CAPM with View Bias Adjustment under Imperfect Information

Wei Hu   ZhenLong Zheng[1]

Nov 14, 2007

[1] Zhenlong Zheng, Department of Finance, School of Economics, Xiamen University, Email: zlzheng@xmu.edu.cn, Tel: +86-13906038903 Fax: +86-592-5920923

Wei Hu, Department of Finance, School of Economics, Xiamen University / Australian School of Business, School of Banking and Finance, University of New South Wales. Email: huwei2406@gmail.com Tel: +61-412-338567

The authors thank Rong Chen, Hai Lin, Kaihao He, Shunming Zhang, Li Wu, and all the seminar participants at Xiamen University for helpful comments and suggestions. We also thank David Feldman for his comments and supports.

CAPM with View Bias Adjustment under Imperfect Information

Nov 14, 2007

Abstract

In this paper, we extend the traditional CAPM theory by introducing a new concept of risk-reward measurement based on view bias adjustment under imperfect information. The main result is that the generalized expectation of the excess rate of return can still be described in a single beta representation, except that the systematic risk is now the weighted average of exposed risk and potential risk. Meanwhile imperfect information can induce instantaneous profit by repackaging portfolios, and we name it information premium. Empirical study indicates that this new concept can help to explain the equity premium puzzle in a way that people have pessimistic view bias, when there is no perfect information in the postwar US, and it explains the momentum by the fact that view bias reciprocates from pessimism to optimism, and it might be a mean reversion process.

Keywords: CAPM, imperfect information, view bias, mean reversion
JEL classifications: G12, G14, C51

1. Introduction

1.1 Goal and structure of paper

The goal of this paper is to generalize the traditional CAPM model under imperfect information. People only know the possible results of an uncertainty, they do not know the exact probabilities of each state, but just have a vague assessment. Under imperfect information, if people are pessimistic, they will relatively amplify the possibilities of left tail events, and the reverse happens for optimistic people. We establish a general framework to measure risk and reward which embodies the traditional measurements, variance and mean, as special cases in which people are view neutral. Please note that view neutral is different from the concept of risk neutral. It refers that people are neither pessimistic nor optimistic under perfect information. If we explain that the pessimism or optimism comes from unknowing, view neutrality means that people know the true probabilities or at least they have a clear subjective probability assessment.

We examine if these new measurements are coherent, that is, if they satisfy the additivity (sub-additivity), homogeneity and risk-free condition properties. The result is that if there is no risk source dimensional receding, the three properties are satisfied. From these properties, we get the following result, when portfolios are repackaged, the reward remains the same, if and only if people are view neutral. When pessimism exists, a person who has the information of probability distribution of each individual security can repackage portfolios to do no-risk arbitrage; when optimism exists, they can split the package to earn positive excess return. There is no conflict between this result and the traditional no arbitrage theory, since under this general framework of risk-reward measurement, information is just another factor, which influences asset pricing besides time and risk. We define the premium as the information premium under imperfect information environment.

We generalize CAPM theory based on the new risk-reward concepts, using general expectation of utility as the optimization object instead. The pricing equation now is

$$\mu_i + \frac{\sigma_i}{\sqrt{dt}}\Theta - r^f = \frac{\Theta^2 \tilde{\sigma}_{iM} + \Phi \sigma_{iM}}{\Theta^2 \tilde{\sigma}_M^2 + \Phi \sigma_M^2}\left(\mu_M + \frac{\sigma_M}{\sqrt{dt}}\Theta - r^f\right) \tag{1}$$

where r^f denotes the risk-free rate, μ_i is the expected return on the i^th^ portfolio, μ_M is the expected return on market portfolio, $\frac{\sigma_i}{\sqrt{dt}}$ and $\frac{\sigma_M}{\sqrt{dt}}$ are the standard deviation of annualized continuously compounded rate of return of the i^{th} security and market portfolio respectively, σ_{iM} is the covariance of the i^{th} security return and market portfolio return, $\tilde{\sigma}_M^2$ and $\tilde{\sigma}_{iM}$ are the potential variance of the market portfolio and potential covariance of the i^{th} security return and market portfolio return respectively. The items of Θ and Φ are two adjustors which are constant when the extent of the unknowing is given.

According to this new general CAPM model, the excess return expectation is still proportional to beta, although here it is a kind of adjusted excess return expectation, and beta is a weighted average of exposed risk and potential risk. When perfect information is available, people will be view neutral, and then the model will degenerate into a classical CAPM model.

Expanding the stochastic discount factor theory in some sort, we get a beta pricing model which can explain the equity premium puzzle in a way that people are pessimistic when there is no perfect information in postwar US. In the empirical study, assuming a reasonable constant risk aversion, we calculate the implied view bias, which solves the beta pricing equation. We get a time series with a mean of 0.497, a slight deviation from neutral 0.5. However if we neglect the view bias factor, a huge risk aversion is needed, that is what we say, the equity premium puzzle.

We make the hypothesis that view bias reciprocates from pessimism to optimism, and it is a mean reversion process, then we could explain the momentum in the following way. Since the systematic risk is the weighted average of exposed risk and potential risk, if people are pessimistic or optimistic under imperfect information, the sum of exposed risk and potential risk dominates; if perfect information is available, people are view neutral, then the exposed risk dominates. Therefore, if view bias is a mean reversion process, beta coefficient will vary between exposed beta and the summation back and forth, and then the momentum is captured. In the empirical study, we do not test the mean reversion hypothesis directly, but test if the periodicities of both view bias and momentum ranking are compatible instead.

Since there is no existing econometrics method to test VCAPM (view bias based CAPM), we develop VOLS (We make this name in accordance with the mathematical model VCAPM), including resetting the assumptions, looking for new estimators, developing the new asymptotic tools based on general mean (WLLN and CLT) to do hypothesis test. At last, we try to use GMM integrated with VOLS method to do VCAPM empirical analysis. That would be another contribution of this paper.

The paper is organized the following way. Chapter 1 is mainly about the introduction of this paper. In chapter 2, we focus on defining new concepts. Chapter 3 is about the deduction of VCAPM and how it can be used to explain the momentum phenomena. In chapter4, we try to explain equity premium puzzle from view bias perspective. Then the next chapter, we work on a new econometrics method feasible for VCAPM empirical analysis. In chapter 6, we do the empirical analysis. Finally, we summarize the conclusions in chapter 7.

1.2 Motivation and literature review

This paper is motivated from the following five considerations, and the momentum effect is our starting point. Momentum firstly documented by Jegadeesh and Titman (1993), which shows that past winners continue to outperform the past losers, while the beta estimate for the winner portfolio is even lower. Fama and French (1996) find that among several CAPM anomalies, momentum is the only one unexplained by the three-factor model. Behavioral explanations gain empirical supports, e.g., Hong and Stein (1999) focus on investor irrationality in processing information, namely risk does not play any significant role in capturing the momentum payoff. Concurrently risk based explanations are being challenged by a number of researches, see Jegadeesh and Titman (2002), Criffin, Ji, and Martin (2003), Cooper, Gutierrez, and Hameed (2004), Zhang

(2005). The following theories give rationality-based explanations. Conrad and Kaul (1998) point out that cross-sectional variation in expected returns can possibly explain momentum. Berk, Green, and Naik (1999) show that momentum, along with several other anomalies, can be generated from the life-cycle variation of firms' endogenously chosen projects. Johnson(2002) show that momentum can arise from a positive relation between expected returns and growth rates. Literature on parameter uncertainty is growing. Coles, Loewenstein and Suay (1995) investigate the effects of parameter uncertainty on equilibrium pricing. Brav and Heaton (2002) stress that parameter uncertainty can play an important role in explaining asset pricing anomalies, and it is difficult to distinguish behavioral models and rational models with structural uncertainty. Wang (2006) contribute to explain the dependence of momentum profits on the lagged market performance, suggests that test results are consistent with an extended version of the CAPM, a theory in which investors rationally process information. New information revealed around an important corporate event in the ranking period induces high beta uncertainty. Their confidence about their previous beta estimate is blown away. The beta adjustment creates momentum in stock returns. Eric Ghysels and Eric Jacquier (2006) combine data-driven and parametric methods and turn to beta dynamics for resolutions as well. Since the explanations are still not satisfactory. Can we explain the momentum in another way?

Another dramatic challenge is the equity premium puzzle by Mehra and Prescott (1985). In post war U.S. data, the slope of average return-beta line is much higher than reasonable risk aversion and consumption volatility estimates suggest. Mehra (2003) summarize the efforts in point, including the alternative assumptions about preferences, disaster states, survivorship bias, incomplete markets, and market imperfections. Guidolin (2006) suggest that in the presence of infrequent but observable structural breaks, the representative agent is on a rational learning path, which is modeled as a binomial lattice, concerning that the real consumption growth process can generate high equity premia and low risk-free interest rates without requiring a strong curvature of the utility function.

Thirdly, the idea that risk and uncertainty are relevant for economic analysis was clearly suggested by Frank Knight (1921) in his famous book Risk, Uncertainty and Profit. If the probabilities of the events occurring are rarely known, we usually have to act on our best beliefs and subjective assessments. CAPM studies how people make their choices between consumptions and investments, and how to choose among portfolios when possible states of the world and the corresponding probabilities of each state are known. If people are blind to the probabilities, how do they make choice, and how could we model the process of making choice?

Fourthly, many people show their stable risk aversion in the psychological tests, but they still will take risks when they are very optimism. Weinstein (1980) find unrealistic optimism in estimating the chances of future life events. Hey (1984) define an optimist as someone "…who thinks that [some event] E is more likely when faced with the prospect {£100 if E happens; £0 if E does not happen} than when faced with the prospect {£100 if E does not happen; £0 if E happens}". Abel (2002) redefines pessimism as the case in which the rational learning predictive distribution is first-order stochastically dominated by the full information distribution. Muren (2006) design an experiment to test if individuals show (unrealistic) optimism when determining their subjective probabilities about exogenous circumstances. If risk loving means to increase the subjective probability of good states, what is the difference between risk loving and optimism?

Finally, expected utility maximization axiom implies many assumptions, and people maximize the expectation of utility is one of them. The subjective probability will determine the risk measurements instead of objective probability. Acerbi (2002) and Basset et al (2004) take efforts on distortion risk measures, which model the pessimism or optimism factor through a distortion function. Gourieroux and Liu (2006) provide a unified statistical framework for the analysis of distortion risk measures, derive the general formula for calculating the asymptotic distribution of the nonparametric estimator of the distortion risk measures, and provide the closed form expressions for special examples such as VaR, Tail-VaR and proportional hazard-distortion risk measure. Bassett et al. (2004) revise the expected utility maximization axiom with a modified probability measure, namely Choquet expected utility Axiom. It indicates that a general form of pessimistic portfolio optimization based on the Choquet approach may be formulated as a problem of linear quantile regression. Nevertheless, the cost is that the expectation does not exist, and the martingale theory is not available any more, therefore many existing asset pricing theories are not useful then. Another reason why we introduce new measures, not using quantile or other distortion measures is that we are studying how people make their choices under imperfect information. In other words, the information or the probability of each state is the object being examined. We hope the information is being lost within control, and we can model the whole process.

Our contribution is that we sew up the pieces together by introducing new concepts. Setting out from motivation 3 and 4, we define a new measurement of risk and reward of return to capture people's view bias variation under imperfect information. If people are blind to the probabilities of each state of the world, they may be optimism or pessimism, and then the measurement of risk and reward should be different from the measurements we used before, i.e., mean and variance. Since the new measurements embody mean and variance as special cases, we name them general mean and general variance. Then according to motivation 5, we revise the object of the optimization problem as the general expected utility maximization. After redoing the Merton problem, we get a view bias based CAPM, named VCAPM, using which we explain the momentum phenomena and equity premium puzzle to some extent.

2. A General Framework of Risk-Reward Measurement

2.1 Definition

Consider a random variable X whose probability density function is $f_X(x)$, and $\theta \in (0,1)$ is the view bias coefficient.

The expected return of X can be obtained by minimizing the following objective function with respect to q,

Definition 2.1.1 [General mean]: $E_\theta^1(X) := \arg\min_q \left[(1-\theta)\int_{X<q} (x-q)^2 f_X(x)dx + \theta\int_{X>q} (x-q)^2 f_X(x)dx \right]$,

where X is absolute-integrable, i.e., $\int |x| f_X(x)dx < \infty$.

If θ equals to 50%, the above definition is the ordinary mathematical expectation. If people are pessimistic, θ is less than 50%, from the definition, they will pay more attention to the unpleasant states, and vice versa.

Solving the minimization problem by taking the first order derivative, we can rewrite the definition this way

$$E_\theta^1(X) := q^* = \int \pi_X(\theta) x f_X(x)dx \tag{2}$$

Where

$$\pi_X(\theta) := \frac{(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx} \tag{3}$$

At first glance, definition (2) with (3) is very artificial, but by understanding how it forms, it is more acceptable. Virtually quantile is generated from a very similar minimization problem, but using $|x-q|$ rather than $(x-q)^2$ in the objective function. We will go deep into the details of the definition in Appendix I.

From definition (3), we know that $\pi_X(\theta) f_X(x)$ is a probability measure, and $E_\theta^1(X)$ is a mathematical expectation of X, adjusted by view bias θ. The superscript stands for the risk source dimension, we will show its' necessity in section 2.3, where the properties of general mean and variance would be discussed. Since $E_\theta^1(X)$ is an implicit function, and there is no closed form, we can only get the result through numerical method.

Simultaneously we get the risk definition:

Definition 2.1.2 [General variance]: $D_\theta^1(X) := \int (x - E_\theta^1(X))^2 \pi_X(\theta) f_X(x)dx$

It is easy to prove that

$$D_\theta^1(X) = E_\theta^1(X^2) - \left[E_\theta^1(X)\right]^2 \tag{4}$$

Proposition 2.1.1: $E_\theta^1(X)$ is a monotonously increasing function with respect to θ.

Please see appendix I for the proof. Through numerical method, we get a monotonously increasing function of general mean with respect to θ. What's more, we find that $D_\theta^1(X)$ is invariable with respect to θ at least for normal distribution. Please see Figure 1,

<INSERT FIGURE 1 ABOUT HERE>

When two random variables are involved, we define the conditional mean first, using which we define the general mean later.

Definition 2.1.3 [Conditional mean (Linear)]:

$$E_\theta^2(Y \mid X) := \arg\min_{g \in A} E_{\theta,X,Y}^2 [Y - g(X)]^2$$
$$= \arg\min_{g \in A} E\left[\pi_X \pi_{Y|X} (Y - g(X))^2\right]$$

Where $A = \{g : \Re^2 \to \Re \mid g(X) = \beta_0 + \beta_1 X\}$, and

$$\pi_{Y|X}(\theta) := \frac{(1-\theta)_{sign(\beta_1)(Y-\beta_0-\beta_1 X)<0} + \theta_{sign(\beta_1)(Y-\beta_0-\beta_1 X)>0}}{\int [(1-\theta)_{sign(\beta_1)(Y-\beta_0-\beta_1 X)<0} + \theta_{sign(\beta_1)(Y-\beta_0-\beta_1 X)>0}] f_{Y|X}(y \mid x) dy}$$

Definition 2.1.4 [General mean]: If $E\left[\pi_X \pi_{Y|X} h(X,Y)\right] = E\left[\pi_Y \pi_{X|Y} h(X,Y)\right]$, then the general mean $E_\theta^2[h(X,Y)]$ is well defined as

$$E_\theta^2[h(X,Y)] := E\left[\pi_X \pi_{Y|X} h(X,Y)\right] = \int\int \pi_X \pi_{Y|X} h(X,Y) f(x,y) dx dy$$

The conditional mean of higher dimension is not well defined. However, if the content, which the expectation operator is being taken onto，is additive separable, the general mean is well defined as follows.

Definition 2.1.5 [General mean (additive separable)]:

$$\mathrm{E}_\theta^{n+}[h_1(X_1,Y_1) + h_2(X_2,Y_2) + \cdots + h_m(X_m,Y_m)]$$
$$:= E[\pi_{X_1,Y_1|X_1} h_1(X_1,Y_1) + \pi_{X_2,Y_2|X_2} h_2(X_2,Y_2) + \cdots + \pi_{X_m,Y_m|X_{m1}} h_m(X_m,Y_m)]$$
$$= E[\pi_{Y_1,X_1|Y_1} h_1(X_1,Y_1) + \pi_{Y_2,X_2|Y_2} h_2(X_2,Y_2) + \cdots + \pi_{Y_m,X_m|Y_{m1}} h_m(X_m,Y_m)]$$

Where the superscript, '+' indicates the additive separable, 'n' stands for the n risk-source dimension, and here it is no more than twice of m. In the general mean framework, moments are defined based on conditional moments. From the econometrics analysis in chapter 5, we can see that the moments, other than conditional moments are more useful to describe the relationships of multi-variables. That is the reason why we combine GMM and VOLS to estimate VCAPM.

Definition2.1.6 [General Variance]:

$$D_{\theta}^{n+}(Z) := E_{\theta}^{n+}(Z - E_{\theta}^{n+}(Z))^2$$

Definition 2.1.7 [General Covariance]:

$$Cov_{\theta}^{2}(X,Y) := E_{\theta}^{2}(XY) - E_{\theta}^{1}(X)E_{\theta}^{1}(Y)$$

2.2 Further Analysis

CAPM is the backbone of academic finance, but recent empirical tests have challenged the CAPM by identifying several powerful anomalies, e.g., the momentum effect and the equity premium puzzle which attract many researches. We introduce a new component, view bias under imperfect information, besides risk preference to try to give some explanation. We set out from a simple psychological test.

<INSERT TABLE 1 ABOUT HERE>

Suppose there are four states in the world. The payoff and the probabilities of two strategies X and Y under each state are shown in table 1. Setting out from the traditional mean variance efficiency, if people know both the payoff and probability, the risk averse will prefer Y strategy, since the expectation of Y is greater than X, meanwhile the variance of Y is less than that of X.

Now, let us go through another situation.

<INSERT TABLE 2 ABOUT HERE>

When people are blind to the probabilities of each state, their dicisions will depend on whether they are pessimistic or optimistic. When they are pessimistic, they will choose the strategy following the maxmin principle. The minimum of X is one, and Y is -1000, people select the maximum of them. That is, strategy X is preferred. The intrinsical enlightenment here is that under imperfect information, new measurements of risk and return are needed, accordingly the object of the optimization should be revised as well.

However, if the minimum payoffs of each strategy are no difference, e.g., zero for the price of a security, or negative infinity for the continuous compounded growth rate, the maxmin principle is of not much help. So many people turn to quantile to measure risk and reward, but the cost is that the expectation does not exist, and the martingale theory is not available any more. That is why we introduce general mean, which embodies the traditional measurements, variance and mean, as special cases when people are view neutral. When people are pessimistic, they will relatively amplify the possibilities of left tail events, and vice versa. If we explain that pessimism or optimism comes from unknowing, view neutrality means people know the true probabilities or at least they have a clear subjective probability assessment. Figure 2 gives us intuition.

< INSERT FIGURE 2 ABOUT HERE >

The relation between $E_\theta^1(x)$ and $E(x)$ is determined as follows,

$$E_\theta^1(x)\begin{cases} > E(x) & \theta > 50\% \\ = E(x) & \theta = 50\% \\ < E(x) & \theta < 50\% \end{cases} \tag{5}$$

Another reason why we introduce general mean, not using quantile, is that we are studying how people make their choices under imperfect information. In other words, the information or the probability of each state is the object being examined. We hope the information is being lost within control, and we can model the whole process. General mean is of superiority from this perspective. We use two existing information measurements defined as follows,

Definition 2.2.1: The differential entropy of a continuous random variable X with density f is:

$$H(X) = H(f) = -\int f_X(x) \log f_X(x) dx$$

Definition 2.2.2: For two continuous random variables with density f and h, define the relative distance from f and h to be

$$D(f \parallel h) = \int f_X(x) \ln\left(\frac{f_X(x)}{h_X(x)}\right) dx$$

Proposition 2.2.1: For any probability densities f and h, if $h_X(x) = \pi_X(\theta) f_X(x)$, where $\pi_X(\theta)$ defined as formula (3) then

$$\frac{dD(f \parallel h)}{d\theta} := \begin{cases} > 0 & \theta > 50\% \\ = 0 & \theta = 50\% \\ < 0 & \theta < 50\% \end{cases} \tag{6}$$

The father of information theory, Shannon, named $H(X)$ entropy. It quantifies how far from deterministic a random variable is. And $D(f \parallel h)$ is the relative entropy named as Kullback-leibler distance. From Gibbs Inequality,

Lemma 2.2.1: For any probability densities f and h, $D(f \parallel h) \geq 0$ with equality if and only if $f = h$ almost everywhere.

We know that the probability information is lost as h deviates from f. In this paper, we care more about how view bias θ take effects in the information losing process. Therefore, we prove the proposition 2.2.1, and get the result that the more deviation from 50% the view bias has, the greater the probability information is lost within general mean framework. We also prove that,

Proposition 2.2.2: $0 \le D(f \| h_{general\ \ mean}) \le D(f \| h_{quantile}) = +\infty$ with the first equality if and only if $\theta = 50\%$.

That means using quantile to measure the reward of return, the probability information is lost out of view bias θ's control.

We make comparisons between risk preference and view bias, and summarize the results in table 3

<INSERT TABLE 3 ABOUT HERE>

Extending stochastic discount factor theory, we get the pricing equation within general mean framework,

$$P(x) = E_{\theta}^{1}(mx) = \int_{\Omega} \pi(\theta) mxdf(x)$$

where x is the payoff of an asset, m is a change of measure by risk preference, $\pi(\theta)$ is a change of measure by view bias. The total effect is the combination of those two factors. More important, they are separable, and that would make the analysis much easier.

The similarity of risk preference and view bias is that both of them are changes of measure from real to subjective probability, and for both of them, only means are altered, variances remain the same. The dissimilarities can be detailed from the following aspects.

First, risk aversion changes the measure from real probabilities to subjective probabilities. This is some kind of shift, which maintains all the probability information. Pessimism is also equivalent to paying more attention to the unpleasant states, but through the way of amplifying the weight of unpleasant states, and shrinking the probabilities of the other states, it is some kind of reshaping. The probability information of the latter is lost gradually as unknowing becomes more and more serious. So view bias adjustment is an in-continuous change of measure actually, that make it possible to be statistically separated from risk preference, which is another change of measure, but continuously.

Next, the concept of risk preference exists no matter information is perfect or not. It describes whether the marginal utility of one dollar in unpleasant state is greater or not. It is people's character, which does not vary a lot. However, the pessimism and optimism are meaningful only under the imperfect information. It scales the attitude in face of the unknowing, and it should vary when situation changes. What's more, unnecessarily the pessimistic investor is risk averse, and vice versa.

The last dissimilarity is that risk preference causes no instantaneous profit when portfolios are repackaged, however that is not true for view bias. We can understand it better after going through the next section.

2.3 Properties

We want to examine whether the measurements are coherent, namely, whether they satisfy three properties: additivity (sub-additivity), homogeneity and risk free conditions. In this section, we are to explain two doubtful points here. The first, under what condition, considering the first two order general moments is enough, or make it more explicit, under what condition, general mean variance efficiency strategy and expected utility maximization strategy are compatible? The second, if the coherence of return measurement is not satisfied, is no arbitrage violated? Should it still be used to measure the reward of return here?

Normal distributions are the only stable ones whose first two order moments can fully reveal their distribution information. Stable here refers that if the asset returns are multivariate normal distribution, the return of the portfolio made up of the aforementioned assets is still normal distributed. Therefore, if view bias is given, risk sources being normal distributed, although not necessary, can guarantee the compatibility. From another perspective, with no distribution constraints, quadratic utility can also satisfy the compatibility.

Coherence for risk measurement is just a criterion to evaluate which one is more reasonable. Whereas we seldom see the discussion on the coherence of reward measurement, for that is a necessary condition to be a measurement. If it is not satisfied, arbitrage opportunity exists, and then the assets would be mispriced. However, in this paper, after imperfect information and view bias being introduced, coherence for return measurement is not a necessity any more. It is not a violation of no arbitrage opportunity, since the premium is not the compensation for bearing risk, but providing more information. We name it information premium.

i. Additivity:

Proposition 2.3.1: a) $E_\theta^{n+}(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} E_\theta^1(X_i)$ b) If $X_i, \quad i=1,2,\cdots n$ are perfectly correlated random variables, then $E_\theta^{n+}(\sum_{i=1}^{n} X_i) = E_\theta^1(\sum_{i=1}^{n} X_i)$, or the equality exists, if and only if $\theta = 50\%$.

Proposition2.3.1 is obvious from definition 2.1.5, and it is easy to find an example to prove that if both the perfect correlation and $\theta = 50\%$ are not satisfied, it might be that

$$E_\theta^1(\sum_{i=1}^{n} X_i) \neq \sum_{i=1}^{n} E_\theta^1(X_i) \tag{7}$$

Hence, if there is no risk-source dimensional receding, additivity is satisfied. We define the information premium as

$$\left| E_\theta^{n+}(\sum_{i=1}^{n} X_i) - E_\theta^1(\sum_{i=1}^{n} X_i) \right| \tag{8}$$

The superscript is very useful to remind us if there is any risk sources dimensional receding. Since it is confusing, we mark the power outside the bracket from now on. For example, $\left[E_\theta^{n+}(\sum_{i=1}^{n} X_i)\right]^2$ stands for the square of n dimensional general mean of the sum of n random variables.

Proposition 2.3.2: if $X_i \sim N(\mu_i,\sigma_i)$, $\exists\rho_{ij} \neq 1$, the correlation between X_i and X_j is ρ_{ij}, then

$$E_\theta^{n+}(\sum_{i=1}^{n} X_i) - E_\theta^{1}(\sum_{i=1}^{n} X_i)\begin{cases} > 0 & \theta > 50\% \\ = 0 & \theta = 50\% \\ < 0 & \theta < 50\% \end{cases} \tag{9}$$

We understand it this way, when portfolios are repackaged, the return remains the same, if and only if people are view neutral. If pessimism exists, people can assemble them into portfolio to do riskless arbitrage; if optimism exists, they can split the package to earn non-zero excess return. There is no conflict between this result and the traditional no arbitrage theory, since under this general framework of risk-reward measurement, information is another factor that influences asset pricing besides time and risk. People who earn non-zero excess return through repackaging must know more about the probabilities of each individual security. The premium is not the compensation for bearing risk, but providing more information. We name it information premium.

Within the general framework of risk-reward measurement, we reexamine the market completeness in contingent claim market, Arrow security market and ordinary security market. Market completeness expands itself from security level to portfolio level. To make sure each elementary adopted consumption process obtainable, there should be no portfolio repackaging constraints. Getting information premium through repackaging can improve the welfare of two parties, and it is a Pareto equilibrium allocation process.

ii. **Homogeneity:**

$$E_\theta^{2}(\alpha Z) = \alpha E_\theta^{2}(Z)\,; \quad E_\theta^{n+}(\alpha Z) = \alpha E_\theta^{n+}(Z) \tag{10}$$

Homogeneity is satisfied.

iii. **Risk-free condition:**

$$E_\theta^{2}(r+Z) = r + E_\theta^{2}(Z)\,; \quad E_\theta^{n+}(r+Z) = r + E_\theta^{n+}(z) \tag{11}$$

Risk-free condition is satisfied.

We also prove that,

i. Sub-additivity:

$$D_\theta^{2+}(X+Y) = D_\theta^{1}(X) + D_\theta^{1}(Y) + 2COV_\theta^{2}(X,Y) \tag{12}$$

ii. **Homogeneity:**

$$D_\theta^{n+}(\alpha Z) = \alpha^2 D_\theta^{n+}(Z) \tag{13}$$

iii. **Risk-free condition:**

$$D_{\theta}^{n+}(r+Z) = D_{\theta}^{n+}(Z) \tag{14}$$

We find many properties that the mean and variance hold are not held in general sense; this is merely because the former are special cases of the latter. For example,

$$Y = \rho X \Rightarrow E_{\theta}^{1}(Y) = \rho E_{\theta}^{1}(X), \quad iff \quad \rho \geq 0 \tag{15}$$

We will discuss more details in the following part where it is needed. Anyway we need to be very careful when performing mathematical calculations, and cannot directly use general mean or variance concepts into the formula that we have already got through the mean variance process.

3. CAPM Model Based on View Bias Adjustment (VCAPM)

3.1 View bias adjusted CAPM with constant state variables

In this section, we derive VCAPM formula, assuming that the state variables, such as interest rate, climate, etc., which influence the drift rate or volatility of prices' diffusion processes are constant. The main thought is still from Merton, R (1973b), but we do it in the general framework of risk-reward measurements.

Define:
$W(t) =$ Total wealth at time t
$P_i(t) =$ Price of the ith asset at time t (i=1, ⋯, n)
$C(t) =$ Consumption per unit time at time t
$w_i(t) =$ Proportion of total wealth in the ith asset at time t(i=1, ⋯, n)

Note $[\sum_{i=1}^{n} w_i(t)] \equiv 1$.
Assumption 1: Time interval between each decision is infinitesimal.
Assumption 2: Prices follow diffusion processes.
Assumption 3: Only consumption and portfolio process are controllable.
Assumption 4: There is no exogenous endowment.
Assumption 5: Investors are homogenous.
Assumption 6: Information is imperfect, and pessimism or optimism view bias might exist.

From section 2.3, general mean variance efficiency strategy and general expected utility maximization strategy are compatible, if risk sources are normal distributed. We model the consumption and the portfolio selecting process as follows,

$$J[W(t),t] \equiv \max_{\{C_{(\tau)}, w_{(\tau)}\}} E_{(\theta,\ t)}^{n+}\{\int_{t}^{T} U_1[C(\tau),\tau]d\tau + U_2[W(T),T]\} \tag{16}$$

St: boundary condition: $J[W(T),T] = U_2[W(T),T]$ (17)

budget equations: $$W(t)=\sum_{i=1}^{n} w_i(t_0)\frac{P_i(t)}{P_i(t_0)}[W(t_0)-C(t_0)h] \tag{18}$$

assumption1: $$t\equiv t_0+h,\quad h\to 0 \tag{19}$$

assumption2: $$\frac{dP_i(t)}{P_i(t)}=\mu_i(t)dt+\sigma_i(t)\sqrt{dt}\,\omega_i,\qquad i=1,2,\cdots,n$$

$$V_{(n\times n)}=[\sigma_{il}],\qquad \sigma_{il}=\sigma_i\sigma_l\rho_{il},\qquad i,l=1,2,\cdots,n \tag{20}$$

Using stochastic dynamic programming technique, we get the optimized relationship of view bias adjusted excess return and risk amount between an individual security and market portfolio.

$$\frac{\mu_i(t)+\frac{\sigma_i(t)}{\sqrt{dt}}E^1_{(\theta,\ t)}(\omega)-r^f}{\mu_M(t)+\frac{\sigma_M(t)}{\sqrt{dt}}E^1_{(\theta,\ t)}(\omega)-r^f}=\frac{\sigma^\theta_{iM}(t)}{\sigma^\theta_{MM}(t)}=\frac{[E^1_{(\theta,\ t)}(\omega)]^2\sigma_i(t)\sigma_M(t)\delta_{iM}(t)+[E^1_{(\theta,\ t)}(\omega^2)]\sigma_i(t)\sigma_M(t)\rho_{iM}(t)}{[E^1_{(\theta,\ t)}(\omega)]^2\sigma^2_M(t)\delta_{MM}(t)+[E^1_{(\theta,\ t)}(\omega^2)]\sigma^2_M(t)},\quad i=1,2,\cdots,n \tag{21}$$

where

$$\sigma_i(t)\sigma_M(t)\delta_{iM}(t)=\sum_{j=1}^{n}w_j\sigma_i(t)\sigma_j(t)\sqrt{1-\rho^2_{ij}(t)}\times sign(\rho_{ij}(t))\,;\quad \sigma_M{}^2(t)\delta_{MM}(t)=\sum_{j=1}^{n}w_j\sigma_M(t)\sigma_j(t)\sqrt{1-\rho^2_{jM}(t)}\times sign(\rho_{jM}(t))$$

$$\sigma_i(t)\sigma_M(t)\rho_{iM}(t)=\sum_{j=1}^{n}w_j\sigma_i(t)\sigma_j(t)\rho_{ij}(t)\,;\qquad \sigma_M{}^2(t)=\sigma_M{}^2(t)\rho_{MM}(t)=\sum_{j=1}^{n}w_j\sigma_M(t)\sigma_j(t)\rho_{jM}(t) \tag{22}$$

and ω is any standard normal distributed random variable. Drop t, and define

$$\Theta:=E^1_{(\theta,t)}(\omega)\,;\quad \Phi:=E^1_{(\theta,t)}(\omega^2)\,;\quad \tilde{\sigma}_{iM}:=\sigma_i\sigma_M\delta_{iM}\,;\quad \tilde{\sigma}^2_M:=\sigma^2_M\delta_{MM}\,; \tag{23}$$

Equation (21) is simplified as follows,

$$\mu_i+\frac{\sigma_i}{\sqrt{dt}}\Theta-r^f=\beta^\theta\left(\mu_M+\frac{\sigma_M}{\sqrt{dt}}\Theta-r^f\right),\qquad i=1,2,\cdots,n$$
$$where\quad \beta^\theta:=\frac{\Theta^2\tilde{\sigma}_{iM}+\Phi\sigma_{iM}}{\Theta^2\tilde{\sigma}^2_M+\Phi\sigma^2_M} \tag{24}$$

3.2 Model Specification

From equation (24), expected excess return is adjusted from $\mu-r^f$ to $\mu+\frac{\sigma}{\sqrt{dt}}\Theta-r^f$, where Θ is affected by people's view bias θ under imperfect information. From Figure 2 and formula (5), if investors are pessimistic, i.e., $\Theta<0$, they will devalue their expectations of excess return, and vise versa.

As for the risk measurement, great changes take place in this formula. If perfect information exists, we regard market variance as benchmark. The proportion of covariance σ_{iM} to the benchmark is the amount of risk. That is just a special case when $\Phi=1$ and $\Theta=0$, namely people are view neutral when perfect information exists.

If imperfect information exists, as view bias being apart from 50%, $\Phi > 1$ and $\Theta^2 > 0$, the term $\tilde{\sigma}_{iM}$ and $\tilde{\sigma}_{MM}$ take effects in the formula. From definition (23) and equation (22), we could explain σ_{iM} as the exposed risk and $\tilde{\sigma}_{iM}$ the potential risk. The systematic risk is the weighted average of both of them. Divided by the new benchmark, systematic market risk, the general beta coefficient is obtained.

For the sake of not diluting the whole picture, also because the concepts come up during the course of CAPM proving, we put the details in appendix III, although a deep understanding of potential risk is necessary. Here we just give an intuitive explanation.

< INSERT FIGURE 3 ABOUT HERE >

Market portfolio consists of n individual securities. If we want to examine the risk return trade-off between the ith security and market portfolio, from equation (21) and (22), we need to study the relationship between the ith and any jth security, where j varies from 1 to n, and then sum them together. Figure 3 provides us an intuitive sketch map, J and I stands for the jth and ith security respectively. For convenience, suppose σ_j is unit one, if σ_{iA0} and σ_{iB} are unit one as well, they will fall on the unit circle (see OI_B and OI_{A0}). Since their correlations ρ_{ij} are different, their exposures to J are different (see OE_{AB} and OE_{A0}), meanwhile, their exposures to other random process, which is orthogonal to J (perpendicular to J axis), are OP_B and OP_{A0}. Under perfect information, $\Theta = 0$, the projection of OP_B and OP_{A0} are zero, so they will not be priced. If σ_{iA0} is greater than unit one, e.g., OI_A, the price of OI_A and OI_B are exactly the same, and the difference of OP_A and OP_B takes no effects. Sum all J together, the systematic risk, which is the sum of the exposures to all J, will be priced, and we name the other part, which is the sum of the exposures to other orthogonal random processes, idiosyncratic risk, and that part should not be priced.

Under imperfect information, if view bias is given, $\Phi > 1$, and $\Theta^2 > 0$. From equation (21) and (22), it should be an affine transformation. Considering OI_A and OI_B, the projections on J are OS_A and OS_B respectively, and the distance between OP_A and OP_B shows its' difference in the projection within imperfect information pricing system (see S_AS_B). Here, not only the risk exposed to J but also the risk exposed to other random process, which is orthogonal to J, take effects. OS_B is the weighted average of OE_{AB} and OP_B. Utilize the fact that $D_\theta^1(X)$ is invariable with respect to θ for normal distribution, as well as employ equation (4), we have

$$\lim_{\substack{\theta\to 100\% \\ or.\theta\to 0\%}} \beta^\theta = \lim_{\substack{\theta\to 100\% \\ or.\theta\to 0\%}} \left(\frac{\Theta^2 \tilde{\sigma}_{iM} + \Phi \sigma_{iM}}{\Theta^2 \tilde{\sigma}_M^2 + \Phi \sigma_M^2} \right) = \lim_{\substack{\theta\to 100\% \\ or.\theta\to 0\%}} \left(\frac{\Theta^2 \tilde{\sigma}_{iM} + (\Theta^2+1) \sigma_{iM}}{\Theta^2 \tilde{\sigma}_M^2 + (\Theta^2+1) \sigma_M^2} \right) = \frac{\tilde{\sigma}_{iM} + \sigma_{iM}}{\tilde{\sigma}_M^2 + \sigma_M^2}, \quad i = 1, 2, \cdots, n \tag{25}$$

Since OP_B takes effect gradually as people's view bias departs from neutral inch by inch, we name it potential risk. The systematic risk of security OI_B is the sum of all J's OS_B, and idiosyncratic risk is the sum of I_BS_B. See securities OI_B and OI_{A1}, they are of the same systematic risk OS_B, and different idiosyncratic risks, I_BS_B and $I_{A1}S_B$, but only the systematic risks will be priced.

However, according to the analysis in section 2.2, view bias scales the attitude in face of unknowing, and it should vary whenever situation changes. Therefore, it should not always be a constant. If it is a mean reversion process as follows,

$$d\theta = (p - q\theta)dt + \sigma dW\,; \quad dW = \varepsilon_t \sqrt{dt} \tag{26}$$

Φ and Θ vary with θ. From equation (24), β^θ is determined by more of exposed risk or the summation of exposed and potential risks now and again, and then β^θ appears to be a fluctuation. That is why many researches on momentum set out from beta coefficient mean reversion analysis, but virtually maybe it is not the case.

More interestingly, since β^θ is a weighted average, and the weightings are Θ^2 and Φ, from definition (23), we have

$$\beta^\theta := \frac{\Theta^2 \tilde{\sigma}_{iM} + \Phi\sigma_{iM}}{\Theta^2 \tilde{\sigma}_M^2 + \Phi\sigma_M^2} = \beta^{1-\theta} \tag{27}$$

Therefore, the periodicity of the systematic risk amount is just half as long as the view bias. While the periodicity of general expected excess return $\mu + \frac{\sigma}{\sqrt{dt}}\Theta - r^f$ is in accordance with the view bias. That characteristic to be empirically verified is a weak prediction of view bias mean reversion, and might yield useful applications in explaining the anomalies.

4. New approach to explain equity premium puzzle

In post war U.S. data, the slope of average return-beta line is much higher than what a reasonable risk aversion and consumption volatility estimates suggest. Over the last 50 years, the real stock returns have averaged 9% with a standard deviation of about 16%. While the real return on treasury bills has been about 1%. Aggregate nondurable and services consumption growth had a mean and standard deviation of about 1%. These facts with

$$\left| \frac{E(R^{mv} - R^f)}{\sigma(R^{mv})} \right| \approx a\sigma(\Delta \ln c) \tag{28}$$

can only be reconciled if investors have a risk-aversion coefficient of 50. Or in continuous form,

$$\mu_P + \frac{D_t}{P_t} - r^f = \alpha\sigma_C\sigma_P\rho_{CP} \tag{29}$$

Considering the aggregate consumption has about 0.2 correlation with the market return, risk aversion needs to be 250 to explain the formula (29).

We try to explain it from the perspective of imperfect information. By modifying the continuous model, we get a formula as follows,

$$\mu_P + \frac{D_t}{P_t} + \frac{\sigma_P}{\sqrt{dt}}\Theta - r^f = \alpha\sigma_C\sigma_P\Phi \tag{30}$$

where Θ, Φ are two adjustors, which are functions of view bias θ. The great gap can be explained by people's fear of unknowing in the postwar. If we do not take the case that consumption growth and market returns are perfectly correlated, we have

$$\mu_P + \frac{D_t}{P_t} + \frac{\sigma_P}{\sqrt{dt}}\Theta - r^f = \alpha\sigma_C\sigma_P\left(\Theta^2\sqrt{1-\rho_{CP}^2} \times sign(\rho_{CP}) + \Phi\rho_{CP})\right) \tag{31}$$

Extending the original derivation in some sort, we obtain equation (30), (31). For more details, please see appendix V.

Equity premium puzzle lies in the unmatched real data and CAPM model, our intention is to explain the great discrepancy by introducing one more argument, view bias, to reconcile the real data and revised CAPM model to some extent.

5. Econometrics Model

5.1 Necessity of developing new econometrics model

When we trying to do the empirical analysis, we find there is no existing econometrics method to solve that problem. Because the model correct specification is of the following form, $E_\theta(\varepsilon \mid X) = 0$. (From now on, there is no dimensional receding, and the superscripts are omitted.) However $E_\theta(\varepsilon \mid X) = 0 \Leftrightarrow E(\varepsilon \mid X) = 0 \quad iff \quad \theta = 50\%$, in most cases, we set θ other than 50% to try to explain the anomalies, then the condition $E(\varepsilon \mid X) = 0$ fail, and the model is mis-specified in a traditional linear regression point of view. We usually use the instrumental variable regression to solve the problem. However given θ, which is different from 50%, then $E_\theta(\varepsilon \mid X) = 0 \Rightarrow E(\varepsilon \mid X) \neq 0$, while the converse may not be true. Because $E(\varepsilon \mid X) \neq 0$ contains no information about θ, and obviously we cannot deduce whether or not $E_\theta(\varepsilon \mid X) = 0$. That is, using instrumental variable method, we neglect the information about θ; the 2SLS estimator would not be efficient. So we develop VOLS, including resetting the assumptions, looking for new estimators, developing the new asymptotic tools based on general mean (WLLN and CLT) to do hypothesis testing. At last, we try to use GMM integrated with VOLS method to do VCAPM empirical analysis.

5.2 Comparison of VOLS with existing models

1. VOLS vs. OLS: θ can be regarded as weights, so it shares some characteristics of weighted OLS.

2. VOLS vs. WLS: In VOLS, the weights θ or $(1-\theta)$ is assigned through an implicit function. The dividing point or dividing line are the general mean or general conditional mean, they are exactly what we want to estimate. While WLS are being used to eliminate heteroscedasticity, the weights are given beforehand. If we use WLS to describe investors' pessimism or optimism view bias, we have to assign the weights artificially.

3. VOLS vs. Quantile regression: Both of them can describe the tail anomalies. Given θ, no information lost using VOLS method. After doing some modification, the analysis framework of VOLS is very similar to OLS.

5.3 VCAPM estimation

The reason we prefer to use general mean expression rather than expand it into ordinary mean with an adjustment item is that it makes the economic model more compact, and by any possibilities, we could deduce a terse expression of the conclusion. However, the estimating process is to search for the root of an implicit function. Not only is it time consuming, but also the existing asymptotic tools, such as LLN, CLT cannot be directly used any more. The small sample properties does not exists, however the large sample properties including consistency and asymptotic normality can be proved.

The VOLS estimator is

$$\begin{aligned}\hat{\beta} &= \left(\mathrm{X}'\hat{\Pi}_X \mathrm{X}\right)^{-1} \mathrm{X}'\hat{\Pi}_X \hat{\Pi}_{Y|X} Y \\ &= \left(\frac{\mathrm{X}'\hat{\Pi}_X \mathrm{X}}{n}\right)^{-1} \frac{\mathrm{X}'\hat{\Pi}_X \hat{\Pi}_{Y|X} Y}{n}\end{aligned} \tag{32}$$

Where,

$$\hat{\Pi}_X = \begin{pmatrix} \hat{\pi}_X & & 0 \\ & \ddots & \\ 0 & & \hat{\pi}_X \end{pmatrix}_{n\times n}, \quad \hat{\Pi}_{Y|X} = \begin{pmatrix} \hat{\pi}_{Y|X} & & 0 \\ & \ddots & \\ 0 & & \hat{\pi}_{Y|X} \end{pmatrix}_{n\times n} \tag{33}$$

And,

$$\hat{\pi}_{Y|X} = \frac{\theta_{sign(\hat{\beta}_1)(Y_t-\hat{g}(X_t))>0} + (1-\theta)_{sign(\hat{\beta}_1)(Y_t-\hat{g}(X_t))<0}}{\sum_{t=1}^{n}\left((\theta_{X_t>\hat{q}} + \theta_{X_t<\hat{q}})(\theta_{sign(\hat{\beta}_1)(Y_t-\hat{g}(X_t))>0} + (1-\theta)_{sign(\hat{\beta}_1)(Y_t-\hat{g}(X_t))<0})\right) \Big/ \sum_{t=1}^{n}\left(\theta_{X_t>\hat{q}} + (1-\theta)_{X_t<\hat{q}}\right)} \tag{34}$$

We develop the hypothesis testing by the case of conditional homoskedasticity and heteroskedasticity. We test the unconditional VCAPM. The unconditional refers that $\beta_t = \beta$ is time invariant. There is only one equation, but with two parameters to be estimated. Fortunately, we know that the unconditional VCAPM is just a special case of conditional VCAPM. Therefore, we solve the problem by estimating an unconditional VCAPM model using conditional GMM, which is restricted by VOLS. The econometrics model is as follows,

$$E(\pi_{r_{i,t}} r_{i,t}^e - \beta_0 - \beta_1 \pi_{r_{M,t}} r_{M,t}^e \mid I_t) = 0 \tag{35}$$

$$\text{St: } E_\theta^2(x_k \mid x_j) = \alpha_0 + \alpha_1 x_j \quad where \quad x_j, x_k \in \{r_{i,t}^e, r_{M,t}^e, r_{i,t-1}^e, r_{M,t-1}^e\}$$

From model (35), we get model (36),

$$\Rightarrow \begin{cases} E(\pi_{r_{i,t}^e} r_{i,t}^e - \beta_0 - \beta_1 \pi_{r_{M,t}} r_{M,t}^e) = 0 \\ E(\pi_{r_{i,t-1}^e} \pi_{r_{i,t}^e | r_{i,t-1}^e} r_{i,t}^e r_{i,t-1}^e - \beta_0 \pi_{r_{i,t-1}^e} r_{i,t-1}^e - \beta_1 \pi_{r_{i,t-1}^e} \pi_{r_{M,t}^e | r_{i,t-1}^e} r_{M,t}^e r_{i,t-1}^e) = 0 \\ E(\pi_{r_{M,t-1}^e} \pi_{r_{i,t}^e | r_{M,t-1}^e} r_{i,t}^e r_{M,t-1}^e - \beta_0 \pi_{r_{M,t-1}^e} r_{M,t-1}^e - \beta_1 \pi_{r_{M,t-1}^e} \pi_{r_{M,t}^e | r_{M,t-1}^e} r_{M,t}^e r_{M,t-1}^e) = 0 \end{cases} \tag{36}$$

$$\text{St: } E_\theta^2(x_k \mid x_j) = \alpha_0 + \alpha_1 x_j \quad where \quad x_j \in \{r_{i,t-1}^e, r_{M,t-1}^e\}, x_k \in \{r_{i,t}^e, r_{M,t}^e\}$$

Now, the model is overidentified, thus we can do hypothesis test. The null hypothesis is

$$H_0 : \beta_0 = 0 \tag{37}$$

The mathematical derivation is rather lengthy. Considering that we are not concentrating on the econometrics methodology in this paper, we just present the proof of weak law of large numbers (WLLN) of general mean for i.i.d Samples in appendix VI, and omit the other econometrics proofs.

6. Empirical Study

Firstly, we solve equation

$$\mu_P + \frac{D_t}{P_t} + \frac{\sigma_P}{\sqrt{dt}}\Theta - r^f = \alpha\sigma_C\sigma_P\left(\Theta^2\sqrt{1-\rho_{CP}^2} \times sign(\rho_{CP}) + \Phi\rho_{CP})\right) \tag{38}$$

and get people's view bias. The daily market return including distribution from NYSE, AMEX or NASDAQ is obtained from CRSP. We get the monthly average of market return, which includes the distribution, and get the volatility of market return excluding distribution. Then the annualized volatility is obtained by multiplying $\sqrt{252}$. The 30 T-bill from CRSP divided by 30.4 is being used as risk free rate. The U.S. quarterly aggregate nondurable goods and service consumption per capita is obtained from John Compbell's website. We interpolate it into a monthly time series, and use Gauss-Laguerre quadrature to approximate the function of Θ and Φ with respect to θ. The convergence tolerance is E-7; the maximum number of iteration is less than seven. That performance indicates that it is a practical method in econometrics. After doing θ's sensitivity analysis, see figure 4, we find that the greater from view neutral the deviation is, the more sensitive the general mean is, with respect to view bias.

<INSERT FIGURE 4 ABOUT HERE>

We assume that the risk aversion is constant 3, and then the monthly view bias time series could be figured out, see figure 5. The data range is from May.1926 to Sep.1999. The mean is 0.497, a slight deviation from view neutral, the maximum and minimum are 0.528 and 0.467 respectively. We draw the view bias periodogram within the time span from Jun. 1963 to Dec. 1970. The periodicity is around 60 months, Please see figure 6.

<INSERT FIGURE 5 ABOUT HERE>

<INSERT FIGURE 6 ABOUT HERE>

Secondly, we work on ranking each stock in deciles. We run the monthly common stock returns with share code 10/11 from NYSE, AMEX and NASDAQ. The sample range is from Jun.1963 to Dec. 1970. Following many studies, the ranking periods have the length of six months. For any given month, the rank of a certain stock is determined based on the past 6-month returns. A difference from the previous studies is that instead of constructing momentum portfolios, we select some representative stocks randomly according to the permanent company number, and then do the spectral analysis. We get the periodicities of each selected stock.

Thirdly, we make comparisons between the periodicities of view bias and each representative common stock. We find that quite many of them are compatible. 13 of 20 stocks have a distinct periodicity, which are either around 30 or 60 months, namely the half of or equal to the periodicity of view bias. The result supports the hypothesis

$$\beta^{\theta} := \frac{\Theta^2 \tilde{\sigma}_{iM} + \Phi \sigma_{iM}}{\Theta^2 \tilde{\sigma}_M^2 + \Phi \sigma_M^2} = \beta^{1-\theta} \tag{39}$$

Please see figure 7.

<INSERT FIGURE 7 ABOUT HERE>

Fourthly, we solve equation (38), but assume a constant neutral view bias; thereafter we get the risk preference time series with mean of 309 and median of 30. That is the equity premium puzzle. We plot a scatter diagram of view bias and risk preference, and then group the data set by the sign of correlations ρ_{CP}. We run regression analysis by group only keeping the records whose risk aversion is within the range of (-300, 300). The result indicates that they are strongly correlated for both two groups. The risk preference and view bias are negative correlated when ρ_{CP} is positive, and are positive correlated when ρ_{CP} is negative. In other words, if the market return and consumption growth move in the same direction, a risk averse investor is more like a pessimistic investor. If the market return and consumption growth move in a counter-direction, a risk averse investor is more like an optimistic investor.

<INSERT FIGURE 8 ABOUT HERE>

Finally, we take average of the implied view bias into VCAPM to estimate the view bias based beta, and to test if the intercept is zero. We use GMM method with restriction to VOLS. The reason why we select the stock (PERMNO 27705) whose name is INTERNATIONAL RECTIFIER CORP as sample individual security is because its' periodicity is compatible with what view bias suggests. So we consider that its' excess return is influenced by the movement of view bias. The sample range is from July 1963 to Nov.1965, total 30 months. After eliminating five abnormal records, we run classical OLS regression to estimate CAPM model (model A). The result indicates that there is no linear relationship between the excess return of individual security and market index. The R square is only 0.11, and t-test of both the intercept and slope could not be passed. That is how the traditional CAPM model being challenged. The average of implied view bias during the sample period is 0.494. We run four VOLS regressions.

$$E_{\theta}^{2}(x_k \mid x_j) = \alpha_0 + \alpha_1 x_j \quad where \quad x_j \in \{r_{i,t-1}^e, r_{M,t-1}^e\}, x_k \in \{r_{i,t}^e, r_{M,t}^e\} \tag{40}$$

Only those two whose regressand is the market excess return show us some linear relationship.

$$E_{\theta}^{2}(r_{M,t}^e \mid r_{i,t-1}^e) = \alpha_0 + \alpha_1 r_{i,t-1}^e; \quad E_{\vartheta}^{2}(r_{M,t}^e \mid r_{M,t-1}^e) = \alpha_0 + \alpha_1 r_{M,t-1}^e \tag{41}$$

We estimate GMM in two ways; the first one (model B) is from model (36)

$$\Rightarrow \begin{cases} E(\pi_{r_{i,t}^e} r_{i,t}^e - \beta_0 - \beta_1 \pi_{r_{M,t}} r_{M,t}^e) = 0 \\ E(\pi_{r_{i,t-1}^e} \pi_{r_{i,t}^e \mid r_{i,t-1}^e} r_{i,t}^e r_{i,t-1}^e - \beta_0 \pi_{r_{i,t-1}^e} r_{i,t-1}^e - \beta_1 \pi_{r_{i,t-1}^e} \pi_{r_{M,t}^e \mid r_{i,t-1}^e} r_{M,t}^e r_{i,t-1}^e) = 0 \\ E(\pi_{r_{M,t-1}^e} \pi_{r_{i,t}^e \mid r_{M,t-1}^e} r_{i,t}^e r_{M,t-1}^e - \beta_0 \pi_{r_{M,t-1}^e} r_{M,t-1}^e - \beta_1 \pi_{r_{M,t-1}^e} \pi_{r_{M,t}^e \mid r_{M,t-1}^e} r_{M,t}^e r_{M,t-1}^e) = 0 \end{cases} \tag{42}$$

The second (model C) is as follows,

$$\Rightarrow \begin{cases} E(\pi_{r_{i,t}^e} r_{i,t}^e - \beta_0 - \beta_1 \pi_{r_{M,t}} r_{M,t}^e) = 0 \\ E(r_{i,t}^e r_{i,t-1}^e - \beta_0 \pi_{r_{i,t-1}^e} r_{i,t-1}^e - \beta_1 \pi_{r_{i,t-1}^e} \pi_{r_{M,t}^e \mid r_{i,t-1}^e} r_{M,t}^e r_{i,t-1}^e) = 0 \\ E(r_{i,t}^e r_{M,t-1}^e - \beta_0 \pi_{r_{M,t-1}^e} r_{M,t-1}^e - \beta_1 \pi_{r_{M,t-1}^e} \pi_{r_{M,t}^e \mid r_{M,t-1}^e} r_{M,t}^e r_{M,t-1}^e) = 0 \end{cases} \tag{43}$$

Because the following two VOLS regression is not statistically significant,

$$E_{\theta}^{2}(r_{i,t}^e \mid r_{i,t-1}^e) = \alpha_0 + \alpha_1 r_{i,t-1}^e; \quad E_{\vartheta}^{2}(r_{i,t}^e \mid r_{M,t-1}^e) = \alpha_0 + \alpha_1 r_{M,t-1}^e \tag{44}$$

we are trying to neglect the terms of $\pi_{r_{i,t-1}^e} \pi_{r_{i,t}^e \mid r_{i,t-1}^e}$ as well as $\pi_{r_{M,t-1}^e} \pi_{r_{i,t}^e \mid r_{M,t-1}^e}$ in model (36), and get model C. Fortunately, both of model B and C show us a statistically significant view bias based beta, p-values are no more than 0.0001. More important, there is no evidence indicates that we should reject the zero intercept null hypothesis. The p-value of intercept B and C is 0.015 and 0.002 respectively.

<INSERT TABLE 5 ABOUT HERE>
<INSERT FIGURE 9 ABOUT HERE>
<INSERT TABLE 6 ABOUT HERE>
<INSERT TABLE 7 ABOUT HERE>

7. Conclusion

From the perspective of the mathematical model, after view bias adjustment, the general expectation of excess return can still be described in a single beta representation, except the systematic risk is the weighted average of exposed risk and potential risk. Empirical study indicates that in post war US, because of people's fear of unknowing, there exists a great discrepancy between the predicted return and what the consumption suggests, The momentum phenomena can be explained by the fact that stock ranking is affected by view bias movements. Quite many of the stocks share a compatible periodicity with view bias. From the econometrics perspective, VCAPM needs to be estimated through GMM method, which is restricted to VOLS. The empirical study provides us an illustration, for the same data set, although it cannot be described by CAPM, however sometimes it could be well captured by VCAPM. We take it as the first light of this research.

Appendix I:

1. General mean can be rewritten as

$$E_\theta^1(X) := q^* = \int \pi_X(\theta) x f_X(x)dx \tag{45}$$

Where

$$\pi_X(\theta) := \frac{(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx} \tag{46}$$

Proof.

Definition [General mean]: $E_\theta^1(X) := \underset{q}{\operatorname{argmin}}\left[(1-\theta)\int_{X<q}(x-q)^2 f_X(x)dx + \theta\int_{X>q}(x-q)^2 f_X(x)dx\right]$

Where X is absolute-integrable, i.e, $\int |x| f_X(x)dx < \infty$.

Leibnitz's Rule for improper integral:

$\frac{d}{d\varphi}\int_{a(\varphi)}^{+\infty} f(x,\varphi)dx = \int_{a(\varphi)}^{+\infty}\frac{\partial}{\partial\varphi} f(x,\varphi)dx - f(a(\varphi),\varphi)\frac{d}{d\varphi}a(\varphi)$ is valid for φ in $[a,b]$ when f and $\frac{\partial f}{\partial\varphi}$ are continuous for all $x \ge a(\varphi)$ and $a \le \varphi \le b$, and $a'(\varphi)$ is continuous in $[a,b]$, and when also there exists a function $M(x)$ independent of φ, such that $\left|\frac{\partial f(\varphi,x)}{\partial x}\right| \le M(x)$ for all $x \ge a(\varphi)$ and $a \le \varphi \le b$, and such that the integral $\int_{a(\varphi)}^{+\infty} M(x)dx$ converges.

It can be proved that the theory remain valid if interval $[a,b]$ for φ is replaced by an arbitrary interval, finite, infinite, open or closed or neither.

To take the first order derivative, using the above Leibnitz's rule for the following improper integral, we get

$$\begin{aligned}
&\frac{d\left((1-\theta)\int_{X<q}(x-q)^2 f_X(x)dx + \theta\int_{X>q}(x-q)^2 f_X(x)dx\right)}{dq} \\
&= \frac{d\left(\int_{-\infty}^{q}(1-\theta)(x-q)^2 f_X(x)dx\right)}{dq} + \frac{d\left(\int_{q}^{+\infty}\theta(x-q)^2 f_X(x)dx\right)}{dq} \\
&= \left((1-\theta)(q-q)^2 f_X(q)\frac{dq}{dq} + \int_{-\infty}^{q}\frac{d\left((1-\theta)(x-q)^2 f_X(x)\right)}{dq}dx\right) + \left(-\theta(q-q)^2 f_X(q)\frac{dq}{dq} + \int_{q}^{+\infty}\frac{d\left(\theta(x-q)^2 f_X(x)\right)}{dq}dx\right) \\
&= \int_{-\infty}^{q}\frac{d\left((1-\theta)(x-q)^2 f_X(x)\right)}{dq}dx + \int_{q}^{+\infty}\frac{d\left(\theta(x-q)^2 f_X(x)\right)}{dq}dx \\
&= 0
\end{aligned}$$

$$q^*\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx = \int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx$$

$$E_\theta^1(X) := q^* = \frac{\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx} \quad (47)$$

If we can verify that for all $x \in (-\infty, q^*)$ and $q^* \in (-\infty, +\infty)$, there exits a function of $M_1(x)$

$$\left|\frac{d(1-\theta)(x-q)^2 f_X(x)}{dx}\right| = \left|(1-\theta)\left(2(x-q) f_X(x) + (x-q)^2\frac{df_X(x)}{dx}\right)\right| \le M_1(x) \quad (48)$$

and such that the integral $\int_{-\infty}^{q^*} M_1(x)dx$ converges, and for all $x \in (q^*, +\infty)$ and $q^* \in (-\infty, +\infty)$, there exists a function of $M_2(x)$

$$\left|\frac{d\theta(x-q)^2 f_X(x)}{dx}\right| = \left|\theta\left(2(x-q) f_X(x) + (x-q)^2\frac{df_X(x)}{dx}\right)\right| \le M_2(x) \quad (49)$$

and such that the integral $\int_{q^*}^{+\infty} M_2(x)dx$ converges, then definition (47) is proved.

To see this, first setting out from an optimization

$$E(X) := \arg\min_q \left[\int_{-\infty}^{+\infty} (x-q)^2 f_X(x)dx\right], \tag{50}$$

where X is absolute-integrable, i.e, $\int |x| f_X(x)dx < \infty$, we can get the definition of mean as follows

$$\frac{d\left(\int_{-\infty}^{+\infty}(x-q^*)^2 f_X(x)dx\right)}{dq^*} = \int_{-\infty}^{+\infty}\frac{d\left((x-q^*)^2 f_X(x)\right)}{dq^*}dx = \int_{-\infty}^{+\infty}2(x-q^*)f_X(x)dx = 0 \tag{51}$$

That is $q^* = \int_{-\infty}^{+\infty} x f_X(x)dx$. Then we know that there must be a $M(x)$ where

$$\left|\frac{d(x-q)^2 f_X(x)}{dx}\right| = \left|2(x-q)f_X(x) + (x-q)^2\frac{df_X(x)}{dx}\right| \le M(x) \tag{52}$$

for all $x \in (-\infty, +\infty)$, such that the integral $\int_{-\infty}^{+\infty} M(x)dx$ converges. Let $M_1(x) \equiv M(x)$, then it is easy to prove that

$$\left|\frac{d(1-\theta)(x-q)^2 f_X(x)}{dx}\right| = \left|(1-\theta)\left(2(x-q)f_X(x) + (x-q)^2\frac{df_X(x)}{dx}\right)\right| \le M_1(x) \tag{53}$$

Since $\int_{-\infty}^{+\infty} M_1(x)dx < \infty$ and $M_1(x) \ge 0$, for any $q^* \in (-\infty, +\infty)$, $\int_{-\infty}^{q^*} M_1(x)dx$ converges as well. Following the same logic, Let $M_2(x) \equiv M(x)$, it is easy to prove that $\int_{q^*}^{+\infty} M_2(x)dx$ converges, then the general mean is well defined as follows.

$$E_\theta^1(X) := q^* = \frac{\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx} \tag{54}$$

It is proved.

2. Proposition 2.1.1: $E_\theta^1(X)$ is a monotonously increasing function of θ.

Proof.

We know that setting out from optimization

$$Quantile_\theta(X) \equiv \arg\min_q \left[\int_{-\infty}^{+\infty} |x-q| f_X(x)dx\right] \tag{55}$$

$$\frac{d\left[\int_{-\infty}^{q^*}(q^*-x)f_X(x)dx + \int_{q^*}^{+\infty}(x-q^*)f_X(x)dx\right]}{dq^*} = 0 \tag{56}$$

$$\int_{-\infty}^{q^*} f_X(x)dx - \int_{q^*}^{+\infty} f_X(x)dx = 0$$

$$F_X(q) - [1-F_X(q)] = 0$$

$$F_X(q) = 50\% \tag{57}$$

We get the definition of quantile as $q = F_X^{-1}(50\%)$. We know that there must exits $M(x)$, where

$$\left|\frac{d[(q^*-x)f_X(x)]}{dx}\right| = \left|\left(-f_X(x) + (q^*-x)\frac{df_X(x)}{dx}\right)\right| \le M(x) \tag{58}$$

such that $\int_{-\infty}^{+\infty} M(x)dx < \infty$ converges.

Now we claim that for all $x \in (-\infty, q^*)$ and $q^* \in (-\infty, +\infty)$, there exits a function of $M_1(x)$

$$\left|\frac{d((1-\theta)xf_X(x))}{dx}\right| = \left|(1-\theta)\left(f_X(x) + x\frac{df_X(x)}{dx}\right)\right| \le M_1(x) \tag{59}$$

and such that the integral $\int_{-\infty}^{q^*} M_1(x)dx$ converges, and for all $x \in (q^*, +\infty)$ and $q^* \in (-\infty, +\infty)$, there exists a function of $M_2(x)$,

$$\left|\frac{d(\theta xf_X(x)}{dx}\right| = \left|\theta\left(f_X(x) + x\frac{df_X(x)}{dx}\right)\right| \le M_2(x) \tag{60}$$

and such that the integral $\int_{q^*}^{+\infty} M_2(x)dx$ converges.

To see this, just let $M_1(x) = M_2(x) \equiv M(x)$, then it is easy to be verified.

And because

$$\left|\frac{d((1-\theta)f_X(x)}{dx}\right| = \left|(1-\theta)\frac{df_X(x)}{dx}\right| \le \left|\frac{df_X(x)}{dx}\right| \equiv N(x) \tag{61}$$

$$\left|\frac{d(\theta f_X(x))}{dx}\right| = \left|\theta\frac{df_X(x)}{dx}\right| \le \left|\frac{df_X(x)}{dx}\right| \equiv N(x) \tag{62}$$

$$\int_{-\infty}^{+\infty} N(x)dx = \int_{-\infty}^{+\infty}\left|\frac{df_X(x)}{dx}\right|dx = \int_{-\infty}^{\text{mod}}\frac{df_X(x)}{dx}dx - \int_{\text{mod}}^{+\infty}\frac{df_X(x)}{dx}dx = 2f_X(\text{mod}) < \infty \tag{63}$$

Since $N(x) \ge 0$, for any $q^* \in (-\infty, +\infty)$,

$$\int_{-\infty}^{q^*} N(x)dx < \infty \text{, and } \int_{q^*}^{+\infty} N(x)dx < \infty \tag{64}$$

Now we are ready to prove $E_\theta^1(X)$ is a monotonously increasing function of θ

Since

$$E_\theta^1(X) := q^* = \frac{\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx} \tag{65}$$

We get

$$q^* - \frac{\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx} = 0 \qquad (66)$$

$$\frac{dq^*}{d\theta} - \frac{d}{d\theta}\left(\frac{\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx}\right)$$

$$= \frac{dq^*}{d\theta} - \frac{1}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)}\frac{d}{d\theta}\left(\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right) - \left(\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)\frac{d}{d\theta}\left(\frac{1}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)}\right)$$

$$= \frac{dq^*}{d\theta} - \frac{(1-2\theta)q^* f_X(q^*)\frac{dq^*}{d\theta} + \int x[1_{X>q^*} - 1_{X<q^*}] f_X(x)dx}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)} + \frac{\left(\int x[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)^2} \times \left((1-2\theta) f_X(q^*)\frac{dq^*}{d\theta} + \int [1_{X>q^*} - 1_{X<q^*}] f_X(x)dx\right)$$

$$= \frac{dq^*}{d\theta} - \frac{(1-2\theta)q^* f_X(q^*)\frac{dq^*}{d\theta} + \int x[1_{X>q^*} - 1_{X<q^*}] f_X(x)dx}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)} + \frac{q^*}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)} \times \left((1-2\theta) f_X(q^*)\frac{dq^*}{d\theta} + \int [1_{X>q^*} - 1_{X<q^*}] f_X(x)dx\right)$$

$$= 0$$

$$\frac{dq^*}{d\theta} = \frac{\int (x-q^*)[1_{X>q^*} - 1_{X<q^*}] f_X(x)dx}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)} = \frac{\int |x-q^*| f_X(x)dx}{\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)} > 0 \qquad (67)$$

Proposition 2.1.1 is proved.

3. Proposition 2.2.1: for any probability densities f and h, if $h_X(x) = \pi_X(\theta) f_X(x)$, where $\pi_X(\theta)$ defined as formula (3) then

$$\frac{dD(f \| h)}{d\theta} := \begin{cases} > 0 & \theta > 50\% \\ = 0 & \theta = 50\% \\ < 0 & \theta < 50\% \end{cases} \qquad (68)$$

Proof.

$$
\begin{aligned}
\frac{dD(f\,\|\,h)}{d\theta} &= \frac{d}{d\theta}[\int f_X(x)\ln\left(\frac{f_X(x)}{h_X(x)}\right)dx] = \frac{d}{d\theta}[\int f_X(x)\ln\left(\frac{f_X(x)}{\pi_X(\theta)f_X(x)}\right)dx] \\
&= \frac{d}{d\theta}[\int -f_X(x)\ln\pi_X(\theta)dx] = -\left(\int f_X(x)\frac{1}{\pi_X(\theta)}\frac{d}{d\theta}[\pi_X(\theta)]dx\right) - \frac{f_X(q^*)\ln(1-\theta)}{scale}\frac{dq^*}{d\theta} + \frac{f_X(q^*)\ln(\theta)}{scale}\frac{dq^*}{d\theta} \\
&= -\left(\int f_X(x)\frac{1}{\pi_X(\theta)}\frac{d}{d\theta}[\frac{(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}}{\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx}]dx\right) + \frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right) \\
&= -\left(\int f_X(x)\frac{1}{\pi_X(\theta)}\left(\frac{1_{X>q^*}-1_{X<q^*}}{\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx} - \frac{\left((1-\theta)1_{X<q^*}+\theta 1_{X>q^*}\right)\int[(1_{X>q^*}-1_{X<q^*}]f_X(x)dx}{\left(\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx\right)^2}\right)dx\right) + \frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right) \\
&= -\int \frac{(1_{X>q^*}-1_{X<q^*})f_X(x)}{(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}}dx + \int \frac{(1_{X>q^*}-1_{X<q^*})f_X(x)}{\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx}dx + \frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right) \\
&= \int (1_{X>q^*}-1_{X<q^*})f_X(x)\left(\frac{1}{\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx} - \frac{1}{(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}}\right)dx + \frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right) \\
&= \int_{X>q^*}\left(\frac{f_X(x)}{\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx}\right)dx - \int_{X<q^*}\left(\frac{f_X(x)}{\int[(1-\theta)1_{X<q^*}+\theta 1_{X>q^*}]f_X(x)dx}\right)dx + \int_{X<q^*}\left(\frac{f_X(x)}{1-\theta}\right)dx - \int_{X>q^*}\left(\frac{f_X(x)}{\theta}\right)dx + \frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right)
\end{aligned}
$$

If $\theta \geq 50\%$, then

$$(1-\theta) \leq \int[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx \leq \theta \tag{69}$$

$$\frac{1}{\theta} \leq \frac{1}{\int[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx} \leq \frac{1}{1-\theta} \tag{70}$$

$$\frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right) \geq 0 \tag{71}$$

If $\theta \leq 50\%$, then

$$\theta \leq \int[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx \leq (1-\theta) \tag{72}$$

$$\frac{1}{1-\theta} \leq \frac{1}{\int[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]f_X(x)dx} \leq \frac{1}{\theta} \tag{73}$$

$$\frac{f_X(q^*)}{scale}\frac{dq^*}{d\theta}\ln\left(\frac{\theta}{1-\theta}\right) \leq 0 \tag{74}$$

Therefore, we get the following result

$$\frac{dD(f\,\|\,h)}{d\theta} = \begin{cases} >0 & \theta > 50\% \\ =0 & \theta = 50\% \\ <0 & \theta < 50\% \end{cases} \tag{75}$$

Proposition 2.2.1 is proved.

4. Proposition 2.2.2: $0 \leq D(f\,\|\,h_{general\ \ mean}) \leq D(f\,\|\,h_{quantile}) = +\infty$ with the first equality if and only if $\theta = 50\%$.

Proof.

$$Quantile_\theta(X) := \arg\min_q \left[(1-\theta)\int_{X<q} |x-q|\, f_X(x)dx + \theta\int_{X>q} |x-q|\, f_X(x)dx\right] \quad (76)$$

$$\theta = F_X(q^*) = \int 1_{X<q^*} f_X(x)dx = 1 - \int 1_{X>q^*} f_X(x) \quad (77)$$

$$1 = \frac{F_X(q^*)}{\theta} = \int \left(\frac{1}{\theta}\right)_{X<q^*} f_X(x)dx = \frac{1}{\theta} - \int \left(\frac{1}{\theta}\right)_{X>q^*} f_X(x)dx \quad (78)$$

Setting out from quantile definition, we get the above formula. That formula is equivalent to take the expectation of 1 with the probability $\left(\frac{1}{\theta}\right)_{X<q^*} f_X(x)$ or $\left(\frac{1}{\theta}\right)_{X>q^*} f_X(x)$. From definition of relative entropy,

$$D(f \,\|\, h_q) = \int f_X(x)\ln\left(\frac{f_X(x)}{h_{qX}(x)}\right)dx = -\int f_X(x)\ln\left(\frac{1}{\theta}\right)_{X<q} dx = +\infty \quad (79)$$

That is also an illustration of the aforementioned theory, i.e., defining

$$B = Supp(f) = \{x : f_X(x) > 0\} \quad (80)$$

in either case if the support $Supp\ (f) \not\subseteq Supp\ (h)$, then $D(f \,\|\, h) = \infty$.

$$\begin{aligned}
D(f \,\|\, h_g) &= \int f_X(x)\ln\left(\frac{f_X(x)}{h_{gX}(x)}\right)dx = -\int f_X(x)\ln(\pi_X(\theta))dx \\
&= -\int f_X(x)\ln\left(\frac{(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}}{\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx}\right)dx \\
&= -\int f_X(x)\left(\ln[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] - \ln\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right)\right)dx \\
&= -\int f_X(x)\left(\ln[(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}]\right)dx + \ln\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right) \\
&= \ln\left(\int [(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}] f_X(x)dx\right) + \int f_X(x)\left(\ln[\frac{1}{(1-\theta)1_{X<q^*} + \theta 1_{X>q^*}}]\right)dx
\end{aligned} \quad (81)$$

Therefore, we have

$$D(f \,\|\, h_q) > D(f \,\|\, h_g) \quad (82)$$

Proposition 2.2.2 is proved.

Appendix II:

Proposition 2.3.2: if $X_i \sim N(\mu_i, \sigma_i)$, the correlation between X_i and X_j is ρ_{ij}, and $\exists \rho_{ij} \neq 1$ then

$$E_\theta^{n+}(\sum_{i=1}^n X_i) - E_\theta^1(\sum_{i=1}^n X_i) \begin{cases} > 0 & \theta > 50\% \\ = 0 & \theta = 50\% \\ < 0 & \theta < 50\% \end{cases} \quad (83)$$

Proof.

Case 1: If $\theta = 50\%$, then

$$E_{\theta}^{n+}(\sum_{i=1}^{n} X_i) - E_{\theta}^{1}(\sum_{i=1}^{n} X_i) = E(\sum_{i=1}^{n} X_i) - E(\sum_{i=1}^{n} X_i) = 0 \tag{84}$$

Case 2: If $\theta > 50\%$, and $\rho_{ij} \neq 1$

a. If $n = 2$, $X_i \sim N(\mu_i, \sigma_i)$, then we have

$$X_1 + X_2 \sim N(\mu_1 + \mu_2, {\sigma_1}^2 + {\sigma_2}^2 + 2\sigma_1\sigma_2\rho_{12}) \tag{85}$$

$$E_{\theta}^{2}(\sum_{i=1}^{2} X_i) - E_{\theta}^{1}(\sum_{i=1}^{2} X_i) = (\sigma_1 + \sigma_2)E_{\theta}^{1}[\omega] - E_{\theta}^{1}[\omega]\sqrt{{\sigma_1}^2 + {\sigma_2}^2 + 2\sigma_1\sigma_2\rho_{12}} \tag{86}$$

Since $\rho_{12} \neq 1$, and from proposition 2.1.1, $E_{\theta}^{1}[\omega]$ is a monotonously increasing function with respect to θ, $E_{\theta}^{1}[\omega] > 0$, thus

$$E_{\theta}^{2+}(\sum_{i=1}^{2} X_i) - E_{\theta}^{1}(\sum_{i=1}^{2} X_i) > 0 \tag{87}$$

b. Suppose $n = k$, then

$$E_{\theta}^{n+}(\sum_{i=1}^{n} X_i) - E_{\theta}^{1}(\sum_{i=1}^{n} X_i) > 0 \tag{88}$$

c. If $n = k + 1$, then

$$E_{\theta}^{n+}(\sum_{i=1}^{n} X_i) = E_{\theta}^{(k+1)+}(\sum_{i=1}^{k+1} X_i) = E_{\theta}^{k+}(\sum_{i=1}^{k} X_i) + E_{\theta}^{1}(X_{k+1}) > E_{\theta}^{1}(\sum_{i=1}^{k} X_i) + E_{\theta}^{1}(X_{k+1}) > E_{\theta}^{1}(\sum_{i=1}^{k+1} X_i) = E_{\theta}^{1}(\sum_{i=1}^{n} X_i) \tag{89}$$

Case 3: If $\theta < 50\%$ and $\rho_{ij} \neq 1$, then $E_{\theta}^{1}[\omega] < 0$, ceteris paribus, we therefore get

$$E_{\theta}^{n+}(\sum_{i=1}^{n} X_i) < E_{\theta}^{1}(\sum_{i=1}^{n} X_i) \tag{90}$$

Case4: If $\forall \rho_{ij} = 1$ then

$$E_{\theta}^{n+}(\sum_{i=1}^{n} X_i) - E_{\theta}^{1}(\sum_{i=1}^{n} X_i) = nE_{\theta}^{1}(X) - nE_{\theta}^{1}(X) = 0 \tag{91}$$

So, if $X_i \sim N(\mu_i, \sigma_i)$, and $\exists \rho_{ij} \neq 1$, we obtain

$$E_{\theta}^{n+}(\sum_{i=1}^{n} X_i) - E_{\theta}^{1}(\sum_{i=1}^{n} X_i) \begin{cases} > 0 & \theta > 50\% \\ = 0 & \theta = 50\% \\ < 0 & \theta < 50\% \end{cases} \tag{92}$$

Proposition 2.3.2 is proved.

Appendix III:

Here we are going to explain, within general expectation framework, how to construct two standard normal distributions X, Y whose correlation is ρ. When perfect information exists, usually Y is structured as follows,

$$X, Z \overset{iid}{\sim} N(0,1); \quad Y = \rho X + \sqrt{1-\rho^2} Z \tag{93}$$

We can prove that no matter $\rho > 0$ or $\rho < 0$, the character of Y, that is $Y \sim N(0,1)$ is maintained, and $\rho_{XY} = \rho$, the point we want to carry, is satisfied.

But in θ -adjusted framework, $\rho > 0$ or $\rho < 0$ matters. For the maintenance of mathematical character needs the sign of the second term in accordance with the sign of ρ. In other words, it should be $Y = \rho X + sign(\rho) \times \sqrt{1-\rho^2} Z$. The reason is as follows,

Case1: $\rho > 0$

Suppose $X, Z \overset{iid}{\sim} N(0,1)$, Y is structured as $Y = \rho X + \sqrt{1-\rho^2} Z$, then $Y \sim N(0,1)$, we get

$$E_\theta^1(Y) = \Theta \tag{94}$$

$$E_\theta^2(\rho X + \sqrt{1-\rho^2} Z) = E_\theta^1(\rho X) + E_\theta^1(\sqrt{1-\rho^2} Z) = \rho E_\theta^1(X) + \sqrt{1-\rho^2} E_\theta^1(Z) = \rho\Theta + \sqrt{1-\rho^2}\Theta \tag{95}$$

If Y is structured as $Y = \rho X - \sqrt{1-\rho^2} Z$, then still $Y \sim N(0,1)$, we get

$$E_\theta^1(Y) = \Theta \tag{96}$$

$$E_\theta^2(\rho X - \sqrt{1-\rho^2} Z) = E_\theta^1(\rho X) - E_\theta^1(\sqrt{1-\rho^2} Z) = \rho E_\theta^1(X) - \sqrt{1-\rho^2} E_\theta^1(Z) = \rho\Theta - \sqrt{1-\rho^2}\Theta \tag{97}$$

In the general expectation framework, we need to maintain more character other than $Y \sim N(0,1)$. If X and Y are positively correlated, people's view bias are concordant. That is if he is pessimistic, he tends to amplify the left tails of X and Y, and vise versa. So one dimension distribution of $(\rho X \pm \sqrt{1-\rho^2} Z)$ and the joint distribution of X and Y are coherent. $E_\theta^1(Y)$ is the expectation of Y, $E_\theta^2(\rho X \pm \sqrt{1-\rho^2} Z)$ is the expectation of $(\rho X \pm \sqrt{1-\rho^2} Z)$, which is integrated by the joint probability of X and Z. From section2.3, if there is risk-source dimensional receding, additivity is not satisfied. And it is sure enough that $(\rho \pm \sqrt{1-\rho^2}) \neq 1$, unless $\rho = 1$, and joint distribution recedes to one dimension. However, according to the monotonously increasing property of general mean, they should be of the same sign. Since both of them are being used to measure the reward of an uncertainty. Obviously,

$$Y = (\rho X + \sqrt{1-\rho^2} Z) \tag{98}$$

should be the only choice.

Case2: $\rho = -1$

$Y = -X$, equation $E_\theta^1(Y) = E_\theta^1(X) = -E_\theta^1(-X)$ exists. Now we are explaining it. Since $X \sim N(0,1)$, normal distribution is symmetry, $Y \sim N(0,1)$, so we obtain, $E_\theta^1(X) = \Theta$ as well as $E_\theta^1(Y) = \Theta$. From the property Homogeneity in section2, we obtain $E_\theta^1(-X) = -E_\theta^1(X)$. It is amazing and confusing. But the following equation shows that the main problem lies in $\rho = -1$, people's view bias are not concordant any more.

$$E_\theta^1(Y) = \int_{-\infty}^{\infty} y\pi_y(\theta) f(y)dy = \int_{-\infty}^{\infty} -x\pi_{-x}(\theta) f(-x)d(-x) = \int_{-\infty}^{\infty} -w\pi_w(1-\theta) f(w)dw = -(-\Theta) = \Theta = E_\theta^1(X) \tag{99}$$

Looking at X while thinking of Y is not the same as looking at Y itself. That is why a joke says that the pessimistic father worry about his elder son when the sun is shining and his younger son when it is raining, because the former is selling umbrella, and the latter is a fisher.

Case3: $\rho < 0, \quad \rho \neq -1$

$$E_{\theta}^{1}(Y) = -[-\rho E_{\theta}^{1}(X) + \sqrt{1-(-\rho)^2}\, E_{\theta}^{1}(Z)] = \rho\Theta - \sqrt{1-\rho^2}\,\Theta \tag{100}$$

$$Y = (\rho X - \sqrt{1-\rho^2}\, Z) \tag{101}$$

Combine equation (98) and (101),

$$Y = \rho X + sign(\rho) \times \sqrt{1-\rho^2}\, Z \tag{102}$$

We can see that the risk of Y is generated from two parts. One is correlated with X , the other part is uncorrelated with X . In other words, when variance is standardized to 1, risk ρ is exposed to X , $sign(\rho)\sqrt{1-\rho^2}$ is exposed to Z . Since the second part of power does exist, and taken into formula A.I-20, term $\sigma_X \sigma_Y sign(\rho_{XY}) \times \sqrt{1-\rho_{XY}^2}$ comes up, that is why we name this term as the potential risk of Y to X .

Appendix IV:

Next, we are to derive the VCAPM model.

$W(t) =$ Total wealth at time t
$P_i(t) =$ Price of the i^th asset at time t (i=1, ⋯, n)
$S_j(t) =$ Value of the j^th state variable at time t , (j=1, ⋯, m)
$C(t) =$ Consumption per unit time at time t
$w_i(t) =$ Proportion of total wealth in the i^th asset at time t (i=1, ⋯, n)

Note $[\sum_{i=1}^{n} w_i(t)] \equiv 1$.

Assumption 1: Time interval between each decision is infinitesimal.
Assumption 2: Prices follow diffusion processes.
Assumption 3: Only consumption and portfolio process are controllable.
Assumption 4: There is no exogenous endowment.
Assumption 5: Investors are homogenous.
Assumption 6: Information is imperfect, and pessimism or optimism view bias might exist.

We model the consumption and portfolio choosing process as follows,

$$J[W(t), S(t), t] \equiv \max_{\{C_{(\tau)}, w_{(\tau)}\}} E_{(\theta,\ t)}^{n+}\{\int_t^T U_1[C(\tau), \tau] d\tau + U_2[W(T), T]\} \tag{103}$$

St: boundary condition: $J[W(T),S(T),T]=U_2[W(T),T]$

budget equations: $$W(t)=\sum_{i=1}^{n}w_i(t_0)\frac{P_i(t)}{P_i(t_0)}[W(t_0)-C(t_0)h] \qquad (104)$$

assumption1: $t\equiv t_0+h,\quad h\to 0$

assumption2: $$\frac{dP_i(t)}{P_i(t)}=\mu_i(S,t)dt+\sigma_i(S,t)\sqrt{dt}\omega_i,\qquad i=1,2,\cdots,n$$

$$V_{(n\times n)}=[\sigma_{il}],\qquad \sigma_{il}=\sigma_i\sigma_l\rho_{il},\qquad i,l=1,2,\cdots,n$$

$$dS_j(t)=f_j(S,t)dt+g_j(S,t)\sqrt{dt}q_j,\qquad j=1,2,\cdots,m$$

$$\Omega_{(m\times m)}=[g_jg_k\eta_{jk}],\qquad j,k=1,2,\cdots,m$$

$$\Gamma_{(n\times m)}=[\varepsilon_{ij}],\qquad \varepsilon_{ij}=\sigma_ig_j\pi_{ij},\qquad i=1,2,\cdots,n;\qquad j=1,2,\cdots,m$$

From equation (103),

$$J[W(t_0),S(t_0),t_0]=\max_{\{C_{(\tau)},w_{(\tau)}\}}E^{n+}_{(\theta,\ t_0)}\{\int_{t_0}^{t}U_1[C(\tau),\tau]d\tau+J[W(t),S(t),t]\} \qquad (105)$$

Define $t\equiv t_0+h$, by Taylor's theorem and the mean value theorem for integrals, equation (105) can be rewritten as,

$$\begin{aligned}J[W(t_0),S(t_0),t_0]=&\max_{\{C_{(\tau)},w_{(\tau)}\}}E^{n+}_{(\theta,\ t_0)}\{U_1[C(\bar{t}),\bar{t}]h+J[W(t_0),S(t_0),t_0]+\frac{\partial J[W(t_0),S(t_0),t_0]}{\partial t}h\\&+\frac{\partial J[W(t_0),S(t_0),t_0]}{\partial W}[W(t)-W(t_0)]+\sum_{j=1}^{m}\frac{\partial J[W(t_0),S(t_0),t_0]}{\partial S_j}[S_j(t)-S_j(t_0)]\\&+\frac{1}{2}\frac{\partial^2 J[W(t_0),S(t_0),t_0]}{\partial W^2}[W(t)-W(t_0)]^2+\frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 J[W(t_0),S(t_0),t_0]}{\partial S_kS_j}[S_k(t)-S_k(t_0)][S_j(t)-S_j(t_0)]\\&+\sum_{j=1}^{m}\frac{\partial J[W(t_0),S(t_0),t_0]}{\partial W\partial S_j}[W(t)-W(t_0)][S_j(t)-S_j(t_0)]+O(h^2)\}\end{aligned} \qquad (106)$$

where $\bar{t}\in[t_0,t]$, take limit as $h\to 0$, take the θ-adjusted expectation operators onto each term, and subtracting $J[W(t_0),S(t_0),t_0]$ of both sides,

$$\begin{aligned}0=&\max_{\{C_{(t)},w_{(t)}\}}\{U_1[C(t),t]dt+\frac{\partial J[W(t),S(t),t]}{\partial t}dt\\&+\frac{\partial J[W(t),S(t),t]}{\partial W}E^{n+}_{(\theta,\ t)}[dW(t)]+\sum_{j=1}^{m}\frac{\partial J[W(t),S(t),t]}{\partial S_j(t)}E^{1}_{(\theta,\ t)}[dS_j(t)]\\&+\frac{1}{2}\frac{\partial^2 J[W(t),S(t),t]}{\partial W^2}E^{n+}_{(\theta,\ t)}[dW(t)]^2+\frac{1}{2}\sum_{k=1}^{m}\sum_{j=1}^{m}\frac{\partial^2 J[W(t),S(t),t]}{\partial S_k\partial S_j}E^{2}_{(\theta,\ t,\eta_{jk})}[dS_k(t)dS_j(t)]\\&+\sum_{j=1}^{m}\frac{\partial J[W(t),S(t),t]}{\partial W\partial S_j}E^{(n+1)+}_{(\theta,\ t,\Gamma_j)}[dW(t)dS_j(t)]+O(dt^2)\}\qquad \Gamma_j\ \textit{is the}\ j^{th}\ \textit{row of}\ \Gamma\ \textit{matrix}\end{aligned} \qquad (107)$$

Now we specify how to get each θ-adjusted expectation term. By subtracting $W(t_0)$ on both sides, the budget equation is rewritten as,

$$W(t)-W(t_0)=[\sum_{i=1}^{n} w_i(t_0)\frac{P_i(t)-P_i(t_0)}{P_i(t_0)}][W(t_0)-C(t_0)h]-C(t_0)h \tag{108}$$

The θ-adjusted expectation of the limit process as $h\to 0$ is

$$\begin{aligned} E^{n+}_{(\theta,\ t)}[dW(t)] &= \{\sum_{i=1}^{n} w_i(t)W(t)E^{1}_{(\theta,\ t)}[\frac{dP_i(t)}{P_i(t)}]-C(t)dt\}+O(dt^2) \\ &= \left\{\sum_{i=1}^{n}\left(w_i(t)W(t)[\mu_i(S,t)+\frac{\sigma_i(S,t)}{\sqrt{dt}}E^{1}_{(\theta,\ t)}(\omega_i)]\right)-C(t)\right\}dt \end{aligned} \tag{109}$$

Applying the same limit process to other terms,

$$\begin{aligned} [W(t)-W(t_0)]^2 &= \{[\sum_{i=1}^{n} w_i(t_0)\frac{P_i(t)-P_i(t_0)}{P_i(t_0)}][W(t_0)-C(t_0)h]-C(t_0)h\}^2 \\ &= \{\sum_{l=1}^{n}\sum_{i=1}^{n} w_i(t_0)w_l(t_0)\frac{P_i(t)-P_i(t_0)}{P_i(t_0)}\frac{P_l(t)-P_l(t_0)}{P_l(t_0)}(W(t_0)^2-2W(t_0)C(t_0)h+C(t_0)^2h^2) \\ &\quad -2\sum_{i=1}^{n} w_i(t_0)\frac{P_i(t)-P_i(t_0)}{P_i(t_0)}[W(t_0)-C(t_0)h]C(t_0)h+C(t_0)^2h^2\} \\ &= \sum_{l=1}^{n}\sum_{i=1}^{n} w_i(t_0)w_l(t_0)(\mu_i(S,t)\mu_j(S,t)h^2+\sigma_i(S,t)\sigma_j(S,t)h\omega_i\omega_j+\mu_i(S,t)\sigma_j(S,t)h\sqrt{h}\omega_j \\ &\quad +\mu_j(S,t)\sigma_i(S,t)h\sqrt{h}\omega_i)(W(t_0)^2-2W(t_0)C(t_0)h+C(t_0)^2h^2) \\ &\quad -2\sum_{i=1}^{n} w_i(t_0)[\mu_i(S,t)h+\sigma_i(S,t)\sqrt{h}\omega_i][W(t_0)C(t_0)h-C(t_0)^2h^2] \\ &= \sum_{l=1}^{n}\sum_{i=1}^{n}\{w_i(t_0)w_l(t_0)W(t_0)^2[\sigma_i(S,t)\sigma_j(S,t)h\omega_i\omega_j]\}+O(h^2) \end{aligned}$$

$$E^{n+}_{(\theta,\ t)}[dW(t)]^2=\sum_{l=1}^{n}\sum_{i=1}^{n}\{w_i(t_0)w_l(t_0)W(t_0)^2\sigma_i(S,t)\sigma_l(S,t)E^{2}_{(\theta,\ t,\rho_{il})}[\omega_i\omega_l]\}dt \tag{110}$$

$$E^{1}_{(\theta,\ t)}[dS_j(t)]=f_j(S,t)dt+g_j(S,t)\sqrt{dt}E^{1}_{(\theta,\ t)}(q_i) \tag{111}$$

$$\begin{aligned} E^{2}_{(\theta,\ t)}[dS_k(t)dS_j(t)] &= E^{2}_{(\theta,\ t)}\{[f_k(S,t)dt+g_k(S,t)\sqrt{dt}q_k(t)][f_j(S,t)dt+g_j(S,t)\sqrt{dt}q_j(t)]\} \\ &= g_k(S,t)g_j(S,t)E^{2}_{(\theta,\ t,\eta_{jk})}[q_k(t)q_j(t)]dt+O(dt^2) \end{aligned} \tag{112}$$

$$\begin{aligned} E^{(n+1)+}_{(\theta,\ t)}[dW(t)dS_j(t)] &= E^{(n+1)+}_{(\theta,\ t)}\left(\{\sum_{i=1}^{n} w_i(t)W(t)[\mu_i(S,t)dt+\sigma_i(S,t)\sqrt{dt}\omega_i]-C(t_0)dt\}[f_j(S,t)dt+g_j(S,t)\sqrt{dt}q_j]\right) \\ &= [\sum_{i=1}^{n} w_i(t)W(t)\sigma_i(S,t)g_j(S,t)E^{2}_{(\theta,\ t,\pi_{ij})}(\omega_i q_j)]dt \end{aligned} \tag{113}$$

Take formula (109) to (113) into equation (107), we get the following HJB function:

$$
\begin{aligned}
0 = \max_{\{C_{(t)}, w_{(t)}\}} \{ & U_1[C(t), t] + \frac{\partial J[W(t), S(t), t]}{\partial t} \\
& + \frac{\partial J[W(t), S(t), t]}{\partial W} \left\{ \sum_{i=1}^{n} \left(w_i(t) W(t) [\mu_i(S,t) + \frac{\sigma_i(S,t)}{\sqrt{dt}} E^1_{(\theta,\ t)}(\omega_i)] \right) - C(t) \right\} + \sum_{j=1}^{m} \frac{\partial J[W(t), S(t), t]}{\partial S_j(t)} [f_j(S,t) + \frac{g_j(S,t)}{\sqrt{dt}} E^1_{(\theta,\ t)}(q_i)] \\
& + \frac{1}{2} \frac{\partial^2 J[W(t), S(t), t]}{\partial W^2} \sum_{l=1}^{n} \sum_{i=1}^{n} \{ w_i(t) w_l(t) W(t)^2 \sigma_i(S,t) \sigma_l(S,t) E^2_{(\theta,\ t, \rho_{il})} [\omega_i \omega_l] \} \\
& + \frac{1}{2} \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 J[W(t), S(t), t]}{\partial S_k \partial S_j} g_k(S,t) g_j(S,t) E^2_{(\theta,\ t, \eta_{jk})} [q_k(t) q_j(t)] \\
& + \sum_{j=1}^{m} \frac{\partial J[W(t), S(t), t]}{\partial W \partial S_j} [\sum_{i=1}^{n} w_i(t) W(t) \sigma_i(S,t) g_j(S,t) E^2_{(\theta,\ t, \pi_{ij})} (\omega_i q_j)] \}
\end{aligned}
\tag{114}
$$

Suppose n^th asset is risk free asset, HJB is

$$
\begin{aligned}
0 = \max_{\{C_{(t)}, w_{(t)}\}} \{ & U_1[C(t), t] + \frac{\partial J[W(t), S(t), t]}{\partial t} \\
& + \frac{\partial J[W(t), S(t), t]}{\partial W} \left\{ \left[\left(\sum_{i=1}^{n-1} \left(w_i(t) [\mu_i(S,t) + \frac{\sigma_i(S,t)}{\sqrt{dt}} E^1_{(\theta,\ t)}(\omega_i)] - r^f \right) + r^f \right) W(t) \right] - C(t) \right\} \\
& + \sum_{j=1}^{m} \frac{\partial J[W(t), S(t), t]}{\partial S_j(t)} [f_j(S,t) + \frac{g_j(S,t)}{\sqrt{dt}} E^1_{(\theta,\ t)}(q_i)] \\
& + \frac{1}{2} \frac{\partial^2 J[W(t), S(t), t]}{\partial W^2} \sum_{l=1}^{n-1} \sum_{i=1}^{n-1} \{ w_i(t) w_l(t) W(t)^2 \sigma_i(S,t) \sigma_l(S,t) E^2_{(\theta,\ t)} [\omega_i \omega_l] \} \\
& + \frac{1}{2} \sum_{k=1}^{m} \sum_{j=1}^{m} \frac{\partial^2 J[W(t), S(t), t]}{\partial S_k \partial S_j} g_k(S,t) g_j(S,t) E^2_{(\theta,\ t)} [q_k(t) q_j(t)] \\
& + \sum_{j=1}^{m} \frac{\partial J[W(t), S(t), t]}{\partial W \partial S_j} [\sum_{i=1}^{n-1} w_i(t) W(t) \sigma_i(S,t) g_j(S,t) E^2_{(\theta,\ t)} (\omega_i q_j)] \}
\end{aligned}
\tag{115}
$$

Let the derivatives of HJB on consumption $C(t)$ and proportion invested in the risky assets, $w_1(t)$ to $w_{n-1}(t)$, equal to zero, the first order conditions are,

$$
U_{1,C}[C^*(t), t] - \frac{\partial J[W(t), S(t), t]}{\partial W} = 0 \tag{116}
$$

$$
\begin{aligned}
& \frac{\partial J[W(t),S(t),t]}{\partial W} [\mu_i(S,t) + \frac{\sigma_i(S,t)}{\sqrt{dt}} E^1_{(\theta,\ t)}(\omega_i) - r^f] + \frac{\partial^2 J[W(t),S(t),t]}{\partial W^2} \sum_{l=1}^{n-1} \{ w_l^*(t) W(t) \sigma_i(S,t) \sigma_l(S,t) E^2_{(\theta,\ t)} [\omega_i \omega_l] \} \\
& + \sum_{j=1}^{m} \frac{\partial J[W(t),S(t),t]}{\partial W \partial S_j} [\sigma_i(S,t) g_j(S,t) E^2_{(\theta,\ t)} (\omega_i q_j)] = 0, \qquad i = 1,2,\cdots,n-1
\end{aligned}
\tag{117}
$$

Define a more compact expression in the following way,

$$
\begin{aligned}
V^{\theta}{}_{(n-1)\times(n-1)} = [\sigma^{\theta}{}_{il}], \quad & \sigma^{\theta}{}_{il} = \sigma_i \sigma_l E^2_{(\theta,\ t)}(\omega_i \omega_l), \quad i, l = 1,2,\cdots,n-1 \\
\Gamma^{\theta}{}_{(n-1)\times m} = [\varepsilon^{\theta}{}_{ij}], \quad & \varepsilon^{\theta}{}_{ij} = \sigma_i g_j E^2_{(\theta,\ t)}(\omega_i q_j), \quad i = 1,2,\cdots,n-1; \quad j = 1,2,\cdots,m
\end{aligned}
\tag{118}
$$

We need to be very careful to differentiate general mean of the product of two standard normal distributed random variables and general covariance of those two, since general mean is not zero any more. Maybe the symbol σ is a little bit confusing, but they are definitely not general variance-covariance matrix here! Then, we can get the optimized portfolio process.

$$w^* = -\frac{J_W[W(t),S(t),t]}{W(t)J_{WW}[W(t),S(t),t]}(V^{\theta}{}_{(n-1)\times(n-1)})^{-1}[\mu_i(S,t)+\frac{\sigma_i(S,t)}{\sqrt{dt}}E^1_{(\theta,\ t)}(\omega_i)-r^f]$$
$$-(V^{\theta}{}_{(n-1)\times(n-1)})^{-1}\Gamma^{\theta}{}_{(n-1)\times m}\frac{J_{SW}[W(t),S(t),t]}{W(t)J_{WW}[W(t),S(t),t]} \tag{119}$$

Leave out the subscript i, write $\mu(S,t)$ and $\sigma(S,t)$ in form of vectors, and sum K homogeneity investors' portfolio proportions, we get the market portfolio proportion.

$$w_M = \frac{\sum_{k=1}^{K} w^k W^k}{\sum_{k=1}^{K} W^k} = \frac{A}{M}(V^{\theta}{}_{(n-1)\times(n-1)})^{-1}[\mu(S,t)+\frac{\sigma(S,t)}{\sqrt{dt}}E^1_{(\theta,\ t)}(\omega)-\gamma^f]+(V^{\theta}{}_{(n-1)\times(n-1)})^{-1}\Gamma^{\theta}{}_{(n-1)\times m}\frac{B}{M} \tag{120}$$

where $A=\sum_{k=1}^{K}(-\frac{J^k_W[W(t),S(t),t]}{J^k_{WW}[W(t),S(t),t]})$; $B=\sum_{k=1}^{K}(-\frac{J^k_{SW}[W(t),S(t),t]}{J^k_{WW}[W(t),S(t),t]})$; $M=\sum_{k=1}^{K}W^k$

The general mean vector of excess return satisfies the following equation,

$$[\mu(S,t)+\frac{\sigma(S,t)}{\sqrt{dt}}E^1_{(\theta,\ t)}(\omega)-\gamma^f] = w_M{}^T V^{\theta}_{(n-1)\times(n-1)}\frac{M}{A}-\Gamma^{\theta}{}_{(n-1)\times m}\frac{B}{A} \tag{121}$$

Assuming that state variables, such as interest rate, climate, etc., which influence the floating rate or volatility of prices' diffusion processes are constants.

Denote $w_M{}^T V^{\theta}_{(n-1)\times(n-1)}$ as $\left(\sigma^{\theta}_{iM}\right)^T$ $\quad i=1,2,\cdots\ n-1$

Write in scalar,

$$[\mu_i(S,t)+\frac{\sigma_i(S,t)}{\sqrt{dt}}E^1_{(\theta,\ t)}(\omega_i)-r^f]=\frac{M}{A}\sigma^{\theta}_{iM}, \qquad i=1,2,\cdots\ n-1 \tag{122}$$

We specify the term of σ^{θ}_{iM} , and get

$$\begin{aligned}\sigma^{\theta}_{iM} &= \sum_{j=1}^{n} w_j\sigma^{\theta}_{ij} = \sum_{j=1}^{n} w_j\sigma_i\sigma_j E^2_{(\theta,\ t,\rho_{ij}>0)}(\omega_i\omega_j)\\ &= \sum_{j=1}^{n} w_j\sigma_i\sigma_j\{[E^1_{(\theta,\ t)}(\omega)]^2\sqrt{1-\rho^2_{ij}}\times sign(\rho_{ij})+[E^1_{(\theta,\ t)}(\omega^2)]\rho_{ij}\}\\ &= [E^1_{(\theta,\ t)}(\omega)]^2\sum_{j=1}^{n} w_j\sigma_i\sigma_j\sqrt{1-\rho^2_{ij}}\times sign(\rho_{ij})+[E^1_{(\theta,\ t)}(\omega^2)]\sum_{j=1}^{n} w_j\sigma_i\sigma_j\rho_{ij}\end{aligned} \tag{123}$$

Define, $\delta_{ij} \equiv sign(\rho_{ij})\times\sqrt{1-\rho_{ij}{}^2}$, and take into (123), we get

$$\begin{aligned}\sigma^{\theta}_{iM} &= [E^1_{(\theta,\ t)}(\omega)]^2\sum_{j=1}^{n} w_j\sigma_i\sigma_j\delta_{ij} + [E^1_{(\theta,\ t)}(\omega^2)]\sum_{j=1}^{n} w_j\sigma_i\sigma_j\rho_{ij}\\ &= [E^1_{(\theta,\ t)}(\omega)]^2\sigma_i\sigma_M\delta_{iM} + [E^1_{(\theta,\ t)}(\omega^2)]\sigma_i\sigma_M\rho_{iM}\end{aligned} \tag{124}$$

If you are interested in details during the derivation of equation (123), please see appendix III. We consider a market portfolio as a whole, and then it is a one-dimension random variable. We use symbol $\overline{M}$ to distinguish it from n dimensional market portfolio,

$$[\mu_{\overline{M}}(S,t)+\frac{\sigma_{\overline{M}}(S,t)}{\sqrt{dt}}E^{1}_{(\theta,\ t)}(\omega_{\overline{M}})-r^{f}]=\frac{M}{A}\sigma^{\theta}_{\overline{M}M}\,,\qquad i=1,2,\quad n-1 \tag{125}$$

Where,

$$\begin{aligned}\sigma^{\theta}_{\overline{M}M}&=\sum_{j=1}^{n}w_j\sigma^{\theta}_{j\overline{M}}=\sum_{j=1}^{n}w_j\sigma_{\overline{M}}\sigma_j E^{2}_{(\theta,\ t,\rho)}(\omega_{\overline{M}}\omega_j)\\&=\sum_{j=1}^{n}w_j\sigma_{\overline{M}}\sigma_j\{[E^{1}_{(\theta,\ t)}(\omega)]^2\sqrt{1-\rho^2_{j\overline{M}}}\times sign(\rho_{j\overline{M}})+[E^{1}_{(\theta,\ t)}(\omega^2)]\rho_{j\overline{M}}\}\\&=[E^{1}_{(\theta,\ t)}(\omega)]^2\sum_{j=1}^{n}w_j\sigma_{\overline{M}}\sigma_j\sqrt{1-\rho^2_{j\overline{M}}}\times sign(\rho_{j\overline{M}})+[E^{1}_{(\theta,\ t)}(\omega^2)]\sum_{j=1}^{n}w_j\sigma_{\overline{M}}\sigma_j\rho_{j\overline{M}}\end{aligned} \tag{126}$$

Define $\delta_{j\overline{M}}\equiv sign(\rho_{j\overline{M}})\times\sqrt{1-\rho_{j\overline{M}}{}^{2}}$, it can be rewritten as follow,

$$\begin{aligned}\sigma^{\theta}_{\overline{M}M}&=[E^{1}_{(\theta,\ t)}(\omega)]^2\sum_{j=1}^{n}w_j\sigma_{\overline{M}}\sigma_j\delta_{j\overline{M}}+[E^{1}_{(\theta,\ t)}(\omega^2)]\sum_{j=1}^{n}w_j\sigma_{\overline{M}}\sigma_j\rho_{j\overline{M}}\\&=[E^{1}_{(\theta,\ t)}(\omega)]^2\sigma_M\sigma_{\overline{M}}\delta_{\overline{M}M}+[E^{1}_{(\theta,\ t)}(\omega^2)]\sigma_{\overline{M}}\sigma_M\rho_{\overline{M}M}\end{aligned} \tag{127}$$

The reason why we use symbol $\overline{M}$ to distinguish one dimensional from n dimensional market portfolio is that the θ-adjusted mean does not satisfy the additivity when there is risk resources dimensional receding. After coefficients come out of θ-adjusted mean, there is no difference between $\overline{M}$ and M , so we can simplify the expression,

$$\begin{aligned}\sigma^{\theta}_{\overline{M}M}&=[E^{1}_{(\theta,\ t)}(\omega)]^2\sigma_M\sigma_M\delta_{MM}+[E^{1}_{(\theta,\ t)}(\omega^2)]\sigma_M\sigma_M\rho_{MM}\\&=[E^{1}_{(\theta,\ t)}(\omega)]^2\sigma^2_M\delta_{MM}+[E^{1}_{(\theta,\ t)}(\omega^2)]\sigma^2_M\end{aligned} \tag{128}$$

Take equations (126) and (128) into (125), and change $\mu_{\overline{M}}(S,t)$, $\sigma_{\overline{M}}(S,t)$ to $\mu_M(S,t)$, $\sigma_M(S,t)$, (125) is re-written as,

$$\frac{\mu_i(S,t)+\frac{\sigma_i(S,t)}{\sqrt{dt}}E^{1}_{(\theta,\ t)}(\omega)-r^f}{\mu_{\overline{M}}(S,t)+\frac{\sigma_{\overline{M}}(S,t)}{\sqrt{dt}}E^{1}_{(\theta,\ t)}(\omega)-r^f}=\frac{\sigma^{\theta}_{iM}}{\sigma^{\theta}_{\overline{M}M}}=\frac{[E^{1}_{(\theta,\ t)}(\omega)]^2\sigma_i\sigma_M\delta_{iM}+[E^{1}_{(\theta,\ t)}(\omega^2)]\sigma_i\sigma_M\rho_{iM}}{[E^{1}_{(\theta,\ t)}(\omega)]^2\sigma^2_M\delta_{MM}+[E^{1}_{(\theta,\ t)}(\omega^2)]\sigma^2_M},\quad i=1,2,\cdots,n \tag{129}$$

where ω is any standard normal distributed random variable. If S is constant, we define,

$$\begin{aligned}&\mu_i:=\mu_i(S,t),\qquad \sigma_i:=\sigma_i(S,t),\qquad \mu_{\overline{M}}:=\mu_{\overline{M}}(S,t)\qquad \sigma_{\overline{M}}:=\sigma_{\overline{M}}(S,t)\\&\Theta:=E^{1}_{(\theta,t)}(\omega)\qquad \Phi:=E^{1}_{(\theta,t)}(\omega^2)\qquad \tilde{\sigma}_{iM}:=\sigma_i\sigma_M\delta_{iM}\qquad \tilde{\sigma}^2_M:=\sigma^2_M\delta_{MM}\end{aligned} \tag{130}$$

Formula (129) is simplified as follows,

$$\frac{\mu_i + \frac{\sigma_i}{\sqrt{dt}}\Theta - r^f}{\mu_{\overline{M}} + \frac{\sigma_{\overline{M}}}{\sqrt{dt}}\Theta - r^f} = \frac{\Theta^2 \tilde{\sigma}_{iM} + \Phi \sigma_{iM}}{\Theta^2 \tilde{\sigma}_M^2 + \Phi \sigma_M^2} =: \beta^\theta \,, \quad i = 1,2,\cdots,n \tag{131}$$

Appendix V

Now we are working on the equity premium puzzle.

Let a generic security have price P_t at any moment in time, and let it pay dividends at the rate $D_t dt$. In an interval dt, the security pays dividends $D_t dt$.The instantaneous total return is,

$$\frac{dP_t}{P_t} + \frac{D_t}{P_t} dt \tag{132}$$

We model the price of risky assets as diffusions, for example,

$$\frac{dP_t}{P_t} = \mu dt + \sigma \sqrt{dt}\omega \tag{133}$$

where μ and σ can be functions of state variables, we can think of a risk-free security as one that has a constant price equal to 1, and pays the risk-free rate as a dividend.

$$P_t = 1 \,, D_t = r^f \tag{134}$$

or as a security that pays no dividend but whose climbs deterministically at a rate

$$\frac{dP_t}{P_t} = r^f dt \tag{135}$$

The utility function is

$$U(\{C_t\}) = E\left(\int_{t=0}^{\infty} e^{-\delta \times t} U(C_t) dt\right) \tag{136}$$

where $U(\{C_t\})$ is the total utility of the whole period, $U(C_t)$ is the utility at t time. The future price and consumptions are random variables. People only choose buying the amount of a certain security at time t, to maximize his or her wealth of whole life. So we model the process this way.

$$\begin{aligned}
\max_{\{\xi_\tau\}} E^2_{(\theta,\ t)}\{\int_t^\infty e^{-\delta\tau} U(C_\tau) d\tau\} &= E^2_{(\theta,\ t)}\{\int_t^{t+h} e^{-\delta\tau} U(C_\tau) d\tau\} + E^2_{(\theta,\ t)}\{\int_{t+h}^\infty e^{-\delta\tau} U(C_\tau) d\tau\} \\
&= E^2_{(\theta,\ t)}\{\int_t^{t+h} e^{-\delta\tau} U(C_\tau) d\tau\} + E^2_{(\theta,\ t)}\{\int_{s=0}^\infty e^{-\delta\times(t+h+s)} U(C_{t+h+s}) ds\} \\
&= E^2_{(\theta,\ t)}\{e^{-\delta\times\hat{t}} U(C_{\hat{t}}) dt\} + E^2_{(\theta,\ t)}\{\int_{s=0}^\infty e^{-\delta\times(t+h+s)} U(C_{t+h+s}) ds\}
\end{aligned} \tag{137}$$

St: $C_t = e_t - \xi_t P_t \,; \quad C_{t+s+h} = e_{t+s+h} + \xi_t D_{t+s+h} h \,; \quad h \to 0$

Although this is a continuous model, and it seems there are infinite periods, from the restrictions, we know it is a two period model virtually, since people only make decision at the starting point, and all the others can be regarded as future times as a whole. Take restrictions into objective, setting derivative with respect to ξ_t equal to zero,

$$\begin{aligned} F &= E^2_{(\theta,\ t)}\{\int_t^\infty e^{-\delta\tau}U(C_\tau)d\tau\} = e^{-\delta\times t}U(C_t)dt + E^2_{(\theta,\ t)}\{\int_{s=0}^\infty e^{-\delta\times(t+s)}U(C_{t+s})ds\} \\ \frac{\partial F}{\partial \xi_t} &= e^{-\delta\times t}U'(C_t)dt\times(-P_t) + E^2_{(\theta,\ t)}\{\int_{s=0}^\infty [e^{-\delta\times(t+s)}U'(C_{t+s})\times(D_{t+s}dt)]ds\} = 0 \end{aligned} \tag{138}$$

We obtain the first order condition for optimal portfolio choice.

$$P_tU'(C_t) = E^2_{(\theta,\ t)}\{\int_{s=0}^\infty [e^{-\delta\times s}U'(C_{t+s})D_{t+s}]ds\} \tag{139}$$

Define:

$$\Lambda_t := e^{-\delta\times t}U'(C_t) \tag{140}$$

Take (140) into (139),

$$P_t\Lambda_t = E^2_{(\theta,\ t)}(\int_{s=0}^\infty \Lambda_{t+s}D_{t+s}ds) \tag{141}$$

In general, substitute $t+h$ for t, we push time forward an infinitesimal interval,

$$P_{t+h}\Lambda_{t+h} = E^2_{(\theta,\ t+h)}(\int_{s=0}^\infty \Lambda_{t+h+s}D_{t+h+s}ds) \tag{142}$$

Take the θ-adjusted expectation operator, since law of iterated expectation still available in this broader sense, we obtain

$$E^2_{(\theta,\ t)}(P_{t+h}\Lambda_{t+h}) = E^2_{(\theta,\ t)}(\int_{s=h}^\infty \Lambda_{t+s}D_{t+s}ds) \tag{143}$$

Integral as an additive function of the interval of integration, we have

$$P_t\Lambda_t = E^2_{(\theta,\ t)}(\int_{s=0}^h \Lambda_{t+s}D_{t+s}ds + \int_{s=h}^\infty \Lambda_{t+s}D_{t+s}ds) = E^2_{(\theta,\ t)}(\int_{s=0}^h \Lambda_{t+s}D_{t+s}ds) + E^2_{(\theta,\ t)}(P_{t+h}\Lambda_{t+h}) \tag{144}$$

Applying the mean value theorem for integrals,

$$P_t\Lambda_t = \Lambda_{t+\bar{h}}D_{t+\bar{h}}h + E^2_{(\theta,\ t)}(P_{t+h}\Lambda_{t+h}) \tag{145}$$

Introduce differences,

$$P_t\Lambda_t = \Lambda_{t+\bar{h}}D_{t+\bar{h}}h + E^2_{(\theta,\ t)}[P_t\Lambda_t + (P_{t+h}\Lambda_{t+h} - P_t\Lambda_t)] \tag{146}$$

And canceling $P_t\Lambda_t$ on both sides,

$$0 = \Lambda_{t+\bar{h}}D_{t+\bar{h}}h + E^2_{(\theta,\ t)}(P_{t+h}\Lambda_{t+h} - P_t\Lambda_t) \tag{147}$$

Taking the limit as $h \to 0$

$$0 = \Lambda_tD_tdt + E^2_{(\theta,\ t)}[d(P_t\Lambda_t)] \tag{148}$$

Break up the $d(P_t\Lambda_t)$ using Ito's lemma,

$$0 = \Lambda_tD_tdt + E^2_{(\theta,\ t)}(P_td\Lambda_t + \Lambda_tdP_t + dP_td\Lambda_t) \tag{149}$$

Dividing by $P_t\Lambda_t$

$$0 = \frac{D_t}{P_t}dt + E^2_{(\theta,\ t)}[\frac{d\Lambda_t}{\Lambda_t} + \frac{dP_t}{P_t} + \frac{d\Lambda_t}{\Lambda_t}\frac{dP_t}{P_t}] \tag{150}$$

Taking $\frac{dP_t}{P_t} = r^f dt$ into equation (150), we obtain,

$$0 = E^1_{(\theta,\ t)}[\frac{d\Lambda_t}{\Lambda_t} + r^f dt + \frac{d\Lambda_t}{\Lambda_t} r^f dt] \tag{151}$$

$$r^f dt = -E^1_{(\theta,\ t)}[\frac{d\Lambda_t}{\Lambda_t}] \tag{152}$$

Taking equation (152) into (150), (150) can be rearrange as

$$E^1_{(\theta,\ t)}[\frac{dP_t}{P_t}] + \frac{D_t}{P_t}dt = r^f dt - E^2_{(\theta,\ t)}[\frac{d\Lambda_t}{\Lambda_t}\frac{dP_t}{P_t}] \tag{153}$$

With definition $\Lambda_t \equiv e^{-\delta \times t}U'(C_t)$, we take Taylor expansion,

$$d\Lambda_t = -\delta e^{-\delta \times t}U'(C_t)dt + \delta e^{-\delta \times t}U''(C_t)dC_t + \frac{1}{2}e^{-\delta \times t}U'''(C_t)dC_t^{\ 2} \tag{154}$$

$$\frac{d\Lambda_t}{\Lambda_t} = -\delta dt + \frac{C_t U''(C_t)}{U'(C_t)}\frac{dC_t}{C_t} + \frac{1}{2}\frac{C_t^{\ 2}U'''(C_t)}{U'(C_t)}\frac{dC_t^{\ 2}}{C_t^{\ 2}} \tag{155}$$

Define：

$$\alpha_t \equiv -\frac{C_t U''(C_t)}{U'(C_t)} \text{ is constant } \alpha \text{ ,} \qquad \upsilon_t \equiv \frac{C_t^{\ 2}U'''(C_t)}{U'(C_t)} \tag{156}$$

Taking (156) into (155), then taking the result into (153). When α_t is constant, we drop the subscript, and obtain

$$E^1_{(\theta,\ t)}[\frac{dP_t}{P_t}] + \frac{D_t}{P_t}dt = r^f dt - E^2_{(\theta,\ t)}[(-\delta dt - \alpha\frac{dC_t}{C_t} + \frac{1}{2}\upsilon_t\frac{dC_t^{\ 2}}{C_t^{\ 2}}) \times \frac{dP_t}{P_t}]$$

$$E^1_{(\theta,\ t)}[\frac{dP_t}{P_t}] + \frac{D_t}{P_t}dt - r^f dt = \alpha E^2_{(\theta,\ t)}[\frac{dC_t}{C_t}\frac{dP_t}{P_t}] \tag{157}$$

Taking the diffusions processes into right side of above equation,

$$E^2_{(\theta,\ t)}[\frac{dC_t}{C_t}\frac{dP_t}{P_t}] = E^2_{(\theta,\ t,\rho_{CP})}[(\mu_C dt + \sigma_C\sqrt{dt}\varpi_C) \times (\mu_P dt + \sigma_P\sqrt{dt}\varpi_P)] \tag{158}$$

where ϖ_C, ϖ_P are standard normal distributed, their correlation is ρ_{CP}. Omitting the high order derivatives, we obtain

$$E^2_{(\theta,\ t)}[\frac{dC_t}{C_t}\frac{dP_t}{P_t}] = E^2_{(\theta,\ t,\rho_{CP})}[(\sigma_C\sqrt{dt}\varpi_C)\times(\sigma_P\sqrt{dt}\varpi_P)] = \sigma_C\sigma_P dt E^2_{(\theta,\ t,\rho_{CP})}[\varpi_C\varpi_P]$$
$$= \sigma_C\sigma_P dt\left([E^1_{(\theta,\ t)}(\varpi)]^2\sqrt{1-\rho^2_{CP}}\times sign(\rho_{CP}) + E^1_{(\theta,\ t)}(\varpi^2)\rho_{CP})\right) \quad (159)$$

Taking (159) into (157),

$$E^1_{(\theta,\ t)}[\frac{dP_t}{P_t}] + \frac{D_t}{P_t}dt - r^f dt = \alpha\sigma_C\sigma_P dt\left(\Theta^2\sqrt{1-\rho^2_{CP}}\times sign(\rho_{CP}) + \Phi\rho_{CP})\right) \quad (160)$$

Canceling dt on both sides,

$$\mu_P + \frac{D_t}{P_t} + \frac{\sigma_P}{\sqrt{dt}}\Theta - r^f = \alpha\sigma_C\sigma_P\left(\Theta^2\sqrt{1-\rho^2_{CP}}\times sign(\rho_{CP}) + \Phi\rho_{CP})\right) \quad (161)$$

If the consumption and price are completely correlated, that is, $\rho_{CP} = 1$

$$\mu_P + \frac{D_t}{P_t} + \frac{\sigma_P}{\sqrt{dt}}\Theta - r^f = \alpha\sigma_C\sigma_P\Phi \quad (162)$$

Appendix VI

1. Lemma1: $\forall z_0 \in (-\infty,+\infty)$, $\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right) \xrightarrow{p} \int_{-\infty}^{+\infty}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)f(x)dx$

Proof.

$$E\left(\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)\right)$$
$$= \frac{1}{n}\sum_{t=1}^{n}E\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right) = E\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right) = \int_{-\infty}^{+\infty}\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right)f_Z(z)dz$$
$$\text{Prob}\left(\left|\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right) - E\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)\right| > \varepsilon\right) \le \frac{E\left(\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right) - E\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)\right)^2}{\varepsilon^2}$$
$$= \frac{Var\left(\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)\right)}{n^2\varepsilon^2} = \frac{Var\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)}{n\varepsilon^2} \xrightarrow{p} 0 \quad as \quad n\to\infty \quad (163)$$

Lemma1 is proved.

2. Lemma2: $\frac{1}{n}\sum_{t=1}^{n}Z_t\left(\frac{\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)}\right) \xrightarrow{p} \int_{-\infty}^{+\infty}Z\left(\frac{\theta_{Z>z_0} + (1-\theta)_{Z<z_0}}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right)f_Z(z)dz}\right)f_Z(z)dz$

Proof.

$$E\left[\frac{1}{n}\sum_{t=1}^{n} Z_t\left(\frac{\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}\right)\right]$$

$$= E\left(\frac{\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)Z}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}\right) = \frac{E\left(\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)Z\right)}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}$$

$$\xrightarrow{p} \frac{E\left(\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)Z\right)}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)f(z)dz} = E\left(\frac{\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)Z}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)f_Z(z)dz}\right) \qquad (164)$$

$$= \int_{-\infty}^{+\infty} Z\left(\frac{\theta_{Z>z_0}+(1-\theta)_{Z<z_0}}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)f_Z(z)dz}\right)f_Z(z)dz \quad as \quad n\to\infty$$

by Chebyshev's inequality,

$$\Pr ob\left(\left|\frac{1}{n}\sum_{t=1}^{n} Z_t\left(\frac{\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}\right) - E\left(\frac{\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)Z}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)f_Z(z)dz}\right)\right| > \varepsilon\right)$$

$$\le \frac{nVar\left(\frac{Z_t\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}\right)}{n^2\varepsilon^2} \qquad (165)$$

$$\xrightarrow{p} 0 \quad as \quad n\to\infty$$

$$\therefore \frac{1}{n}\sum_{t=1}^{n} Z_t\left(\frac{\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0}+(1-\theta)_{Z_t<z_0}\right)}\right) \xrightarrow{p} E\left(\frac{\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)Z}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0}+(1-\theta)_{Z<z_0}\right)f_Z(z)dz}\right)$$

Lemma2 is proved.

3. Lemma [Weak Law of Large Numbers of (WLLN) of general mean for i.i.d Samples]: Suppose $\{Z_t\}$ is i.i.d (μ,σ^2) with $E_\theta(Z_t)$, $E|Z_t|<\infty$, and define

$$\bar{z}_n = \frac{\sum_{Z_i>\bar{Z}_n}\frac{1}{n}\theta_{Z_i>\bar{Z}_n}Z_i + \sum_{Z_j<\bar{Z}_n}\frac{1}{n}(1-\theta)_{Z_j<\bar{Z}_n}Z_j}{\sum_{Z_i>\bar{Z}_n}\frac{1}{n}\theta_{Z_i>\bar{Z}_n} + \sum_{Z_j<\bar{Z}_n}\frac{1}{n}(1-\theta)_{Z_j<\bar{Z}_n}}$$

where $n=1,2,\cdots$, then

$$\bar{z}_n \xrightarrow{p} E_\theta(Z_t) \quad as \quad n\to\infty$$

Proof.

Let:

$$z_0 = E\left(\frac{\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right)Z}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right) f_Z(z)dz} \right) = \int_{-\infty}^{+\infty} Z \left(\frac{\theta_{Z>z_0} + (1-\theta)_{Z<z_0}}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right) f_Z(z)dz} \right) f_Z(z)dz \qquad (166)$$

Using Lemma2,

$$\frac{1}{n}\sum_{t=1}^{n} Z_t \left(\frac{\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}}{\frac{1}{n}\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)} \right) \xrightarrow{p} E\left(\frac{\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right)Z}{\int_{-\infty}^{+\infty}\left(\theta_{Z>z_0} + (1-\theta)_{Z<z_0}\right) f_Z(z)dz} \right) = z_0 \quad as \quad n \to \infty$$

$$\therefore \forall \varepsilon, \qquad P\left(\left| \sum_{t=1}^{n} \frac{\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)Z_t}{\sum_{t=1}^{n}\left(\theta_{Z_t>z_0} + (1-\theta)_{Z_t<z_0}\right)} - z_0 \right| \le \varepsilon \right) \to 1$$

$$let \quad g(\bar{z}_n) = \sum_{t=1}^{n} \frac{\left(\theta_{Z_t>\bar{z}_n} + (1-\theta)_{Z_t<\bar{z}_n}\right)Z_t}{\sum_{t=1}^{n}\left(\theta_{Z_t>\bar{z}_n} + (1-\theta)_{Z_t<\bar{z}_n}\right)} - \bar{z}_n = 0 \qquad (167)$$

$$P\left(\left| z_0 - \arg_{\bar{z}_n}\left(g(\bar{z}_n) = 0\right) \right| < \varepsilon \right) \to 1$$

$$\arg_{\bar{z}_n}\left(g(\bar{z}_n) = 0\right) \xrightarrow{p} z_0$$

That is $\bar{z}_n \xrightarrow{p} E_\theta(Z_t) \quad as \quad n \to \infty$.

This completes the proof.

Appendix VII

General mean is an implicit function. Assume random variable is normal distributed. We use the orthogonal polynomials to approximate the integrand, then the integral is evaluated by Gaussian quadrature. Although Equation (172) is still an implicit function with respect to $q*$, it is of a much simpler form, and easy to get the result using root finding technique.

First, we have

$$q* = \frac{\int [(1-\theta)1_{X<q*} + \theta 1_{X>q*}] f_X(x) x dx}{\int [(1-\theta)1_{X<q*} + \theta 1_{X>q*}] f_X(x) dx}, \quad f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \qquad (168)$$

which is equivalent to

$$b* = \frac{\int [(1-\theta)1_{X<b*} + \theta 1_{X>b*}] e^{-x^2} x dx}{\int [(1-\theta)1_{X<b*} + \theta 1_{X>b*}] e^{-x^2} dx}, \text{ where } b* = \frac{q*}{\sqrt{2}} \qquad (169)$$

Then,

$$(1-\theta)\int_{-\infty}^{0} e^{-(x-b)^2} xdx + \theta\int_{0}^{+\infty} e^{-(x+b)^2} xdx = (\theta-1)\int_{0}^{+\infty} e^{-(x-b)^2} xdx + \theta\int_{0}^{+\infty} e^{-(x+b)^2} xdx = 0 \quad (170)$$

Namely,

$$(\theta-1)\int_{0}^{+\infty} e^{-x} e^{-(x^2-2bx+b^2-x)} xdx + \theta\int_{0}^{+\infty} e^{-x} e^{-(x^2+2bx+b^2-x)} xdx = 0 \quad (171)$$

We name e^{-x} the weight function, then $e^{-(x^2-2bx+b^2-x)}x$ is the integrand. By setting $q^* = \sqrt{2}b = (\ln B)/\sqrt{2}$, and using Guass-Laguerre formula, we obtain the results as follows.

$$\sum_{k=1}^{n} a_k[(\theta-1)B^{x_k} + \theta B^{-x_k}] = 0, \quad (172)$$

where

$$a_k = A_k e^{-x_k{}^2 + x_k},\ A_k = \frac{x_k}{(n+1)^2[L_{n+1}(x_k)]^2},\ L_{n+1}(x_k) = \sum_{k=0}^{n}(-1)^k \binom{n+a}{n-k}\frac{x^k}{k!} = 0 \quad (173)$$

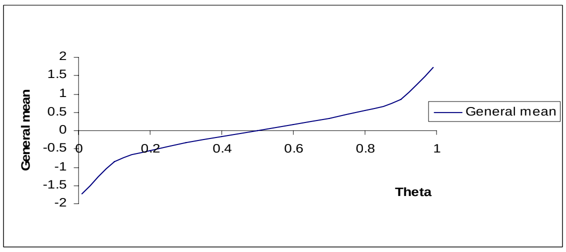
That is, solving equation $L_{n+1}(x_k) = 0$, we get the x_k , $k = 1,\cdots n$, and all the corresponding a_k , $k = 1,\cdots n$. After taking them into equation VI-1, we get B using Newton root finding method, and q^* is obtained. The effort of calculating x_k and a_k is some kind of done once and for ever work. Given different θ, calculating q^* is to solve equation with respect to different θ, the set of x_k and a_k, $k = 1,\cdots n$ are never changed.

See Figure 4, we find that the general mean is more sensitive when people's view bias is more pessimism or optimism. Figure 1 indicates that $E_\theta^1(X)$ is a monotonously increasing function of θ.

Using a similar Guassian quadrature method, Guass-Hermite, to calculate the general mean, we get the same result but with a poorer performance from both of the convergence speed and convergence tolerance perspectives. Guass-Hermite quadrature are used as checking computations.

Figures and tables

Figure 1: General mean of standard normal distribution with respect to view bias



Notes: Using the numerical method in appendix VII, we get a monotonously increasing function of general mean with respect to θ. What's more, we find that $D_\theta^1(X)$ is invariable with respect to θ at least for normal distribution.

Table 1: Perfect information based decision making

X state	s1	s2	s3	s4	sum	Y state	s1	s2	s3	s4	sum
Payoff	1	3	50	100		Payoff	-1000	3	50	100	
Probability	0.05	0.15	0.7	0.1	1	Probability	0.00001	0.09999	0.8	0.1	1
E(X)	0.05	0.45	35	10	45.5	E(Y)	-0.01	0.3	40	10	50.3
D(X)	99.01	270.94	14.18	297.03	681.15	D(Y)	11.03	223.61	0.07	247.11	481.82

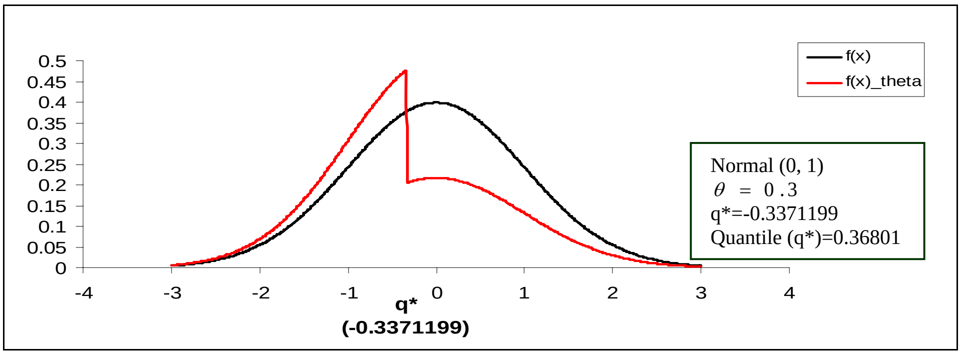
Notes: Suppose there are four states in the world. As shown in table 1, the payoff and the probabilities of two strategies X and Y under each state are known. Setting out from the traditional mean variance efficiency, the risk averse will prefer Y strategy, since the expectation of Y is greater than X, meanwhile the variance of Y is less than that of X.

Table 2: Imperfect information based decision making

X state	s1	s2	s3	s4	Sum	Y state	s1	s2	s3	s4	Sum
payoff	1	3	50	100		payoff	-1000	3	50	100	

Notes: When people are blind to the probabilities of each state, their decisions will depend on whether they are pessimistic or optimistic. When they are pessimistic, they will choose the strategy following the maxmin principle. The minimum of X is one, and Y is -1000, people select the maximum of them. That is, strategy X is preferred.

Figure 2: Probability Adjustment under Imperfect Information (pessimistic investor)



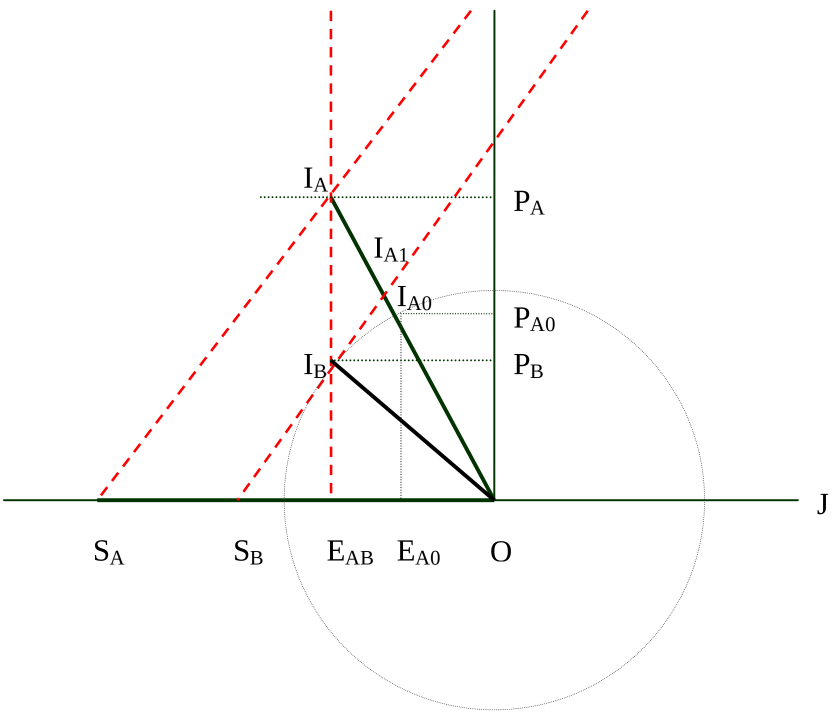
Notes: When people are pessimistic, they will relatively amplify the possibilities of left tail events, and vice versa. In this example, it is a standard normal distribution. When view bias 0.3 is given, the mean of a pessimistic investor with that view bias will hold an expectation of -0.3371199, and the quantile of that expectation is 0.36801.

Table 3: Comparison between Risk Preference and View Bias

$P(x) = E_{\theta}^{1}(mx) = \int_{\Omega} \pi(\theta) mf(x) x dx$	**Risk preference**	**View tendency**
Similarity	Change of measure(real to subjective probability)	
	Only mean is altered, variance remains the same, at least for normal distribution	
Dissimilarity	Shift	Reshape
	Perfect/Imperfect Info	Imperfect Info
	Character	Attitude
	Stable	Variable
	No instantaneous profit from repackaging	Instantaneous profit from repackaging
Relation	separable	

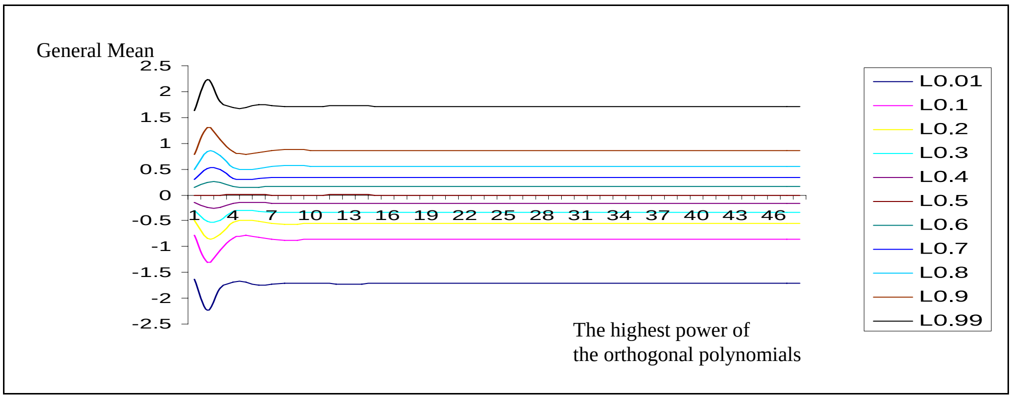
Notes: This table summarized the similarities, dissimilarities, as well as the relation between risk preference and view bias.

Figure 3: View Bias Adjusted Risk-Reward Projection



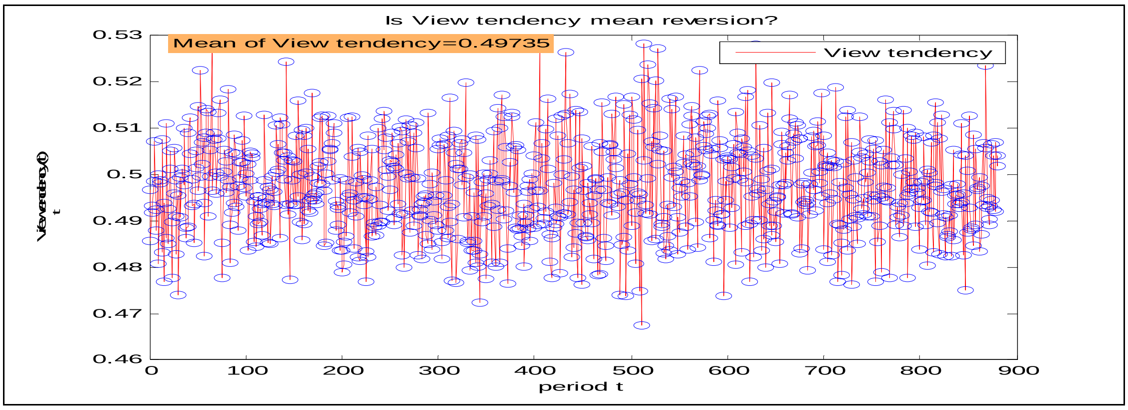
Notes: Figure 3 depicts the relationship between the i^th and any j^th security. Under perfect information, assets are priced under a vertical projection, so the price of OI_A and OI_B are exactly the same. The distinction of the idiosyncratic risks, I_AI_B, causes no pricing difference. Under imperfect information, the assets are priced under a projection from a setting sun. Now, the systematic risk is a linear combination of the exposed risk and some other risk, which will never be priced under a vertical projection. That is why we name it the potential risk. Therefore, in such case, asset I_{A0} and I_B will have the same price OS_B. The distinction of the idiosyncratic risks, $I_{A0}I_B$, causes no pricing difference.

Figure 4: Convergence Procedure of General Mean (Standard Normal Distribution)



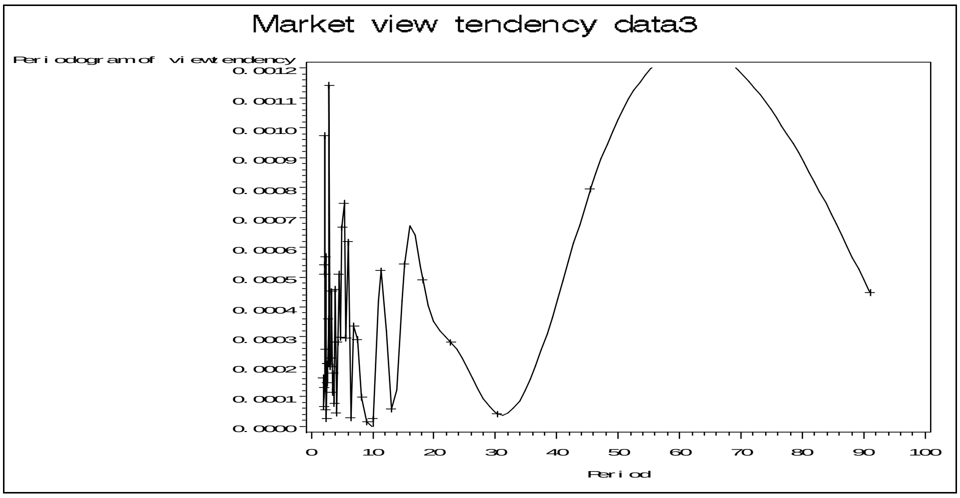
Notes: After doing θ's sensitivity analysis, we find that the greater from view neutral the deviation is, the more sensitive the general mean is, with respect to view bias.

Figure 5: View Bias Trend Chart



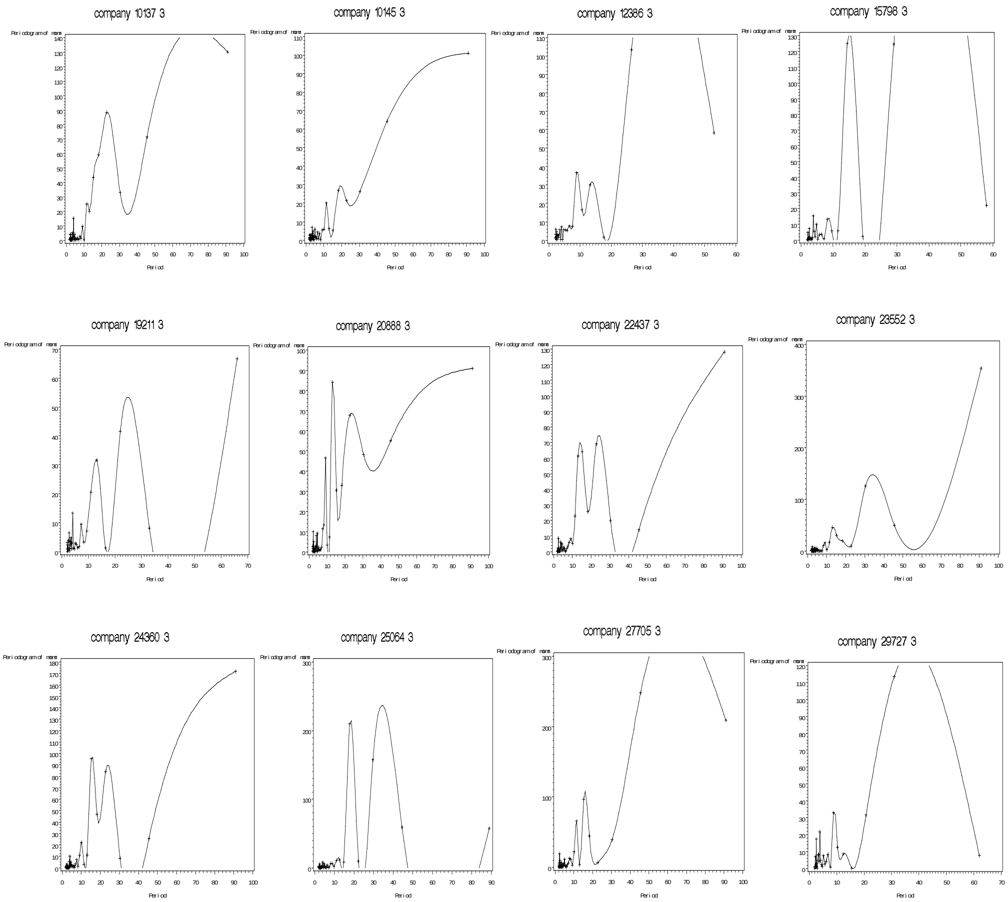
Notes: We assume that the risk aversion is constant 3, then figure out the monthly view bias time series implied by equity premium puzzle. The data range is from May.1926 to Sep.1999. The mean is 0.497, a slight deviation from view neutral, the maximum and minimum are 0.528 and 0.467 respectively.

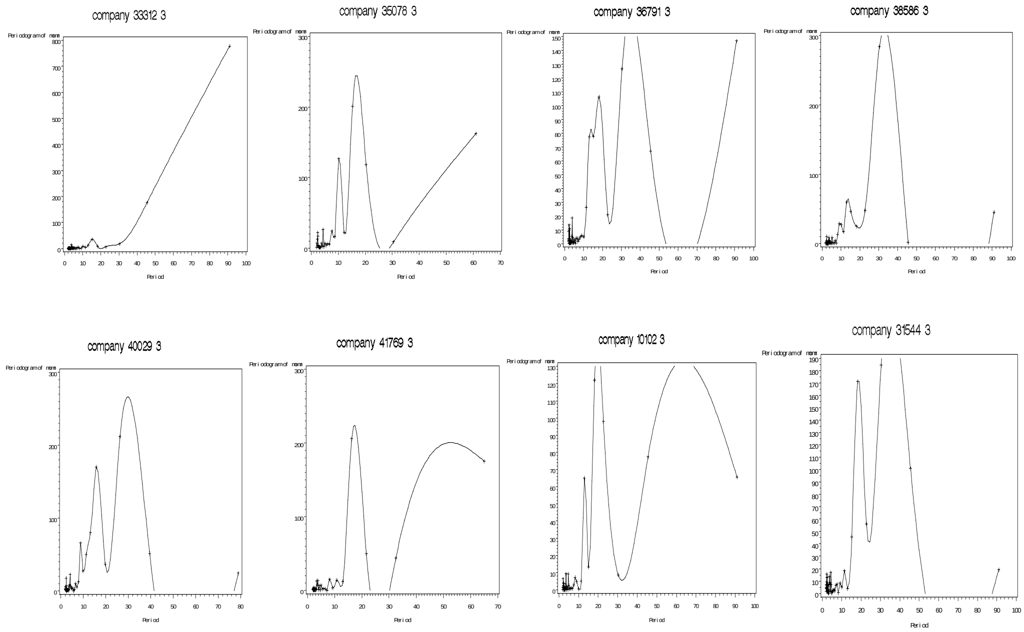
Figure 6: Periodogram of view bias



Notes: We draw the view bias periodogram within the time span from Jun. 1963 to Dec. 1970. The periodicity is around 60 months.

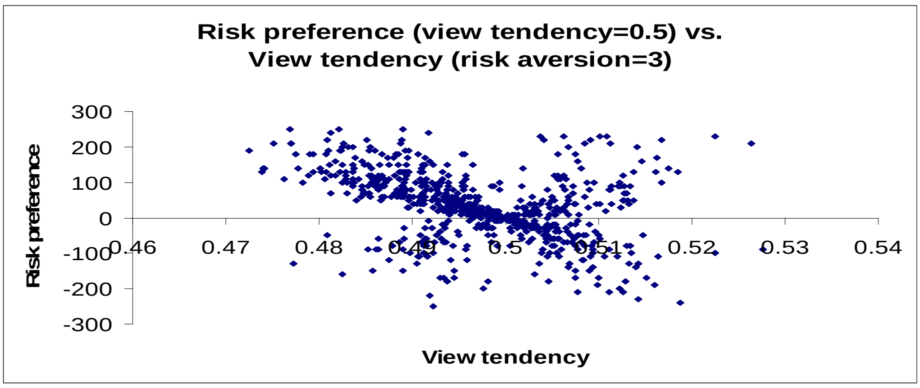
Figure 7 Periodogram of Momentum Ranking

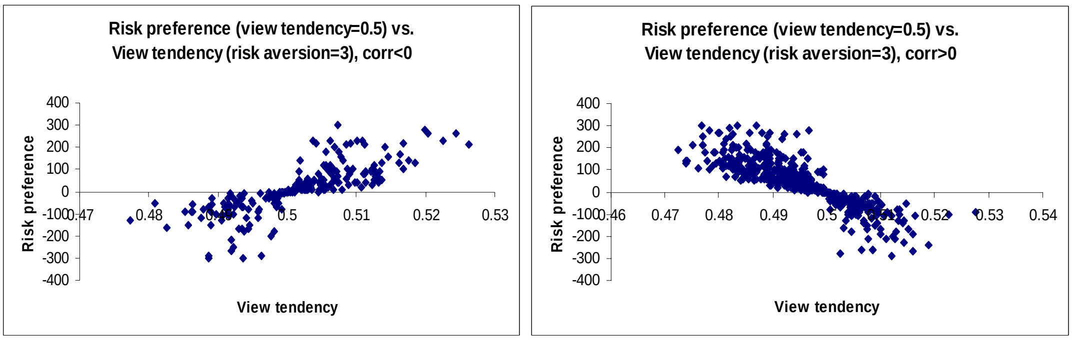




Notes: We run the monthly common stock returns with share code 10/11 from NYSE, AMEX and NASDAQ. The sample range is from Jun.1963 to Dec. 1970. Following many studies, the ranking periods have the length of six months. For any given month, the rank of a certain stock is determined based on the past 6-month returns. The above 20 representative stocks are selected randomly by the permenant company number (PERMNO). We get the periodicities of the ranking of each selected stock. When we make comparisons between the periodicities of view bias and each representative common stock, we find that quite many of them are compatible. 13 of 20 stocks have a distinct periodicity, which are either around 30 or 60 months, namely the half of or equal to the periodicity of view bias.

Figure 8: Scatter Diagrams of Risk preference and View Bias





Notes: We get the time series of risk preference implied by equity premium puzzle assuming a constant neutral view bias. In like wise, we get the time series of view bias also implied by equity premium puzzle but with a constant risk preference 3. Figure 8 depicts the relationship between those two. After grouping the data set by the sign of correlations ρ_{CP} , we get two strongly correlated linear relationship.

Table 4: Linear relationship between Risk Preference and View bias

	Regression Statistics			*Coefficients*	*Standard Error*	*t Stat*	*P-value*	*Significance F*
Corr<0	R Square	0.631877	Intercept	-5013.28810	259.77570	-19.29853	4.627E-49	3.341E-49
	Observations	220	Slope	10016.78149	517.82104	19.34410	3.341E-49	
Corr>0	R Square	0.720758	Intercept	4761.18374	129.88509	36.65689	2.82E-145	6.15E-144
	Observations	514	Slope	-9530.97021	262.17908	-36.35290	6.15E-144	

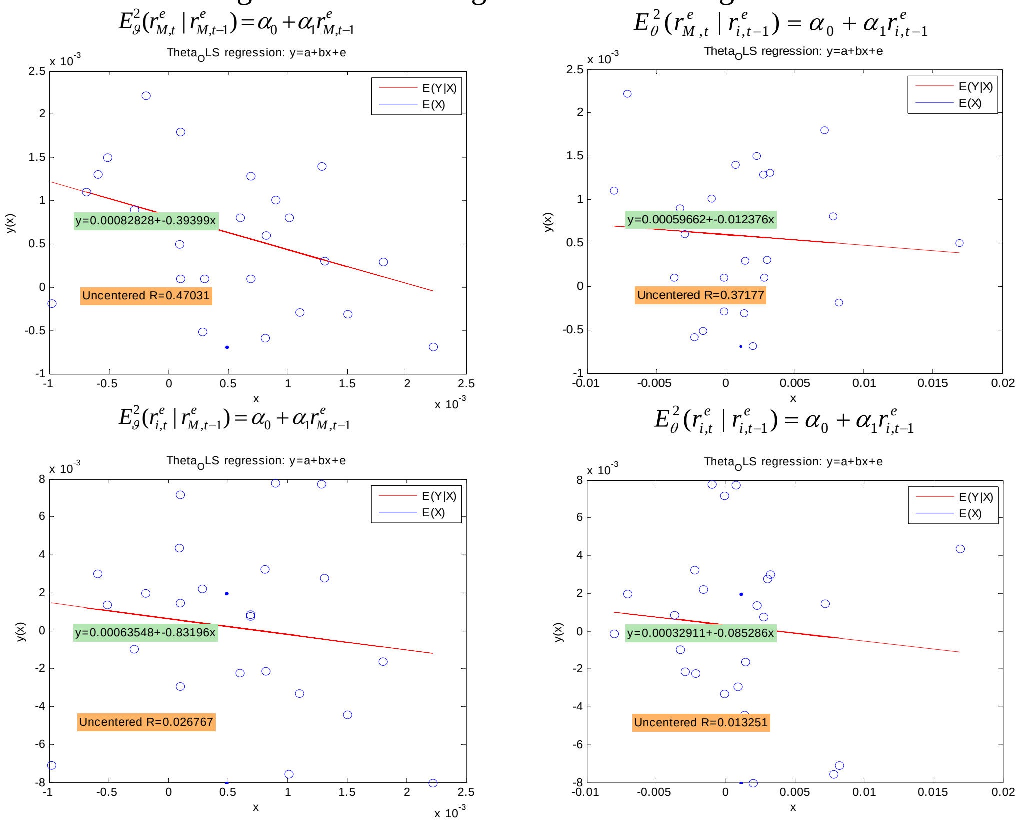
Notes: We run regression analysis by group only keeping the records whose risk aversion is within the range of (-300, 300). The result indicates that they are strongly correlated for both two groups. The risk preference and view bias are negative correlated when ρ_{CP} is positive, and are positive correlated when ρ_{CP} is negative. In other words, if the market return and consumption growth move in the same direction, a risk averse investor is more like a pessimistic investor. If the market return and consumption growth move in a counter-direction, a risk averse investor is more like an optimistic investor.

Table 5: Model A (CAPM) Summary and Parameter Estimates

Model A Summary		OLS Parameter Estimates				
Parameter	2	Parameter	Estimate	Std.Err	t-Value	Pr>| t |
Equations	1	Intercept	-0.00089	0.00105	-0.84000	0.40740
R square	0.110464	Slope	1.85393	1.09700	1.69000	0.10450

Notes: We select the stock (PERMNO 27705) as sample individual security. The sample range is from July 1963 to Nov.1965, total 30 months. We eliminate five abnormal records. The classical OLS regression indicates that there is no linear relationship between the excess return of individual security and market index. R square is only 0.11, and t-test of both the intercept and slope could not be passed. That is how the traditional CAPM model being challenged.

Figure 9: Scatter Diagram and VOLS Regression



Notes: The average of implied view bias during the sample period is 0.494. We run four VOLS regressions. Only those two whose regressand is the market excess return show us some linear relationship, the Uncentered R square is 0.47 and 0.37, respectively.

Table 6: Model B (VCAPM) Summary and Parameter Estimates

Model B Summary		Nonlinear GMM Parameter Estimates				
Parameter	2	Parameter	Estimate	Std.Err	t-Value	Pr>| t |
Equations	3	Intercept	-0.00106	0.00041	-2.62000	0.01500
Number of statements	3	Slope	1.63536	0.35320	4.63000	0.00010

Table 7: Model C (VCAPM) Summary and Parameter Estimates

Model C Summary		Nonlinear GMM Parameter Estimates				
Parameter	2	Parameter	Estimate	Std.Err	t-Value	Pr>| t |
Equations	3	Intercept	-0.00091	0.00026	-3.46000	0.00200
Number of statements	3	Slope	1.69443	0.24460	6.93000	<0.001

Notes: Table 6 and table 7 indicate that both of model B and C show us a statistically significant view bias based beta, p-values are no more than 0.0001. More important, there is no evidence indicates that we should reject the zero intercept null hypothesis. The p-value of intercept B and C is 0.015 and 0.002 respectively.

Reference

Abel, A.B. (2002)., "An Exploration of the Effects of Pessimism and Doubt on Asset Returns", *Journal of Economic Dynamics and Control* 26, 1075-1092.

Acerbi, C. and P. Simonetti, 2002, "Portfolio Optimization with Spectral Measure of Risk", Working paper.

Acerbi, C., 2002, "Spectral Measures of Risk: A Coherent Representation of Subjective Risk Aversion", *Journal of Banking and Finance*, 26: 1505-1518.

Bassett, G. W., R. Koenker, and Kordas, 2004, "Pessimistic Portfolio Allocation and Choquet Expected Utility", *Journal of Financial Econometrics*, 2(4): 477-492.

Berk, Jonathan B., Richard C. Green, and Vasant Naik, 1999, "Optimal investment, growth options, and security returns", *Journal of Finance* 54, 1153-1607.

Bodie, Z., Kane, A. and Marcus A., 2005, *Investments*, 6th Edition, McGraw-Hill, New York.

Brav, Alon, and John B. Heaton, 2002, "Competing theories of financial anomalies", *Review of Financial Studies* 15, 575-606.

Chochrane, John H. 2001, *Asset Pricing*, Princeton University Press, New Jersey 08540

Coles, Jeffrey, Uri Loewenstein, and Jose Suay, 1995, "On equilibrium pricing under parameter uncertainty", *Journal of Financial and Quantitative Analysis* 30, 347-364.

Conrad, Jennifer, and Gautam Kaul, 1998, "An anatomy of trading strategies", *Review of Financial Studies* 11, 489-519.

Cooper, Michael J., Roberto C. Gutierrez, and Allaudeen Hameed, 2004, "Market states and momentum", *Journal of Finance* 59, 1345-1365.

Diacogiannis. George and Feldman, David 2006, "The CAPM Relation for Inefficient Portfolios", Working paper.

Fama, Eugene F., and Kenneth R. French, 1996, "Multifactor explanations of asset pricing anomalies", *Journal of Finance* 51, 55-84.

Feldman, David and Reisman H., 2003, "Simple Construction of the Efficient Frontier," *European Financial Management*, 9, 251-259.

Ghysels, Eric and Jacquier, Eric 2006, "Market Beta Dynamics and Portfolio Efficiency", working paper.

Gourieroux, and Liu 2006 "Sensitivity Analysis of Distortion Risk Measures", working paper.

Gourieroux, C. and J. Jasiak, 2005. "Value-at-Risk", *Handbook of Financial Econometrics*, Forthcoming.

Gourieroux, C., J. P. Laurent, and O. Scaillet, 2000, "Sensitivity Analysis of Values at Risk", *Journal of Empirical Finance*, 7: 225-245.

Griffin, John, Susan Ji, and Spencer Martin, 2003, "Momentum investing and business cycle risk: Evidence from pole to pole", *Journal of Finance* 58, 2515-2547.

Guidolin, Massimo, 2006. "Pessimistic beliefs under rational learning: Quantitative implications for the equity premium puzzle", *Journal of Economics and Business*, Elsevier, vol. 58(2), pages 85-118.

Hey, John D. 1984, "The Economics of Optimism and Pessimism: A Definition and Some Applications", Kyklos, 37(2), 181-205.

Hong, Harrison, and Jeremy Stein, 1999, "A unified theory of underreaction, momentum trading and overreaction in asset markets", *Journal of Finance* 54, 2143-2184.

Huang, C-F. and Litzenberger, R., 1988, *Foundations for Financial Economics*, North-Holland, New York.

Jegadeesh, Narasimhan, and Titman, Sheridan 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency", *Journal of Finance* 48, 65-91.

Jegadeesh, Narasimhan, and Sheridan Titman, 2002, "Cross-sectional and time-series determinants of momentum returns", *Review of Financial Studies* 15, 143-157.

Johnson, Timothy C., 2002, "Rational momentum effects", *Journal of Finance* 57, 585-608.

Knight, Frank H., 1921 *Risk, Uncertainty, and Profit, Boston*, MA: Hart, Schaffner & Marx; Houghton Mifflin Company

Lustig, Hanno and Verdelhan, Adrien 2006 "The Cross-Section of Foreign Currency Risk Premia and Consumption Growth Risk", working paper.

Mehra, R., and E.C. Prescott. 1985. "The Equity Premium: A Puzzle." *Journal of*

Monetary Economics, vol. 15, no. 2 (March):145–161.

Merton, R. C., 1973, “An Intertemporal Capital Asset Pricing Model”, *Econometrica*, 41, 867-887.

Muren, Astri 2006 “Unrealistic Optimism about Exogenous Events: An experimental test”, working paper.

Mehra, Rajnish 2003. “The Equity Premium: Why is it a Puzzle?” *Financial Analysts Journal*, January/February :54-69.

Schmeidler, D., 1989. “Subjective Probability and Expected Utility without Additivity”, *Econometrica*, 57(3): 571-587.

Wang, Keven Q. 2006, “Mean-Reversion and Momentum”, working paper.

Weinstein, Neil D, (1980), “Unrealistic optimism about future life events”, *Journal of Personality & Social Psychology*, 39(5), 806-820.

Zhang, X. Frank, 2005, “Information uncertainty and stock returns”, *Journal of Finance*, forthcoming.