which contains the joint distributions for both return and participation cost. For each probabilistic model, $\forall Q \in \Delta$, we have that $\tilde{c}$ is normally distributed, and

\[
E_Q(\tilde{c}) = \bar{c} + \xi_Q
\]

\[
\text{Var}_Q(\tilde{c}) = \sigma_c^2
\]

where $\xi_Q$ is such that

\[
E_{\mu}(\xi_Q) = 0
\]

\[
\text{Var}_{\mu}(\xi_Q) = \sigma_a^2
\]

Recall that $\mu$ be the probabilistic measure of $\Delta$, and set $P = \bar{Q} = \int_{\Delta} Q d\mu$. Then

\[
E_P(\tilde{c}) = E_{\mu}(E_Q(\tilde{c}) = E_{\mu}[\bar{c} + \xi_Q] = \bar{c}
\]

\[
\text{Var}_{\mu}[E_Q(\tilde{c})] = \text{Var}_{\mu}[\bar{c} + \xi_Q] = \text{Var}_{\mu}(\xi_Q) = \sigma_a^2
\]

\[
\text{Var}_P(\tilde{c}) = E_{\mu}[\text{Var}_Q(\tilde{c})] + \text{Var}_{\mu}[E_Q(\tilde{c})] = \sigma_c^2 + \sigma_a^2
\]

Consequently, the end-of-period return is re-formulated in the following way

\[
r_{w,i} = r_f + w(r_e,i - r_f) - \hat{\kappa}_i
\]

where $\hat{\kappa}_i = \tilde{c}/w_i$. 

---

**Table 4: Estimates with ambiguous idiosyncratic information**

<table>
<thead>
<tr>
<th>Year</th>
<th>$\hat{\delta}_0$</th>
<th>$\hat{\delta}_1$</th>
<th>$E_P(r_{e,i}) - r_f$</th>
<th>$\text{Var}_P(r_e)$</th>
<th>$\text{Var}_P[E_Q(r_e)]$</th>
<th>$\hat{\sigma}_e$</th>
<th>$m$</th>
<th>$\hat{\theta}$</th>
<th>$\hat{c}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>0.29</td>
<td>-0.0088</td>
<td>0.08</td>
<td>0.0256</td>
<td>0.0064</td>
<td>0.090</td>
<td>3.15</td>
<td>11</td>
<td>230</td>
</tr>
<tr>
<td>2007</td>
<td>0.31</td>
<td>-0.0077</td>
<td>0.08</td>
<td>0.0225</td>
<td>0.0066</td>
<td>0.154</td>
<td>5.7</td>
<td>8.5</td>
<td>260</td>
</tr>
</tbody>
</table>
In this case, since the participation cost, again, has nothing to do with the stock return, so it does not affect the optimal holding of stocks, so we still have

\[ \hat{q}_i = \frac{E(r_n|e_i) - r_f}{\lambda \text{Var}(r_n) + \theta \text{Var}_n[E_Q(r_n)]} \]  

But it does affect the maximum value could be achieved:

\[ U(r_{\hat{q},i}) = r_f + \tilde{s}_i - \tilde{k}_i - \gamma_0 \frac{1}{w_i} \]  

where \( \tilde{s}_i = \frac{1}{2} \frac{[E_P(r_n|e_i) - r_f]^2}{\lambda \text{Var}_n(r_n) + \theta \text{Var}_n[E_Q(r_n)]} \) and \( \gamma_0 = \frac{\lambda (\sigma^2 + \sigma_a^2) + \theta \sigma_a^2}{2} \).

Two new features under ambiguous participation cost compared the situation with fixed participation cost: 1) the expected ambiguous participation cost, which is needed to rule out the agent from the stock market, is lower than the fixed one, ceteris paribus; 2) the wealthier DMs could bear a higher expected participation cost than poorer one to hold stock.

The first point is clear. Recall, the value of holding stock with fixed participation cost:

\[ V(r_{\hat{q},i}) = r_f + \tilde{s}_i - \tilde{k}_i \]  

To make the same decision (whether to buy stocks) as before, one only need to compare \( \tilde{k}_i \) and \( \tilde{k}_e + \gamma_0 \frac{1}{w_i} \). Equate the two, we get

\[ \tilde{c} = \tilde{c} + \gamma_0 \frac{1}{w_i} \]  

Assume that agent \( i \) is at the threshold, then the fixed participation cost to prevent him from buying stock is exactly \( \tilde{c} \); however, the expected (ambiguous) participation cost to prevent him from buying stock is \( \tilde{c} = \tilde{c} + \gamma_0 \frac{1}{w_i} < \tilde{c} \).

The second point is also followed by the fact that, if now her wealth is changed to \( w_i' > w_i \), then \( \tilde{c} = \tilde{c} + \gamma_0 \frac{1}{w_i'} > \tilde{c} + \gamma_0 \frac{1}{w_i} \), the cost is lowered, and one will choose to buy stock. Empirically, the ambiguity of participation cost could be further useful to improve the estimates of \( \theta \) if itself could be estimated, but it can be very small since it is a second order effect (through the second moments).
EXTENDED DISCUSSION

In this chapter, first, consider a change in the environment, for instance, the observed asset price (or other signal), and how will the DM react to such change? This question is interesting because it will uncover some dynamics of portfolio formation: the DM learns from the changing observables and update her information by some rule, which, in turn, affects her expectation and evaluation of information quality, and re-balance the portfolio according to the changes in price.

Secondly, I extend the very basic model to a multiple asset setting. This extension will allow us to analyze the problems concerning comparison between two or more different types of assets. and thus, it will be useful to deepen our understanding why some asset is more preferred to others despite comparable or even more attractive return and risk profile.

6.1 CHANGES IN ASSET PRICE AND PORTFOLIO

One interesting question is raised here: how will the changes in observed asset price dynamics (or other signal) affects the stock holding? This question is not only interesting but also important. The existence of momentum trading in stock market is being noticed. To answer this question, it is more appropriate to incorporate a learning structure of DM in a dynamic setting (See Epstein and Schneider, 2007). However, a qualitative analysis here may also offer some insights to understand this question as a first step.

Recall that the optimal holding of stock with robust mean-variance preference is given by the equation:

\[
\tilde{q} = \frac{EP(r_e) - rf}{\tilde{\lambda} VarP(r_e) + \theta Var_p [E_Q(r_e) - rf] + \theta Var_Q [E_Q(r_e)]} = \frac{EP(r_e) - rf}{Z}
\]

where \(Z = \tilde{\lambda} VarP(r_e) + \theta Var_p [E_Q(r_e)]\). As mentioned before, to study the influence of changes in price, it is necessary to incorporate a learning structure; that is to say, the expected return and the variance of expected return depends on the observed price:\n
\[EP(r_e) - rf = f(p)\]

---

1 To simplify the analysis, drop the individual subscripts.

2 In this setting, I want to focus on the analysis of the effect of ambiguity part, so leave the risk part irrelevant to the change in asset price for simplicity; and one of the justifications is that the estimates of \(VarP(r_e)\) are resulted from the long-term historical data and thus a few new observations or signals can barely affect it.
and $Z = \tilde{\lambda} \text{Var}_p(r_e) + \theta g(p)$ where $p$ is the price of the asset. Then, assuming differentiability of the functions, we derive that

$$\frac{d\tilde{q}}{dp} = \frac{f'(p)Z - f(p)\theta g'(p)}{Z^2}. \quad (72)$$

There are several situations such that $\frac{d\tilde{q}}{dp} > 0$, which mean the increase in stock price leads to an increase in optimal holding of stock, if:

- $f'(p) > 0$ and $g'(p) < 0$. This says a situation such that, when observing the increase in stock price, the DM has a higher expected equity premium and lower ambiguity;

- $f'(p) > 0$ and $0 < g'(p) < \frac{f'(p)}{\theta f'(p)} Z$. This is situation where DM’s expected equity premium increases when observing the increase in stock price; and even though the ambiguity increases, its increment is small enough;

- $f'(p) < 0$ and $g'(p) < \frac{f'(p)}{\theta f'(p)} Z < 0$. In this case, the DM’s expected equity premium decrease when observing the increase in stock price; however, at the same time, the ambiguity also decrease, and magnitude of its decrease should be large enough.

Now, consider the decision on whether to hold stock, and also recall the probability of not holding stock is $\Phi(N_A/\sigma_e)$, where $N_A = -f(p) + \sqrt{2\tilde{\kappa}_i} Z$. If $N_A$ becomes smaller, $\Phi(N_A/\sigma_e)$ will become smaller, i.e. more likely to hold stock. So it suffices to analyze the effects of change in stock price $p$ on $N_A$. We have

$$\frac{dN_A}{dp} = -f'(p) + \theta g'(p) \sqrt{\frac{\tilde{\kappa}_i}{2Z}}. \quad (73)$$

Similar to the analysis beforehand, there are some cases where $\frac{dN_A}{dp} < 0$, and thus the increase in stock price leads it more likely to hold stock:

- $f'(p) > 0$ and $g'(p) < 0$. This is exactly the same first situation as before such that, when observing the increase in stock price, the DM has a higher expected equity premium and lower ambiguity;

- $f'(p) > 0$ and $0 < g'(p) < \frac{f'(p)}{\theta} \sqrt{\frac{2Z}{\tilde{\kappa}_i}}$;

- $f'(p) < 0$ and $g'(p) < \frac{f'(p)}{\theta} \sqrt{\frac{2Z}{\tilde{\kappa}_i}} < 0$.

The last two conditions are also similar as before, but only the threshold changes.
6.2 MULTIPLE-ASSET FRAMEWORK

The framework can be easily extended to the case of multiple (more than two) ambiguous assets. The case of multiple assets will offer a more general view of why some asset is more preferred, while others are not despite of comparable or even more attractive return and risk profile. As a motivation, consider some well-documented empirical puzzles, such as, particularly, home bias among others, are not only concerning a lower observed portfolio weight of unfamiliar assets than that of predicted by theory, but also, in fact, many individual investors choose not to hold them.

6.2.1 A general multiple asset framework

Now, the DM can choose to allocate her (unit) wealth into \( n + 1 \) assets. The gross return of asset \( k \), \( r_k \), is square integrable, for each \( k = 1, \ldots, n \). Then the return vector \( \mathbf{r} \) is an \((n \times 1)\) vector of returns on first \( n \) assets, and \( \mathbf{q} \) is the \((n \times 1)\) vector of portfolio weights. And the \((n + 1)\)-th assets is a risk free asset denoted by \( r_f \). For first \( n \) assets, holding each of them is related to a cost \( \kappa_k \), and the \((n \times 1)\) vector \( \mathbf{\kappa} \) is the cost vector.

Then, end-of-period return is

\[
\mathbf{r}_q = r_f + \mathbf{q}^T (\mathbf{r} - r_f \mathbf{1}) - \mathbf{\kappa}^T \mathbf{1}_q
\]

(74)

where \( \mathbf{1} \) is a \((n \times 1)\) unit vector and \( \mathbf{1}_q \) is an indicator vector whose \( j \)-th entry equal to 1 if the \( j \)-th entry of \( \mathbf{q} \) is positive and equal to 0 otherwise. The maximization problem is as usual but with robust mean-variance preference:

\[
\max_{\mathbf{q}} U(q) = E_p(r_q) - \frac{1}{2} \lambda \text{Var}_p(r_q) - \frac{1}{2} \theta \text{Var}_\mu[E_Q(r_q)]
\]

(75)

s.t. \( \mathbf{q} \geq 0 \)

(76)

To directly compute the solution of the above problem is a bit cumbersome because \( \mathbf{1}_q \) also depends on the choice of \( \mathbf{q} \), which makes the function not differentiable.

To solve it, we have to reformulate it into an equivalent problem as follow. Let \( A = \{1, 2, \ldots, n\} \) be the set of first \( n \) assets, and \( \emptyset \) represent risk-free asset. Then, \( \mathcal{E} = \{\emptyset\} \cup B : B \in \mathcal{P}(A) \} \), where \( \mathcal{P}(A) \) is the power set of set \( A \), is the class of all the possible asset combination with positive portfolio weights.

First, fix an asset combination \( E \in \mathcal{E} \) containing \( l \) assets, and denote \( \mathbf{r}_E \) the \((l \times 1)\) return vector of those chosen asset, and \( \mathbf{q}_E \) the \((l \times 1)\) portfolio weight vector associate a positive weight to each asset in \( E \). Then, the end-of-period return is

\[
\mathbf{r}_{q_E} = r_f + \mathbf{q}_E^T (\mathbf{r}_E - r_f \mathbf{1}_l) - \mathbf{\kappa}_E
\]

(77)
where $\mathbf{1}_l$ is the $(l \times 1)$ unit vector and $\kappa_E = \kappa^T \mathbf{1}_l = \sum_{k \in E} \kappa_k$.

Notice that when $E$ is fixed, choice of weights $q_E$ is no long affecting the total participation cost, and thus, it becomes much easier to solve the following maximization problem

$$\max_{q_E} U(r_{q_E}) = E_p(r_{q_E}) - \frac{1}{2} \lambda \text{Var}_P(r_{q_E}) - \frac{1}{2} \theta \text{Var}_\mu[E_Q(r_{q_E})].$$  \hspace{1cm} (78)

Since the cost does not affect the return vector itself and not relevant to the choice of $q_E$, the solution can be simply derived as

$$\hat{q}_E = \left[ \lambda \text{Var}_P(r_E) + \theta \text{Var}_\mu[E_Q(r_E)] \right]^{-1} E_p [r_E - r_f \mathbf{1}_l],$$ \hspace{1cm} (79)

where $\text{Var}_P(r_E)$ is the variance-covariance matrix of $r_E$ under $P$, $\text{Var}_\mu[E_Q(r_E)]$ is the variance-covariance matrix of $E_Q(r_E)$, expected returns of $r_E$ under $\mu$, and finally $E_p [r_E - r_f \mathbf{1}_l]$ is the vector of expected excess returns under $P$. Then denote $V_E = \max_{q_E} U(r_{q_E}) = U(r_{\hat{q}_E})$. Then the DM chooses the maximum of $V_E$ such than $E \in \mathcal{E}$, and denote

$$\hat{E} = \arg \max_{E \in \mathcal{E}} V_E$$ \hspace{1cm} (80)

as the choice correspondence, which is called optimal combination(s) of assets.\(^3\)

Finally, then optimal portfolio weight is just $\hat{q}_E$, which determines the optimal chosen in types of assets and optimal weights for each chosen asset. Set the $k$-th entry of $\hat{q}$ equal to the corresponding value in $\hat{q}_E$ for those chosen assets, and 0 for those not included. A simple counter-positive argument will show that $\hat{q}$ solves the problem and it is the unique solution to the problem if it admits only one solution.\(^4\)

### 6.2.2 A note on an example of two ambiguous assets

When increasing the dimension of types of assets, it becomes much difficult to give analytical results. However, in the case of only two ambiguous assets, it is still manageable in some sense. One of the most interesting case is that if , besides the risk-free asset, there are another two assets, 1 and 2. The asset 2 may be more attractive in the sense that it has a higher shape ratio (after reduction of the participation cost), i.e. $\frac{E_p(r_k) - r_f - \kappa_k}{\text{Var}_P(r_k)}$, $k = 1, 2$. Under mean-variance preference, in a general case, the population of holding asset 2 and portfolio weight of asset 2 should be higher, however, this is not always the case consistent with real financial data.

The home bias might be such a case: in emerging market, it is possible to get a higher expected return for its potentially high growth, but foreign investors usually will not hold an enough high proportion of

---

\(^3\) In fact, $\hat{E} : \mathcal{E} \rightrightarrows \mathcal{E}$ is generally a correspondence. The existence of maximum is guaranteed by the finiteness of $\mathcal{E}$, but the uniqueness is not guaranteed.

\(^4\) Proof is omitted.
such stock in their portfolio. For other motive cases, consider in the same stock market (i.e. similar cost for holding stock), the blue-chip stocks are usually well-analyzed and attract many investors, while some small companies growing fast but less analyzed will attract less investors even though its stock returns are higher. These observations will be consistent under robust mean-variance preference if certain conditions hold.

Consider the following scenario: there are three assets \( o \) represent the risk-free asset with return \( r_f \), and another two ambiguous assets, \( 1 \) and \( 2 \), whose return are \( r_1 \) and \( r_2 \), respectively. Apply the general framework in this case, we will have:

\[
\begin{align*}
V_0 &= r_f \\
V_1 &= r_f + s_1 - \kappa_1 \\
V_2 &= r_f + s_2 - \kappa_2 \\
V_{1,2} &= r_f + s_{1,2} - (\kappa_1 + \kappa_2)
\end{align*}
\]

where:

- \( V_0 \) is the return for investing risk-free asset only; \( V_1 \) and \( V_2 \) is the maximum return for investing in risk-free asset and asset \( 1 \) or asset \( 2 \), respectively; and \( V_{1,2} \) is the maximum return for investing in all three assets;

- \( s_1 = \frac{1}{2} \left[ \frac{E_P(r_1) - r_f}{\lambda Var_P(r_1) + \theta Var_p[E_Q(r_1)]} \right]^2 \) and \( s_2 = \frac{1}{2} \left[ \frac{E_P(r_2) - r_f}{\lambda Var_P(r_2) + \theta Var_p[E_Q(r_2)]} \right]^2 \) is the excess utility from investing in asset \( 1 \) and \( 2 \), respectively; and

\[
s_{1,2} = \frac{1}{2} \left[ E_P(r) - r_f \right]^T \left\{ \lambda Var_P(r) + \theta Var_p[E_Q(r)] \right\}^{-1} \left[ E_P(r) - r_f \right]
\]

is the excess utility for holding both asset \( 1 \) and \( 2 \).

Consider two necessary conditions for the fact that the DM finally only invest risk-free asset and asset \( 1 \), but not asset \( 2 \):

\[
\begin{align*}
V_1 &> V_2 \\
V_1 &> V_{1,2}
\end{align*}
\]

The first condition \( V_1 > V_2 \), which says the maximum utility of investing in risk-free asset and asset \( 1 \) alone is higher than that of investing in risk-free asset and asset \( 2 \) alone, is equivalent to

\[
s_1 - \kappa_1 > s_2 - \kappa_2
\]

which is a comparison between two excess utilities. Then, there could be some possibilities such as

- the inequality will hold if it is too costly to invest in asset \( 2 \);
• fix the parameters for taste, $\tilde{\lambda}$ and $\theta$, and suppose asset 2 has a weakly higher return, i.e. $E_P(r_2) \geq E_P(r_1)$, and its volatility is weakly less than asset 1, i.e. $Var_P(r_2) \leq Var_P(r_1)$ and cost of holding asset 2 is weakly lower than that of asset 1, i.e. $\kappa_2 \leq \kappa_1$, then the inequality will hold only if asset 2 is more ambiguous than asset, i.e. $Var[\mu[E_Q(r_2)]] > Var[\mu[E_Q(r_1)]]$. This result says that even though asset 2 is better than asset 1 in every aspect without considering ambiguity, the enough high ambiguity of asset 2 will finally rule out the possibility of holding it alone (along with risk-free asset).

The second condition $V_1 > V_{1,2}$, which says the maximum utility of investing in risk-free asset and asset 1 alone is higher than that of investing in all three assets, is equivalent to

$$s_1 > s_{1,2} - \kappa_2$$  \hspace{1cm} (89)

which is again a comparison between two excess utilities. The full analysis will depend on a comprehensive discussion on the cost, expected return, the variance-covariance matrix of the asset return, and so forth, which will be too long to accommodate in this paper. Let us focus on only an interesting but general case.

Suppose, again, that asset 2 has a weakly higher return, i.e. $E_P(r_2) \geq E_P(r_1)$, and its volatility is weakly less than asset 1, i.e. $Var_P(r_2) \leq Var_P(r_1)$ and cost of holding asset 2 is also equal to that of asset 1, i.e. $\kappa_2 = \kappa_1 = \kappa$, and $V_1 > V_2 > V_0$, which implies that

$$Var[\mu[E_Q(r_1)]] < Var[\mu[E_Q(r_2)]] < \left[\frac{E_P(r_2) - r_f}{\theta \kappa}\right]^2 - \frac{\lambda}{\theta} Var_P(r_2) \hspace{1cm} (90)$$

which imposes a lower bound and upper bound for the ambiguity of asset 2. In this case, risk-free asset and asset 2 will never be an optimal combination in the presence of asset 1 since it is more ambiguous compared than asset 1, even though it has a (weakly) higher Sharp ratio and (weakly) lower cost. But if there is no asset 1, then asset 2 must be included in optimal combination. The following questions are raised consequently: is it possible to invest all three assets as an optimal under some circumstance? And what’s the necessary condition? And is there some condition that ensures risk-free asset and asset 1 be the only optimal combination?

The first question is easy to answer: consider a situation where there happened to be the case that

$$\tilde{\lambda} Cov_P(r_1, r_2) + \theta Cov[\mu[E_Q(r_1), E_Q(r_2)]] = 0 \hspace{1cm} (91)$$

then we will have

$$s_{1,2} = s_1 + s_2 \hspace{1cm} (92)$$
By the assumption \( V_1 > V_2 > V_0 \),

\[
V_{1,2} = r_f + s_1 + s_2 - \kappa
\]
\[
= V_1 + s_2 - \kappa_2 + \kappa
\]
\[
= V_1 + (V_2 - V_0) + \kappa_1 > V_1
\]  \hspace{1cm} (93)

that is, the optimal combination is holding all three assets, i.e. \( \hat{E} = \{0, 1, 2\} \).

In fact, one can show that there exits a threshold

\[
C^* = \frac{\hat{r}_1 \hat{r}_2 - \sqrt{AB[k_2^2 + (s_1 - s_2)\kappa_2]} + s_1 + \kappa}{s_1 + \kappa} > 0
\]  \hspace{1cm} (94)

such that

\[
\tilde{\lambda}\text{Cov}_P(r_1, r_2) + \theta \text{Cov}_\mu[E_Q(r_1), E_Q(r_2)] \leq C^*
\]  \hspace{1cm} (95)

is a necessary condition for \( V_{1,2} \geq V_1 \) (See the proof in Appendix, A.3). This makes clear that the above case where \( C = 0 \) is just a special case that satisfies the necessary condition. This result says that, if by nature \( \text{Cov}_P(r_1, r_2) \) is fixed, the subjective covariance of the expected returns in DM’s mind should be low enough, then all three assets will be invested.

Consequently, if

\[
C^* < \tilde{\lambda}\text{Cov}_P(r_1, r_2) + \theta \text{Cov}_\mu[E_Q(r_1), E_Q(r_2)] < \tilde{C}
\]  \hspace{1cm} (96)

where

\[
\tilde{C} = \min \left\{ \sqrt{AB}, \frac{\hat{r}_1 \hat{r}_2 + \sqrt{AB[k_2^2 + (s_1 - s_2)\kappa_2]}}{s_1 + \kappa} \right\}
\]  \hspace{1cm} (97)

then, the optimal combination will be only asset 1 with risk-free asset.\(^5\) The result tell us that one of reasons that an attractive asset finally is not included into the portfolio is because correlation of its return with other is too high by nature or in DM’s subjective information and evaluation in her belief.

\(^5\) The comparison between \( \sqrt{AB} \) and \( \frac{\hat{r}_1 \hat{r}_2 + \sqrt{AB[k_2^2 + (s_1 - s_2)\kappa_2]}}{s_1 + \kappa} \) requires more restriction on the assumptions and of marginal interest of this paper.
CONCLUSION

In this paper, I offer a way to explain the enduring puzzle of limited participation in stock market by a simple portfolio choice problem, but under a robust mean-variance preference instead of conventional mean-variance preference. The central idea is that, when the DM is ambiguous about the assets return, which is common in the real life, her behavior will be altered, and in this case, she becomes more conservative for allocating her wealth into stock market. The underlying intuition is aligned with Ellsberg’s thought experiments. Furthermore, one astonishing fact inferred from the empirical data is that the effect of ambiguity in one’s portfolio decision is substantial, or even dominant in some situations.

The ultimate goal of this paper is not only to point out that one of the enduring puzzle can be reconciled by employing new preference model, but to demonstrate that a careful and deep study in the way how one makes decisions under uncertainty can provide us with powerful tools to understand myths puzzling us. Whenever we meet the situation that some theoretical results are inconsistent with what we observed, if the reasoning is solid, we have to question about the untestable assumptions that we have made as the starting point of the reasoning.

In the case of financial market, many models assume rational expectation and a perfectly informed DM; and this leads to inconsistent estimation of parameters because of the neglecting the effect ambiguity and ambiguity aversion, which has been illustrated to substantially affect one’s decision in Ellsberg’s paradox. Indeed, it is very necessary to re-examine the the decision features that DMs employ in the model to carry forward our understanding in many problems and puzzles.
A.1

Table A1: Financial assets and stock holding of US households in 2007

<table>
<thead>
<tr>
<th>deciles of assets</th>
<th>min</th>
<th>max</th>
<th>media</th>
<th>mean</th>
<th>% of hh holding stock</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>90</td>
<td>0</td>
<td>11</td>
<td>1.0%</td>
</tr>
<tr>
<td>2</td>
<td>100</td>
<td>990</td>
<td>400</td>
<td>438</td>
<td>2.7%</td>
</tr>
<tr>
<td>3</td>
<td>1000</td>
<td>2790</td>
<td>1600</td>
<td>1685</td>
<td>4.1%</td>
</tr>
<tr>
<td>4</td>
<td>2800</td>
<td>7210</td>
<td>4750</td>
<td>4782</td>
<td>11.9%</td>
</tr>
<tr>
<td>5</td>
<td>7220</td>
<td>17850</td>
<td>11250</td>
<td>11697</td>
<td>23.1%</td>
</tr>
<tr>
<td>6</td>
<td>17900</td>
<td>46880</td>
<td>28870</td>
<td>29693</td>
<td>35.5%</td>
</tr>
<tr>
<td>7</td>
<td>47000</td>
<td>137800</td>
<td>81675</td>
<td>83707</td>
<td>53.0%</td>
</tr>
<tr>
<td>8</td>
<td>138000</td>
<td>615900</td>
<td>279420</td>
<td>310682</td>
<td>76.0%</td>
</tr>
<tr>
<td>9</td>
<td>617000</td>
<td>4180000</td>
<td>1534000</td>
<td>1772598</td>
<td>88.9%</td>
</tr>
<tr>
<td>10</td>
<td>4190000</td>
<td>6.98 × 10^8</td>
<td>1.5 × 10^8</td>
<td>35252368</td>
<td>95.2%</td>
</tr>
</tbody>
</table>

A.2

This section shows that if we set $\theta \text{Var}_\mu [E_Q(r_e)] = m\hat{\lambda} \text{Var}_P(r_e)$ in the robust mean-variance model for estimation, finally we can get equation (52) and (53):

$$\hat{\lambda} = \frac{1}{1 + m} \hat{\lambda}$$
$$\hat{\epsilon} = \hat{\epsilon}$$

where $\hat{\lambda}$ and $\hat{\epsilon}$ are the estimates from mean-variance model with same data set.

1) We have already know that, from the probit model same as that in chapter 3, we can get the estimates, $\hat{\beta}_0$ and $\hat{\beta}_1$, and thus, recover

$$\hat{\lambda}\hat{\epsilon} = \frac{\hat{\beta}^2\hat{\epsilon}^2}{2(1 + m)\text{Var}_P(r_e)}$$

Recall that in the mean-variance model,

$$\hat{\lambda}\hat{\epsilon} = \frac{\hat{\beta}^2\hat{\epsilon}^2}{2\text{Var}_P(r_e)}$$

So we have

$$\hat{\lambda}\hat{\epsilon} = \frac{\hat{\lambda}\hat{\epsilon}}{1 + m}$$
2) Using the Heckman selection model, we can derive that

\[
E(\tilde{q}_i|\epsilon_i/\sigma_\epsilon > -N_i) = \frac{E_p(r_{xi}) - r_f + \sigma_\epsilon \gamma(N_i)}{\lambda Var_p(r_{xi}) + \theta Var_\mu[E_Q(r_i)]} - \frac{E_p(r_{xi}) - r_f + \sigma_\epsilon \gamma(N_i)}{(1+m)\lambda Var_p(r_{xi})}
\]

where \(N_i = \beta_0 + \beta_1 \sqrt{w_i}\).

Use the estimates from the probit model, and compute \(N_i = \hat{\beta}_0 + \hat{\beta}_1 \sqrt{w_i}\) and the inverse Mills ratio \(\gamma(N_i)\). Again, regress \(q_i\) on \(y_i = \frac{E_p(r_{xi}) - r_f + \sigma_\epsilon \gamma(N_i)}{\lambda Var_p(r_{xi})}\) without constant. We get same estimation of the regression coefficient \(\hat{\gamma}\) as before since all the elements used to construct \(y_i\) and the specification are exactly as the same in the previous estimation. But in this case,

\[
\hat{\gamma} = \frac{1}{(1 + m)\hat{\lambda}}
\]

So we can get

\[
\hat{\lambda} = \frac{1}{(1 + m)\hat{\gamma}} = \frac{\hat{\lambda}}{1 + m}
\]

Thus,

\[
\hat{\epsilon} = \tilde{\epsilon}
\]

A.3

Proof that under the assumptions made in chapter 6, inequity (95) is a necessary condition for holding all three assets.

Proof. 1)

\[
V_1 \leq V_{12} \iff s_1 \leq s_{1,2} - \kappa
\]

To simplify the notation, Let \(\hat{r}_k = E_p(r_{xi}) - r_f, k = 1, 2\) and

\[
\lambda Var_p(r) + \theta Var_\mu[E_Q(r)] = \begin{bmatrix} A & C \\ C & B \end{bmatrix}
\]

where:

\[
A = \lambda Var_p(r_1) + \theta Var_\mu[E_Q(r_1)],
B = \lambda Var_p(r_2) + \theta Var_\mu[E_Q(r_2)],
C = \lambda Cov_p(r_1, r_2) + \theta Cov_\mu[E_Q(r_1), E_Q(r_2)].
\]

Then

\[
\{\lambda Var_p(r) + \theta Var_\mu[E_Q(r)]\}^{-1} = \frac{1}{AB - C^2} \begin{bmatrix} B & -C \\ -C & A \end{bmatrix}.
\]

It follows that

\[
s_{1,2} = \frac{B\hat{r}_1^2 + A\hat{r}_2^2 - 2C\hat{r}_1\hat{r}_2}{AB - C^2}
\]
\[ s_1 = \frac{\tilde{r}_1^2}{A} \]
\[ s_2 = \frac{\tilde{r}_2^2}{B} \]

Then, the inequality (98) is equivalent to

\[ \frac{\tilde{r}_1^2}{A} \leq \frac{B\tilde{r}_1^2 + A\tilde{r}_2^2 - 2C\tilde{r}_1\tilde{r}_2 - \kappa_2}{AB - C^2} \]

\[ \Leftrightarrow (\tilde{r}_1^2 + \kappa_2 A)C^2 - 2\tilde{r}_1\tilde{r}_2 AC + (\tilde{r}_2^2 A - \kappa_2 A^2 B) \geq 0 \]
\[ \Leftrightarrow (s_1 + \kappa_2)C^2 - 2\tilde{r}_1\tilde{r}_2 C + AB(s_2 - \kappa_2) \geq 0 \]

Consider the quadratic equation

\[ (s_1 + \kappa_2)C^2 - 2\tilde{r}_1\tilde{r}_2 C + AB(s_2 - \kappa_2) = 0 \] (99)

and its discriminant is

\[ \Delta = 4\tilde{r}_1\tilde{r}_2 - 4AB(s_1 + \kappa_2)(s_2 - \kappa_2) \]
\[ \Rightarrow \Delta = 4AB[\kappa_2^2 + (s_1 - s_2)\kappa_2] > 0 \]

the strict inequality hold by assumptions. Let

\[ C^* = \frac{\tilde{r}_1\tilde{r}_2 - \sqrt{AB[\kappa_2^2 + (s_1 - s_2)\kappa_2]}}{s_1 + \kappa_2} > 0 \]

is the smaller solution of the quadratic equation (99) and by Velta’s theorem, it is easy to check that both roots should be greater than zero. and now it can be easily seen that

\[ (s_1 + \kappa_2)C^2 - 2\tilde{r}_1\tilde{r}_2 C + AB(s_2 - \kappa_2) \geq 0 \Rightarrow C \leq C^* . \]

2) One thing need to to be further checked is that \( C^* \) is legitimate in the sense that

\[ C^{*2} < AB \]

where \( AB \) is the upper bond for any legitimate \( C^2 \).

Consider another root of the quadratic equation (7.2), \( C^* > C > 0 \). By Velta’s Theorem,

\[ C^* \hat{C} = AB \frac{(s_2 - \kappa_2)}{(s_1 + \kappa_2)} < AB \]

which follows that

\[ C^{*2} < C^* \hat{C} < AB \]

\[ \square \]


