# Option Implied Risk Aversion

**Zhou, Hong Cheng**

**Thesis supervisor: Jørgen Haug**

Master thesis within the major profile of Energy, Natural Resources and the Environment

NORWEGIAN SCHOOL OF ECONOMICS

# Abstract

Risk aversion has been empirically estimated using different equilibrium models [Hansen and Singleton (1982, 1983), Mehra and Prescott (1985)]. However, the results are controversial. Jackwerth (2000) and Ait-Sahalia and Lo (2000) study the risk aversion from a different perspective. They derive the risk aversion function across wealth by using the subjective density and risk-neutral density. Under this method, they avoid using the low-frequency consumption data. In this paper, we use British capital market data to calculate the risk aversion, which can be expressed as a function of risk-neutral density and subjective density. The risk-neutral densities are estimated by two different methods, the double lognormal method of Bahra (1997) as well as the fast and stable method of Jackwerth (2000). The subjective density is generated by the GARCH Monte Carlo method.

Compared to the subjective densities, the estimated risk-neutral densities are leptokurtic with fatter left tails. Using the Kolmogorov-Smirnov test, we find that the risk-neutral densities estimated by different method are statistically different at conventional significance level. The calculated implied risk aversion functions are U-shaped across wealth. This is inconsistent with the theory of finance. Further, the U-shaped risk aversion function is stable under different settings.

# Acknowledgements

I would like to express my deep appreciation to my parents and my friend Jing. They supported me through my whole studying period, from Glasgow to Bergen, from Bergen to Nashville.

I would also like to thank Professor Jørgen Haug, who gave me many helpful advices for selecting the right thesis topic and for organizing the structure of thesis. He supported me throughout my thesis with his patience and knowledge whilst giving me the room to work in my own way.

Last but not least, I would like to thank myself, for working so hard in the past two years. Now I understand the beauty of finance theory.

# Table of Contents

# 1. Introduction

Classical finance theory tries to explain how people make decisions under uncertainty by assuming that people's preference can be represented by a well-behaved utility function. Then one can use asset pricing models to compute the price of financial assets. An important input of the utility function is the so-called risk aversion coefficient, which describes an agent's degree of risk aversion. Given the consumption and capital market data, the risk aversion coefficient can be estimated by equilibrium asset pricing models. However, results of these studies are controversial. Hansen and Singleton (1982, 1983) find that the coefficient of risk aversion is around one, while Mehra and Prescott (1985) arrive at the conclusion that the equity premium can only be explained by an extremely high level of risk aversion.

In the late 1990s, Jackwerth (2000) and Ait-Sahalia and Lo (2000) attack the problem from another perspective. They combine the equilibrium pricing framework with the no arbitrage pricing framework, and derive investors' risk aversion from risk-neutral density (RND) and subjective density. Various RND estimation techniques developed in the 1990s are the foundation for estimating risk aversion function.

Under no arbitrage condition, assets prices equal to the product of their payoffs and the risk-neutral probability discounted by the risk-free rate. Inversely, risk-neutral probability can be derived from asset prices. Option market contains abundant information about investors' future beliefs. Breeden and Litzenberger (1978) and Cox and Ross (1976) develop the theoretical foundation for deriving risk-neutral density from option market. On the other side, Bahra (1997), Jackwerth (2000), and Ait-Sahalia and Lo (1998), among others, propose various empirical RND estimation methods.

In this paper, asset pricing theories related to density estimation and implied risk aversion are briefly reviewed. Subsequently, the RND function and empirical risk aversion will be empirically derived from option prices. Using FTSE 100 index options (ESX), I find that the RNDs estimated by the two methods are statistically different. In addition, unlike subjective

densities, RNDs are leptokurtic with fatter left tails. Next, we compute the implied risk aversion which is a U-shaped function across wealth. After implementing several robustness tests, I conclude that the risk aversion is U-shaped under different empirical settings. Further, I find that the risk aversion function is time-varying.

The contribution of this paper is twofold: firstly, we derived the implied risk aversion function from British dataset; secondly, the RNDs are estimated by two different methods, the double lognormal method [Braha (1997)] as well as the fast and stable method [Jackwerth (2000)].

The structure of the thesis is as follow: section 2 discusses the asset pricing theory and the derivation of the risk aversion function; the density estimation techniques are illustrated in section 3; section 4 provides an empirical example; the last section concludes the thesis.

# 2. Review of Asset Pricing Theory

Asset pricing can be categorized into equilibrium pricing and arbitrage pricing. Equilibrium models price financial assets by demand and supply analysis. No arbitrage pricing models, on the other side, assume that assets with the same future payoffs should have identical price. In this section, both frameworks will be discussed. A more detailed review of asset pricing theory is given in the appendix A.

## 2.1 Equilibrium Pricing

Microeconomists assume that on the demand side consumers' preference can be represented by utility functions, and on the supply side the productions of firms are measured by production functions. The price and quantity of the goods can be found by solving a mathematical optimization problem, given that consumers and firms maximize their utilities and profits respectively. Alternatively, in an exchange economy, the supply of goods is treated as endowment. Exchange economy is frequently used in financial economics. Nevertheless, the static framework is not suitable for asset pricing because financial assets normally live more than one period and have stochastic future payoffs.

Arrow (1964) and Gerard (1959) extend the classical static exchange economy to a two-period stochastic economy, where the consumption and future endowment are stochastic. In order to maximize their lifetime utility, agents allocate their assets intertemporally and hedge against future consumption risks. Their model can be further extended to a multi-period framework [Lucas (1978)].

In equilibrium, asset prices determined by supply and demand at each moment can be used to derive investors' belief [Wang (1993)]. In a standard dynamic exchange economy, e.g. the continuous-time setting of He and Leland (1993), the security market is dynamically complete so that there exists a single representative investor[1] who lives in a finite time interval $[0, T]$. There

---

[1] See Constantinides (1982).

is one riskless bond and one risky stock that represent the market portfolio. The investor is a self-interested, risk averse utility maximizer endowed with one unit of stock at period 0 but no exogenous income. He maximizes his expected utility of consuming a single good on the final date $T$[2]. Suppose the riskless rate is $r$, and the risky asset price follows a diffusion process

$$dS_t = \mu(S_t, t)S_t dt + \sigma(S_t, t)S_t dB_t \quad (2.1)$$

The investor's consumption and investment problem is

$$Max \, E[U(W_T)]$$

$$s.t. \, dW_s = [rW_s + A_s(\mu(S_s, s) - r)]ds + A_s\sigma(S_s, t)dB_s$$

$$W_s \geq 0, 0 \leq s \leq T$$

where $W_s$ is the investor's wealth at period $s$ and $A_s$ is the amount invested in risky stock.

In equilibrium, the investor follows the so-called path-independent strategy and optimally invests all the wealth in the risky stock. This is equivalent to

$$A_s = S_s, \forall s \in [0, T]$$

The concept of equilibrium used here is stronger than what is commonly defined in competitive equilibrium [Merton (1973) and Breeden (1979)]. He and Leland (1993) derive a partial differential equation that must be satisfied by this equilibrium asset price process. Clearly, the set of this equilibrium price process is a subset of the competitive equilibrium price process. However, it still covers an important group of asset price dynamics. For example, Cox and Leland (2000) demonstrate that if the equilibrium asset prices are lognormally distributed then the investor must be following a path-independent strategy. The advantage of this equilibrium is that one can conveniently infer investors' preference from stock prices. He and Leland (1993) show that if the volatility is constant, a decreasing relative risk aversion in terminal wealth of the representative investor is the necessary and sufficient condition for the expected instantaneous return of the market portfolio to be mean reverting. As a matter of fact, under certain assumptions the entire risk aversion function across investor's final wealth can be extracted from asset prices [Ait-Sahalia and Lo (2000) and Jackwerth (2000)].

---

[2] Changing this will not change the results.

## 2.2 No Arbitrage Pricing

Under no arbitrage condition, portfolios with identical future payoffs should have the same price. Therefore, if we know the distribution of future asset payoffs and the value of state prices, asset price can be calculated by

$$P_t = \int_0^\infty f_S(S_T) X(S_T) dS_T \quad (2.2)$$

where $f_S(S_T)$ is the state price density, $S_T$ is the asset price at time $T$, $X(S_T)$ is the future payoff of $S_T$.

Alternatively, the pricing formula can be expressed in terms of risk-neutral density. As we know, state price density equals risk-neutral density discounted by riskless rate. Hence asset price can be expressed as

$$P_t = e^{-rT} \int_0^\infty f_Q(S_T) X(S_T) dS_T \quad (2.3)$$

where $f_Q(S_T)$ is the risk-neutral density for asset $S_T$.

It is possible to transform the diffusion process (2.1) in the probability space $(\Omega, P, \mathcal{F})$ to a stochastic process with a riskless drift term in the risk-neutral probability space $(\Omega, Q, \mathcal{F})$ by Radon-Nykodim derivative

$$\frac{df_Q(S_T)}{df_P(S_T)} = exp\left[-\frac{1}{2}\int_0^t \left(\frac{\mu(S_s, s) - r}{\sigma(S_s, s)}\right)^2 ds - \int_0^t \frac{\mu(S_s, s) - r}{\sigma(S_s, s)} dB_s\right]$$

## 2.3 Implied Risk Aversion

With the preparations showed above, we can start deriving the risk aversion function. In equilibrium, the representative investor optimally invests in the market portfolio and spends all the wealth on the final date $T$, i.e. $S_T = W_T = C_T$. Therefore, the investor's maximization problem can be rewritten as

$$Max \int_0^\infty f_P(S_T) U(S_T) dS_T$$

where $f_P(S_T)$ is the subjective probability density of the stock price on date $T$ across states. Since the investors are defined as rational, their subjective density should be an unbiased forecast of the realized density, and thus consistent with the objective or physical density [Bliss and Panigirtzoglou (2004)].

One does not have to solve the entire maximization problem to get the risk aversion function. Only the first-order condition (F.O.C.) is needed. In this framework, it is easier to derive the risk aversion function by the martingale method rather than the Hamilton-Jacobi-Bellman (HJB) partial differential equation method. For the martingale method, the first step is to transform the intertemporal dynamic maximization problem into a static utility optimization problem, and the second step is to calculate the optimal investment and consumption strategy by martingale representation theorem. Here we only need to implement step one. As the investor's initial endowment is normalized to 1 unit, by applying the risk-neutral valuation method the risky stock price can be expressed as

$$\frac{1}{r^t}\int_0^\infty f_Q(S_T)S_T dS_T = 1$$

where $f_Q(S_T)$ is the risk-neutral probability density of the stock price on date $T$ across states. Now we can transform to a static problem by building up the Lagrangian function

$$\mathcal{L} = \int_0^\infty f_P(S_T)U(S_T)dS_T - \lambda\left[\frac{1}{r^t}\int_0^\infty f_Q(S_T)S_T dS_T - 1\right]$$

where $\lambda$ is the shadow price of the budget constraint. By taking the first-order derivative of the Lagrangian function with respective to the terminal stock price, we have

$$U'(S_T) = \frac{\lambda}{r^t}\frac{f_Q(S_T)}{f_P(S_T)}$$

The F.O.C. above must hold in equilibrium. According to the definition of absolute risk aversion, we also need the second-order derivative of the utility function with respect to $S_T$

$$U''(S_T) = \frac{\lambda}{r^t}\frac{f_Q'(S_T)f_P(S_T) - f_P'(S_T)f_Q(S_T)}{f_P^2(S_T)}$$

The absolute risk aversion function can be written as

$$A(S_T) = -\frac{U''(S_T)}{U'(S_T)} = -\frac{\dfrac{\lambda}{r^t}\dfrac{f_Q'(S_T)f_P(S_T) - f_P'(S_T)f_Q(S_T)}{f_P^2(S_T)}}{\dfrac{\lambda}{r^t}\dfrac{f_Q(S_T)}{f_P(S_T)}} = \frac{f_P'(S_T)}{f_p(S_T)} - \frac{f_Q'(S_T)}{f_Q(S_T)} \quad (2.4)$$

The relative risk aversion function can be written as

$$R(S_T) = -S_T\frac{U''(S_T)}{U'(S_T)} = S_T\left[\frac{f_P'(S_T)}{f_p(S_T)} - \frac{f_Q'(S_T)}{f_Q(S_T)}\right] \quad (2.5)$$

The absolute risk aversion coefficient is now a function of subjective density and risk-neutral density. The result is not surprising. If investors in the stock market are risk-neutral, the RND function depicts the actual beliefs of the investors. If, however, the investors are risk averse, we can study how risk averse they are by analyzing the difference between the RND density and the subjective density. As stressed by Ait-Sahalia and Lo (2000), people can infer any one of the following from the other two: (1) the representative investor's preference; (2) the subjective density; and (3) the risk-neutral density. Market completeness guarantees the uniqueness of the risk-neutral density and further guarantees the uniqueness of the risk aversion function.

Implied risk aversion function is both theoretically and empirically worth studying. The shape of absolute and relative risk aversion across wealth is crucial for choosing a utility function, and further influences the result of equilibrium pricing models. Yet the estimated risk aversion coefficients from different consumption based models are controversial. Measurement errors of the low-frequency consumption data might be one important reason. While applying the new approach, we can use high quality capital market data to estimate the subjective density as well as the risk-neutral density. If the assumptions of this model are appropriate, this approach should provide a more accurate estimated risk aversion function.

# 3. Density Estimation

Absolute risk aversion can be written as a function of subjective density and risk-neutral density. To derive the risk aversion function, one needs to estimate the subjective density and the RND. In this section, methods regarding subjective density and RND estimation will be discussed.

## 3.1 Subjective Density Estimation

Most of the subjective density estimation methods assume that investor's future expectation is reflected by historical information. In other words, these methods assume that future returns or price density are predictable based on past information. The subjective density function can be estimated parametrically or nonparametrically. The nonparametric kernel density estimation (KDE) method is used by Jackwerth (2000), Ait-Sahalia and Lo (2000) and Perignon and Villa (2002). The parametric Monte Carlo simulation method is adopted by Rosenberg and Engle (2002) and Hordahl and Vestin (2005).

In addition, these subjective density estimation methods assume that the density function is stationary, i.e. the underlying stochastic process has not changed during the estimation period [(Bliss and Panigirtzoglou (2004)]. This is an important concern in time series analysis. One needs to make sure that there is no structural change during the estimation period, namely the fundamental factors that influence the underlying data-generating process have not changed. This guarantees the stationarity of the density function.

### 3.1.1 Nonparametric Kernel Density Estimation Method

Nonparametric methods do not make any assumptions about the distribution. Hence, the KDE method does not impose any specific type of stochastic process on the price series, and the estimated subjective density can be in any shape.

The intuition behind the KDE method is similar to that behind histogram, which depicts the density function of a group of data. KDE estimates a smoothed density function by a kernel. The subjective density function estimated by KDE is

$$\hat{f}_P(r) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{r - r_i}{h}\right)$$

where $K(\ )$ is the kernel that integrates to one; $n$ is the number of observations; $h$ is the bandwidth; $r_i$ is the stock return observed in period $i$.
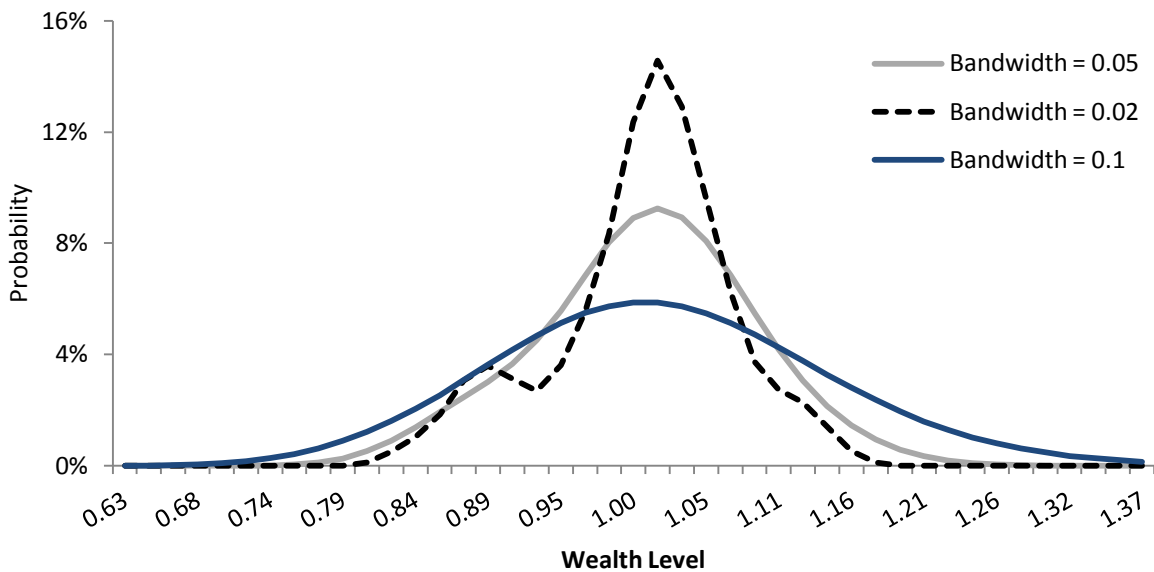


**Figure 1 Subjective Densites across Wealth**
The subjective density is estimated by KDE method. The subjective density describes the distribution of the investor's wealth level on 15.Jun.2012. The wealth level is defined as the stock price on 15.June.2012 divided by the stock price on 16.Apr.2012. The estimation period is from 01.Jun.2007 to 16.Apr.2012. We use a selection of different bandwidths. The kernel is Gaussian.

When implementing the KDE method, one needs to decide the bandwidth, type of kernel and estimation period. Selecting too large a bandwidth will result in an over-smoothed density function, but if the bandwidth is too small the density function might be under-smoothed. Silverman (1986) suggests the bandwidth $h = 1.06\,\sigma/n^{0.2}$ for Gaussian kernel. Yet Jackwerth (2000) uses a bandwidth of $h = 1.8\,\sigma/n^{0.2}$ and claims that other selections will not change the basic finding of his paper. It seems that there is no common rule of thumb for bandwidth selection. Therefore, I will choose different bandwidth under different circumstances. I follow Jackwerth (2000) and Perignon and Villa (2002) who adopt the Gaussian kernel

$$K(r) = \frac{1}{\sqrt{2\pi}} e^{-r^2/2}$$

### 3.1.2 Parametric Monte Carlo Estimation Method

Parametric methods estimate the parameters of a specific time series model. The specification of the model can be determined by, for example, the Box-Jenkins methodology. Next, we run the out-of-sample forecast on the selected model, say ARMA (p, q), with simulated shocks. After repeating a number of times (e.g. 100 000 times), we obtain a group of predictions on a specific day. Then we can derive the subjective density on that day by the kernel density method.

Hordahl and Vestin (2005) use the CIR interest rate model [Cox, Ingersoll and Ross (1985)] and assume that the term structure of short term interest rate follows a mean-reverting process. The discretized mean-reverting process is calibrated using historical data. The subjective density is then obtained by Monte Carlo simulation.

Rosenberg and Engle (2002) criticize the KDE estimation method used by Jackwerth (2000) and Ait-Sahalia and Lo (2000) and indicate that those methods assume that investors form probability beliefs by equally weighting events over the estimation period and disregarding previous event. Rosenberg and Engle further stress that future events depend more on recent information. They generate the subjective density by a discrete-time stochastic volatility model, the ARMA-GARCH model.

## 3.2 Risk-Neutral Density Estimation

It is not possible to derive the implied risk aversion function until we can empirically extract the RNDs. The theory of deriving RND function from option prices dates from the 1970s [Ross (1976), Breeden and Litzenberger (1978)], but the empirical research blooms during the 1990s [Rubinstein (1994), Jackwerth and Rubinstein (1996), Ait-Sahalia and Lo (1998), Bahra (1997) and Malz (1997)]. The application of RND function includes illiquid or exotic options pricing [Rosenberg (1998)], risk management [Ait-Sahalia and Lo (2000)], central bank policy making and event study [Bahra (1997), Castren (2005), Hordahl and Vestin (2005)].

Theory of RND extraction is discussed in subsection 3.2.1.Then I categorize the theory into two groups, and briefly review the estimation methods that fall into each group in subsection 3.2.2 and 3.2.3 respectively.

### 3.2.1 Theory of RND Extraction

Under no arbitrage framework, the price of an asset equals the expected future payoffs under risk-neutral measure $Q$ discounted by risk-free rate. Thus the prices of the call option and the put option[3] can be expressed as

$$C_{Call} = e^{-rt}E^Q[Max(S_T - K)] = e^{-rt}\int_0^\infty f_Q(S_T)Max(S_T - K, 0)\,dS_T \quad (3.1)$$

$$C_{Put} = e^{-rt}E^Q[Max(K - S_T)] = e^{-rt}\int_0^\infty f_Q(S_T)Max(K - S_T, 0)\,dS_T \quad (3.2)$$

Given the risk-neutral density, risk-free rate and future payoffs, we can compute the option price. Conversely, if we know the option price, the risk-neutral density can be backed out using equation (3.1) and (3.2). Estimation methods founded on this theory attempt to find a density function that minimize the distance between the observed option price and the calculated price. These methods are illustrated in section 3.2.2.

On the other hand, Breeden and Litzenberger (1978) prove that the state price can be approximated by a butterfly spread, which can be constructed by selling (buying) two call (put) options with the same strike price $K_1$ while buying (selling) two call (put) options with strike prices $K_0$ and $K_2$ respectively, where $K_0 < K_1 < K_2$ and $K_1 - K_0 = K_2 - K_1$.

The panel A and B of figure 2 show the payoffs of two butterfly spreads with strike prices of 9.5, 10, 10.5, and 9.9, 10 and 10.1 respectively. The payoff of the first butterfly spread at $S_T = K_1$ equals $K_1 - K_0 = K_2 - K_1 = 10.5 - 10 = 10 - 9.5 = 0.5$ while the payoff of the latter one is 0.1. As the distances between the adjacent strike prices $K_1 - K_0$ and $K_2 - K_1$ become narrower, the payoff of the butterfly spread divided by $K_1 - K_0$ or $K_2 - K_1$ is closer to that of the pure security which pays one unit of currency when $S_T = K_1$. This is demonstrated in panel C and D

---

[3] Unless specified, all the options mentioned in the thesis are European-style.
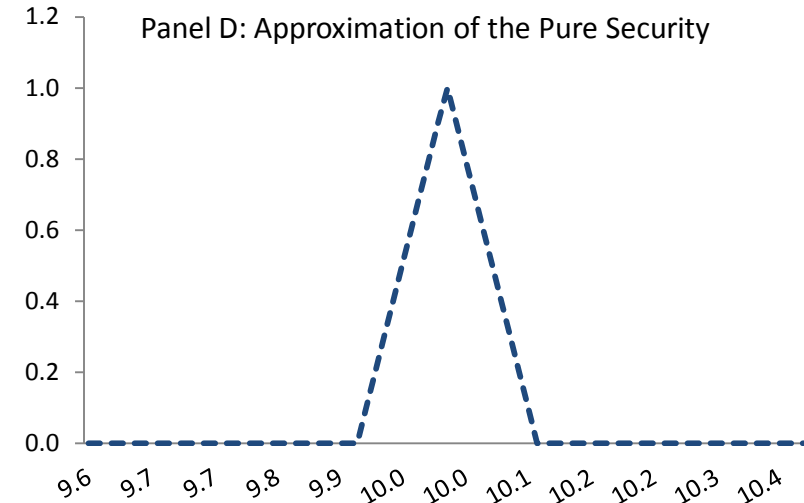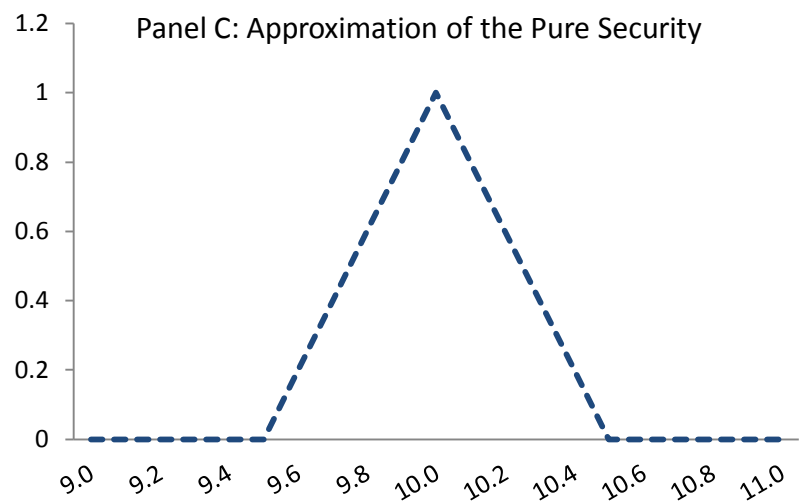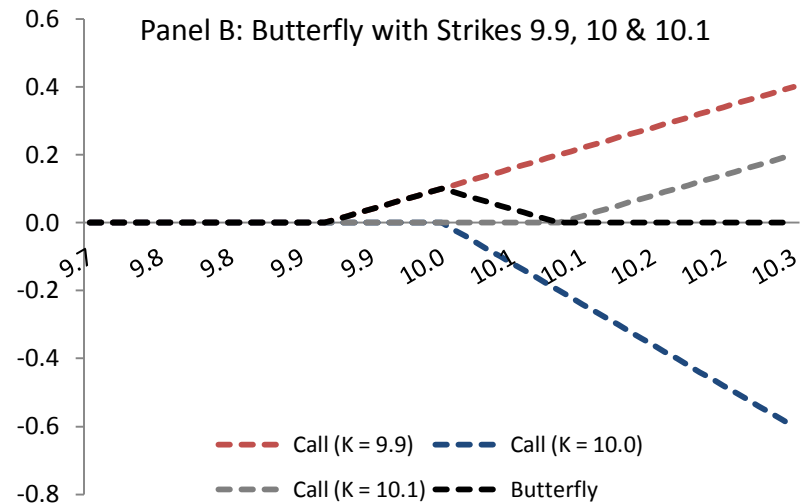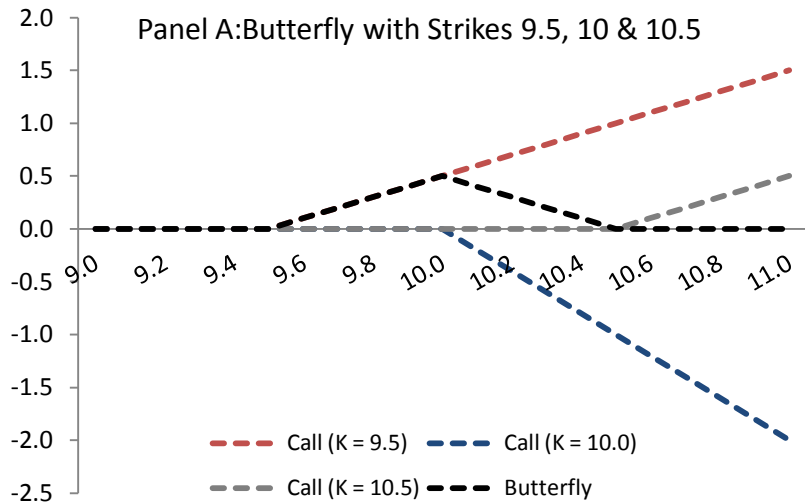
**Figure 2 Payoff of the Butterfly Spread and the Approximation of the Pure Security**
Panel A and B show the payoffs of two butterfly spreads. They are constructed by call options. Panel C and D demonstrate the payoff of the butterfly spreads divided by the distance between the strike prices of two adjcent options.

of figure 2. Therefore, the value of a pure security that pays one unit of currency at $S_T = K_1$ approximately equals the value of the butterfly divided by $K_1 - K_0$ or $K_2 - K_1$.

The price of the pure security, which pays one unit of currency when the terminal price of the underlying asset equals $S_T$, can be expressed as

$$P_0(S_T = K, \Delta K) = \lim_{\Delta K \to 0} \frac{[C_{Call,K+\Delta K} - C_{Call,K}] - [C_{Call,K} - C_{Call,K-\Delta K}]}{\Delta K} \quad (3.3)$$

Equivalently, we have

$$\lim_{\Delta K \to 0} \frac{P_0(S_T = K, \Delta K)}{\Delta K} = \lim_{\Delta K \to 0} \frac{[C_{Call,K+\Delta K} - C_{Call,K}] - [C_{Call,K} - C_{Call,K-\Delta K}]}{(\Delta K)^2} \quad (3.4)$$

The left-hand side of the equation is the state price density while the term on the right-hand side is the second order derivative of the call option with respect to the strike price. We know that the state price density equals the RND discounted by risk-free rate. Therefore, RND function can be written as

$$f_Q(S_T) = e^{rT} \frac{\partial^2 C}{\partial K^2}\bigg|_{K=S_T}$$

If we have a continuous function of option prices in terms of the strike price, we can derive the RND function by taking the second-order derivative with respective to the strike price. To get a positive risk-neutral probability across strike price, the price of the call has to be a convex function of strike price. Methods discussed in section 3.2.3 try to derive the RND function based on the theory of Breeden and Litzenberger (1978).

Theoretically, the RND function generated from these two thoughts are consistent with each other. This is proved by Malz (1997) and Figlewski (2008). If we take the second-order derivative of the equation (3.1), we have

$$\frac{\partial C_{Call}}{\partial K} = \frac{\partial \left[ e^{-rt} \int_K^\infty f_Q(S_T)(S_T - K) \, dS_T \right]}{\partial K} = e^{-rt} \left[ -Kf_Q(K) + Kf_Q(K) - \int_K^\infty -f_Q(S_T) \, dS_T \right]$$

$$= e^{-rt} \int_K^\infty -f_Q(S_T) \, dS_T = -e^{-rt}[1 - F_Q(K)]$$

Reorganize the terms and we have

$$F_Q(K) = e^{rt} \frac{\partial C_{Call}}{\partial K} + 1$$

By taking the second-order derivative with respect to $K$, the desired result can be obtained

$$\frac{\partial^2 C_{Call}}{\partial K^2} = e^{-rt} f_Q(K)$$

If options with any strike price are available, the market is complete. In a complete market, there is one and only one RND function. In theory, estimation methods based on either the risk-neutral pricing or the theory of Breeden and Litzenberger (1978) should give us identical RND function. However, the empirically estimated RND functions may not be the same due to market imperfections and different estimation assumptions.

The methods discussed in the following two subsections are based on the two thoughts discussed above. There are some other ways to categorize the RND estimation methods, e.g. Jackwerth (2004) classifies them as nonparametric and parametric methods according to the statistical properties. Note that there are many approaches to back out the RND function from options market, and in this thesis only a selection of the methods are reviewed. More comprehensive reviews have been done by Jackwerth (2004) and Taylor (2005).

### 3.2.2 Risk-Neutral Pricing: Double Lognormal Method

According to the risk-neutral pricing theory, the price of any asset equals its expected future payoffs under risk-neutral measure discounted by the riskless rate. Consequently, one can compute a theoretical option price using a predetermined density function. By minimizing the sum of squared differences between the theoretical prices and the observed market prices, one can approximate the true RND function. This is done by appropriately choosing the parameters of the predetermined density function. To approximate the RND function, Bahra (1997) uses a mixture of two lognormal density functions, Melick and Thomas (1997) adopt a mixture of three lognormal density functions, Ritchey (1990) estimates the RND function with a mixture of normal distributions, while Bookstaber and McDonald (1987) propose the generalized

distribution method which uses four parameters to capture the mean, volatility, skewness and kurtosis of a density.

On the other side, Rubinstein (1994) derives RNDs from the implied binomial tree model. The implied binomial tree is closely related to the Cox-Ross-Rubinstein binomial tree. Similarly, RNDs can be extracted from the implied binomial tree by solving a series of equations. However, given that only limited number of discrete strike prices is observed, the risk-neutral probabilities obtained are discrete and the probabilities of the tails are not available. So interpolation and extrapolation are required. Further, Jackwerth and Rubinstein (1996) study various RND estimation methods which are based on no arbitrage relation. They introduce non-quadratic objective functions, including the "absolute difference function" and "maximum entropy function". Their method is an innovation of Rubinstein's (1994) implied binomial tree.

In this thesis, Bahra's double lognormal method is applied. Bahra uses a combination of two lognormal density functions, and estimates the parameters in the density function by minimizing the sum of squared differences between the observed option prices and the prices calculated by equations (3.1) and (3.2).

The lognormal mixture method is first proposed by Ritchey (1990). When deciding the number functions, one should consider the trade-off between better fit of the RND function and less degrees of freedom (as one needs to estimate more parameters). The estimated RND function is said to be a weighted average of lognormal density functions

$$f_Q(S_T) = \sum_{i=1}^{k} [\omega_i L(\alpha_i, \beta_i; S_T)]$$

$$\sum_{i=1}^{k} \omega_i = 1$$

Bahra (1997) suggests a mixture of two lognormal density functions so that only five parameters $\alpha_1, \beta_1, \alpha_2, \beta_2, \omega$ will be estimated while the function still fits the actual prices well. The option prices, based on risk-neutral pricing theory, can then be expressed as

$$C_{Call} = e^{-rt} \int_0^\infty [\omega L(\alpha_1, \beta_1; S_T) + (1 - \omega)L(\alpha_2, \beta_2; S_T)] Max(S_T - K, 0) \, dS_T \quad (3.5)$$

$$C_{Put} = e^{-rt} \int_0^\infty [\omega \, L(\alpha_1, \beta_1; S_T) + (1-\omega) L(\alpha_2, \beta_2; S_T)] Max(K - S_T, 0) \, dS_T \quad (3.6)$$

Equations (3.5) and (3.6) can be solved analytically. The derivation, however, is quite lengthy, and is given in appendix B. The expressions for the call and put are

$$C_{Call} = e^{-rt} \left\{ \omega \left[ e^{\alpha_1 + \frac{\beta_1^2}{2}} N(d_1) - KN(d_2) \right] + (1-\omega) \left[ e^{\alpha_2 + \frac{\beta_2^2}{2}} N(d_3) - KN(d_4) \right] \right\} \quad (3.7)$$

$$C_{Put} = e^{-rt} \left\{ \omega \left[ KN(-d_2) - e^{\alpha_1 + \frac{\beta_1^2}{2}} N(-d_1) \right] + (1-\omega) \left[ KN(-d_4) - e^{\alpha_2 + \frac{\beta_2^2}{2}} N(-d_3) \right] \right\} \quad (3.8)$$

where

$$d_1 = \frac{\alpha_1 + \beta_1^2 - ln(k)}{\beta_1}$$

$$d_2 = d_1 - \beta_1$$

$$d_3 = \frac{\alpha_2 + \beta_2^2 - ln(k)}{\beta_2}$$

$$d_4 = d_3 - \beta_2$$

Given the observed and computed prices for put and call, we can estimate those five coefficients by solving the following minimization problem

$$\underset{\alpha_1, \beta_1, \alpha_2, \beta_2, \omega}{Min} \sum_{i=1}^n \left( C_{Call,i} - \hat{C}_{Call,i} \right)^2 + \sum_{i=1}^n \left( C_{Put,i} - \hat{C}_{Put,i} \right)^2$$

$$+ \left[ \omega e^{\alpha_1 + \frac{\beta_1^2}{2}} + (1-\omega) e^{\alpha_2 + \frac{\beta_2^2}{2}} - Se^{(r-d)t} \right]^2$$

$$s.t. \beta_1, \beta_2 > 0 \text{ and } 0 \leq \omega \leq 1$$

The first and second terms minimize the sum of squared errors between the observed and computed call and put options prices across different strike prices. The last term minimize the distance between the mean of implied RND and the futures price. Under no arbitrage condition,

the cost of carry[4] relation must hold and thus $F = Se^{(r-d)t}$. On the other side, the forward price is an unbiased estimation of future spot price and should equal the mean of the implied RND

$$F = Se^{(r-d)t} = E[S_t] = \omega \, exp\left(\alpha_1 + {\beta_1}^2/2\right) + (1 - \omega) \, exp\left(\alpha_2 + {\beta_2}^2/2\right).$$

Advantages of double lognormal approach are obvious. The model is parsimonious and we only need to estimate five coefficients. Therefore, unlike the kernel regression method discussed later, double lognormal approach performs well with limited data. Moreover, unlike some methods that produce negative probabilities or require interpolation and extrapolation, this method will always result in positive probabilities and smooth curve.

However, as the predetermined density function is a mixture of two lognormal densities and the shape of the RND depends only on five coefficients, it might not fully capture the properties of the observed option prices. Sometimes even if the objective function is minimized, the difference between the actual option price and the calculated price is still quite large. Another practical problem is that sometimes the mixed lognormal density could be bimodal if the mean of the first component is significantly different from that of the second component. Furthermore, the method assumes that the options are correctly priced. If the traded options do not fully capture investors' beliefs, it might be inappropriate to use the estimated RND function for asset pricing, policy making, risk management or other purposes. As indicated by Bahra (1997), illiquid option market is not a good choice for estimating RND function.

### 3.2.3 Breeden and Litzenberger (1978): Fast and Stable Method

Inspired by Breeden and Litzenberger (1978), RND function can be estimated by taking the second-order derivative of the option price function with respective to strike price. However, strike prices of options in the market are not continuous. Moreover, the range of the strike prices is limited and the options with extremely low or high strike prices are not available. Without information about option prices that have extremely the low and high strike price, we are unable to accurately estimate the tails of the RND. One direct approach is to fit a predetermined

---

[4] When the underlying asset is a dividend paying stock, dividend yield has to be considered; if the underlying asset is a type of foreign currency, foreign interest rate is added; present value of interest payments have to be subtracted from the initial price when the underlying asset is a fixed-income security; for commodities, storage cost and convenience yield have to be considered. See Hull (2006) for a thorough explanation.

function or a mixture of several predetermined functions to the observed option prices through strike price. This is similar to the double lognormal method presented in section 3.2.2. Bates (1991) uses a cubic spline[5] to fit the observations while Yatchew and Hardle (2006) use model-free least square method to estimate the state price density by fitting the option prices across strike prices.

The advantage of the fitting method is that it is not nested in the Black-Scholes framework so it is not restricted by the assumptions of Black-Scholes model. Nevertheless, as pointed out by Jackwerth (2004), under this approach the tails of the estimated RND function might still be inaccurate. For example, for parametric fitting methods, parameters that minimize the errors between the predetermined function and the actual prices might not be optimal because they do not include the potential unavailable options with extremely low and high strike price. So the shape of the tails actually depends on the observations that do not lie in the tails.

Alternatively, Shimko (1993) suggests fitting the implied volatilities across strike. An important strength of fitting the function $\sigma_{Implied}(K)$ is that unlike option prices fluctuate sharply across strike prices, the plots of implied volatility against strike price are smoother [Jackwerth (2004)]. When fitting the option prices curve, we minimize the distance between the actual prices and the fitted function. Hence we put more weights on expensive in-the-money options than out-of-the-money options. However, by fitting the smoother implied volatility function, this problem can be alleviated.

Before going through the implied volatility fitting method, we briefly explain the concept of implied volatility. In the Black-Scholes option pricing formula, there are five inputs[6], namely underlying asset price, annualized volatility of underlying asset returns, strike price, interest rate and time to maturity. Conversely, given the observed option price, any one of the five variables can be backed out. The most interesting variable among the five is the implied volatility derived from the Black-Scholes model. It is said that implied volatilities have superior predictability for future volatilities compared to historical volatilities, volatilities calculated by GARCH or other

---

[5] A spline is a smoothed polynomial function, where a polynomial is in the form $C(K) = \alpha_0 + \alpha_1 K + \alpha_2 K^2 + \cdots + \alpha_n K^n$.

[6] For options with underlying assets that are dividend paying stocks or currency, dividend yield and foreign interest rate should be considered. Slight changes of the Black-Scholes formula can be made to value options on these assets.

time series models [Mayhew (1995)]. However, when plotting the implied volatility curve across strike prices, we realize that it is a convex function of strike price. This violates one of the Black-Scholes assumptions[7] and further motivates research about stochastic volatility option pricing model [Hull and White (1987), Heston (1993)] and deterministic volatility model [Runbinstein (1994)]. However, the implied volatility fitting method and other related estimation methods use the Black-Scholes formula only as a "transformation mechanism" that first transforms actual option prices into implied volatilities and then converts fitted implied volatility curve into option price curve.

By fitting the volatility smile curve, Shimko's (1993) method assumes that implied volatility curve and strike price have a nonlinear relationship. In the Black-Scholes framework, the function of an option is $C(S_t, T, K, \sigma, r)$. Therefore, we can derive the RND function by taking the second-order derivative regarding strike price. But if the implied volatility is related to strike price $K$, the function of the option can be rewritten as $C(S_t, T, K, \sigma_{Implied}(K), r)$. As strike price will influence the value of the option directly and via volatility, we should find the continuous implied volatility function with respect to strike price and then replace it back into the Black-Scholes formula.

The function $\sigma_{Implied}(K)$ can be fitted by polynomials, splines, kernel regressions or other methods. Many estimation methods consider both the smoothness of the estimated volatility function and the fitness to actual data. Jackwerth and Rubinstein (1996) and Bliss and Panigirtzoglou (2002), for example, introduce the trade-off coefficient or the so-called penalty function to do so. On the other side, the kernel regression is labeled as easy to implement and reliable by Jackwerth (2004). Ait-Sahalia and Lo (1998, 2000) assume that the implied volatility is influenced by all five variables from the Black-Scholes and estimate the function by kernel regression. Due to its data intensity, this method will not be used in the thesis. An illustration of kernel regression with application in finance is provided by Campbell et al. (1997).

---

[7] The explicit assumptions originally made by Black and Scholes (1973) are:
   (1)  The no arbitrage condition holds;
   (2)  The capital market is perfect, e.g. no short sell restrictions, no transaction costs, securities can be subdivided arbitrarily, securities are traded continuously;
   (3)  The underlying security does not pay dividends;
   (4)  The data generating process of the underlying asset is geometric Brownian motion;
   (5)  Both interest rate and the volatility of the asset returns are constant.

Malz (1997) proposed another frequently adopted method that interpolates implied volatilities across deltas. Delta is one of the Greek letters which measures the sensitivity of the underlying asset price $S_t$ to option price $C(S_t, T, K, \sigma, r)$. In other words, it is the first-order derivative of option price to underlying asset price. In the Black-Scholes framework, the delta[8] of a call option can be written as

$$\Delta = \frac{\partial C}{\partial S} = N\left[\frac{ln(S_0/K) + (r + \sigma^2/r)T}{\sigma\sqrt{T}}\right]$$

As delta is a monotonic function of strike price, the function $\sigma_{Implied}(K)$ can be derived via delta. The intuition of this method is the same as those discussed above in this section. But this method is more complicated in the sense that it requires a transformation from delta to strike.

Last but not least, I will introduce Jackwerth's so-called fast and stable method. The fast and stable method will also be used in this thesis. Jackwerth (2000) indicates that the fast and stable method can find a smooth risk-neutral density which also explains the observed option prices.

$$\underset{\sigma_j}{Min} \sum_{j=0}^{n} (\sigma_j'')^2 + \lambda \sum_{i=1}^{m} \left(\frac{\sigma_i - \bar{\sigma}_\iota}{SD_i}\right)^2$$

where $\sigma_j$ and $\sigma_i$ are estimated implied volatility associated with strike price $K_j$ and $K_i$ respectively; $\bar{\sigma}_\iota$ is the implied volatility derived from actual option prices; $\lambda$ is the trade-off parameter; $\sigma_j''$ is the second derivative of the implied volatility curve and can be approximated by $(\sigma_{j-1} - 2\sigma_j + \sigma_{j+1})/\Delta^2$; $SD_i$ is the standard deviation of $\bar{\sigma}_\iota$.

The first part of the objective function finds the smallest squared second-order derivative of the implied volatility curve since we want a smooth implied volatility curve. The second term minimizes the distance between the observed implied volatility and the estimated one. The trade-off parameter is selected manually to balance the need to find a smooth implied volatility curve and the fitness of estimated curve. To solve the objective function above, a group of volatilities $\sigma_j$ and $\sigma_i$ are selected so that the curvature of the volatility curve is minimized and the estimated volatility curve agrees with the observed volatilities.

---

[8] This is for assets without dividend payment.

After deriving the implied volatility curve, we substitute it back to the Black-Scholes formula and calculate the RND function by taking the second derivative with regard to strike price. The closed-form solution is

$$P(S_j) = r^T \left\{ \frac{r^{-T} n(d2_j)}{S_j \sigma_j \sqrt{T}} \left( 1 + 2 S_j \sqrt{T} d1_j \sigma_j' \right) + s d^{-T} \sqrt{T} n(d1_j) \left[ \sigma_j'' + \frac{d1_j d2_j}{\sigma_j} (\sigma_j')^2 \right] \right\}$$

where

$$d1_j = \frac{ln(sd^{-T})/(S_j r^{-T})}{\sigma_j \sqrt{T}} + \frac{1}{2} \sigma_j \sqrt{T}$$

$$d2_j = d1_j - \sigma_j \sqrt{T}$$

$d = 1 + dividend\ yield$

$S = index\ value\ today$

$\sigma_j' = first\ derivative\ of\ implied\ volatility\ with\ respect\ to\ strike\ price, and\ can\ be$

$approximated\ by\ (\sigma_{j+1} - \sigma_{j-1})/2\ \Delta$

Like we stated in the paragraphs above, fitting the option price curve itself by minimizing the sum of squared errors will put more weights on in-the-money options since the prices of these options are higher. Fitting the implied volatility curve can alleviate the problem. Another advantage of this method is that the curvature of the implied volatility curve can be minimized so that we can get a smooth curve. The disadvantage of this method is that the trade-off parameter is selected arbitrarily. A different trade-off parameter will result in a somewhat different RND function. For example, a trade-off parameter that closes to zero might result in a straight line. Since we use the Black-Scholes formula as the transformation mechanism, a constant volatility curve means that risk-neutral density is lognormal.

# 4. Implied Risk Aversion: An Empirical Example

In this section, I estimate the subjective density function by the GARCH Monte Carlo method. The RND function is estimated by Bahra's double lognormal method as well as Jackwerth's fast and stable method. I will then derive the implied risk aversion function using equation (2.4) in section 2.3. Furthermore, to examine the stability of the implied risk aversion function, robustness tests will be implemented.

## 4.1 Data

The empirical example is based on a dataset from Thomson Reuters Datastream. The dataset contains daily prices of FTSE 100 index options (ESX) that expire on June 2012, ranging from 06.Jun.2010 to 01.May.2012. The underlying asset of the FTSE 100 index option is the FTSE 100 index, which measures the price level of 100 largest companies (in terms of market capitalization) listed on the London Stock Exchange. FTSE 100 index options are European style and are traded on NYSE Liffe London. The expiration date of the FTSE 100 index option contract is normally the third Friday of the delivery month, i.e. 15.Jun.2012 in this example. For options contracts expired on June 2012, there are 56 different strike prices, ranging from 1600 to 8800.

In studies related to RND estimations, S&P 500 European style options (SPX) traded on the Chicago Board Options Exchange (CBOE) are frequently used [Jackwerth and Rubinstein (1996), Ait-Sahalia and Lo (1998), Figlewski (2008)]. S&P 500 index option is known as one of the most frequently traded options in the world. There are more than 100 strikes prices for contracts that have the same expiry date. Accordingly, it is convenient to extract the RND functions from S&P 500 index options. Yet fewer researchers choose FTSE 100 index options [Bliss and Panigirtzoglou (2002, 2004)]. This paper can be a supplement to existing literature.
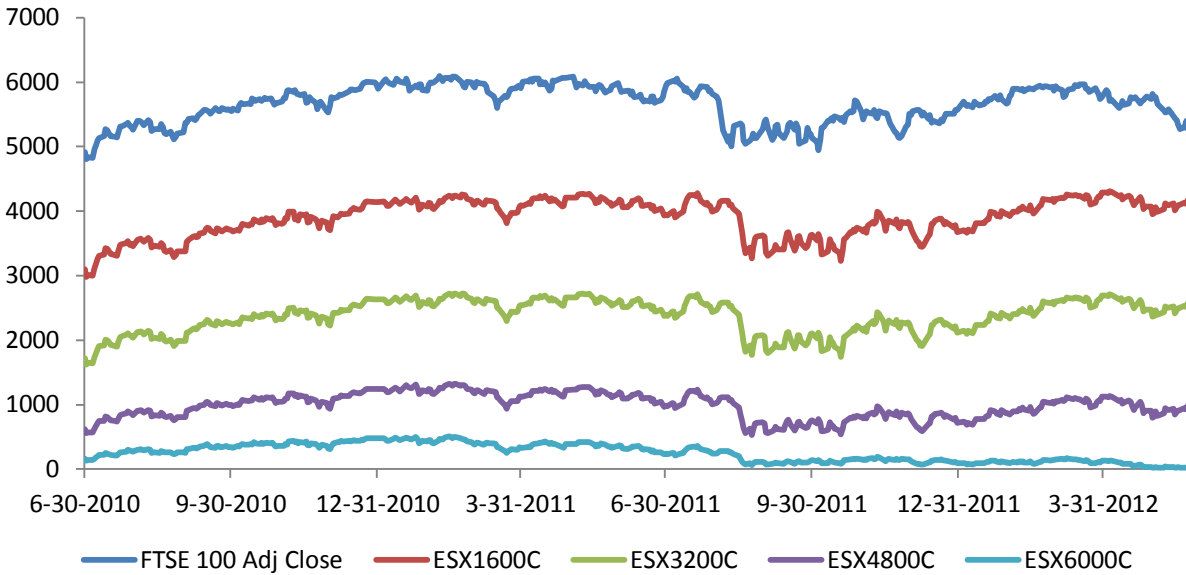
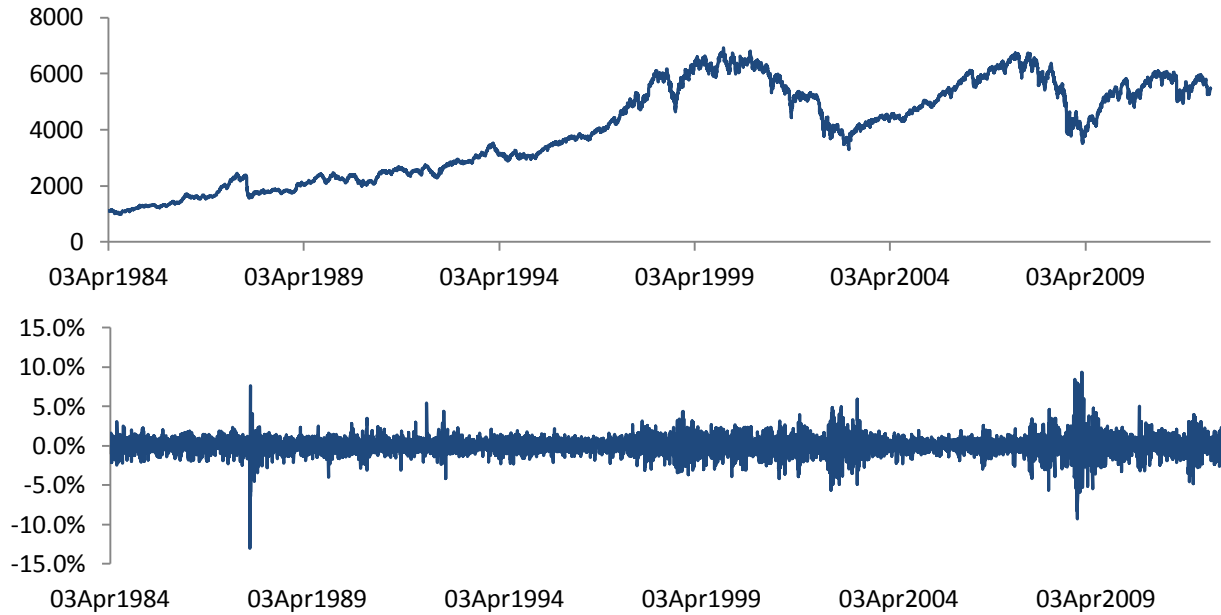**Figure 3 Adjusted Closing Price of FTSE 100 Index and Call Option Prices**
The strike prices of the four options are 1600, 3200, 4800 and 6000 respectively. The sample period is from 06.Jul.2010 to 01.May.2012.

## 4.2 Estimation

### 4.2.1 Subjective Density Estimation

In this empirical example, I estimate the one-month-ahead FTSE 100 subjective density. Let $I_T$ be the information set on date $T$, and $I_T$ is used to estimate the subjective density of the index level $P_{T+N}$, $N$ days after the trading day $T$. Here date $T$ is 16.Apr.2012 and date $T+N$ is 15.Jun.2012. From figure 4, we can see that the behavior of the prices and returns of FTSE 100 is consistent with the stylized facts of financial time series. That is, the FTSE 100 closing prices are non-stationary, and the logarithmic returns demonstrate the volatility clustering effect[9]. Before year 2000 the FTSE 100 index increases gradually and the volatility seems to be stable. After year 2000, however, the index starts to fluctuate, and the return series become more volatile, especially during the period of crash. It seems that the FTSE 100 has recovered from the financial crisis since 2009. Therefore, it is appropriate to use the estimation period 16.Apr.2009 to 16.Apr.2012 to generate the subjective density.

---

[9] I have used augmented Dickey-Fuller test to examine the non-stationarity of prices and the Lagrange multiplier test to examine the volatility clustering effect of returns. The unreported results indicate that these two stylized facts do exist.

**Figure 4 Adjusted Closing Price and Logarithmic Return of FTSE 100 Index**
The FTSE 100 index and the logarithmic return series are from 03.Apr.1984 to 01.May.2012.

I use the GARCH Monte Carlo method to estimate the subjective density function. The following the GARCH specification is used

$$R_t = \beta_0 + \varepsilon_t$$

$$h_{t+1}^2 = \alpha_0 + \alpha_1 \varepsilon_t^2 + \alpha_2 h_t^2$$

$$\varepsilon_{t+1} = h_{t+1} \epsilon_{t+1}$$

$R_t$ is the logarithmic return of FTSE 100 index on date $t$; $\beta_0$ is a constant and $\varepsilon_t$ is a shock for date $t$; the time-dependent standard deviation $h_{t+1}$ can be forecasted by the shock $\varepsilon_t$ and the standard deviation $h_t$, where $\varepsilon_t = R_t - \beta_0$; the shock $\varepsilon_{t+1}$ is the product of $h_{t+1}$ and $\epsilon_{t+1}$, where $\{\epsilon_t\}$ is a sequence of identically and independently distributed (i.i.d.) normal random variables with mean zero and variance one.

The information set $I_T$ is used to estimate the parameters of the model, and simulated series $P_{T+1}, \dots, P_{T+N}$ can be generated with $N$ standardized random residuals. After repeating 200, 000 times, we obtain 200,000 predicted prices $\{P_{T+N}^i\}_{i=1}^{200,000}$. Then the kernel method is used to estimate the subjective density based on these simulated observations.
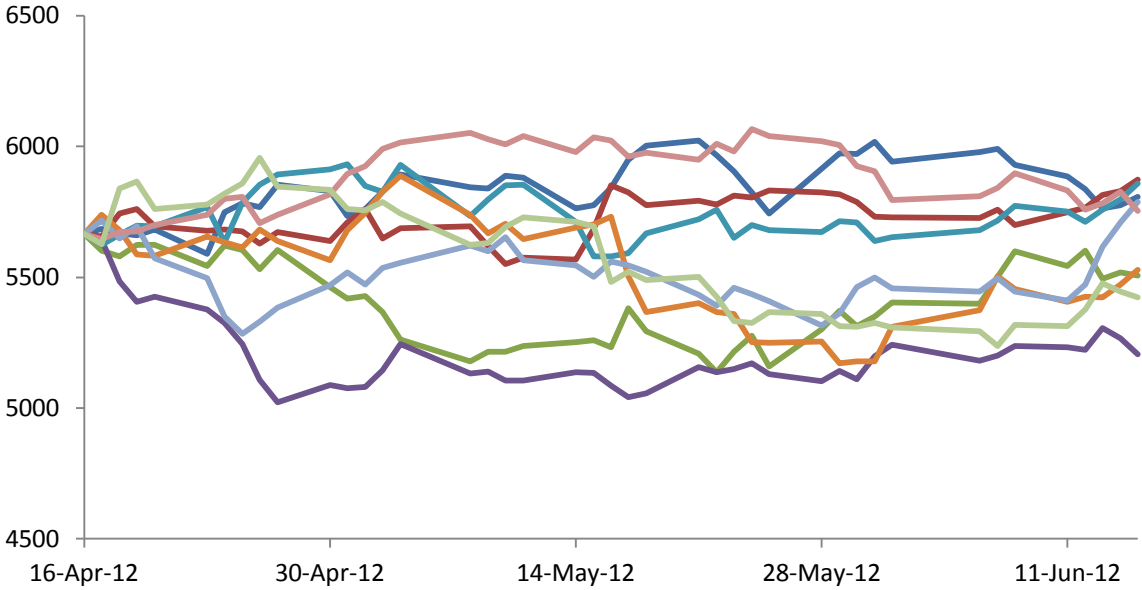
**Figure 5 FTSE 100 Index Simulation**
The estimation period is from 16.Apr.2009 to 16.Apr.2012. The GARCH (1, 1) model is used to forecast the FTSE 100 index level on 15.Jun.2012 with simulated shocks.
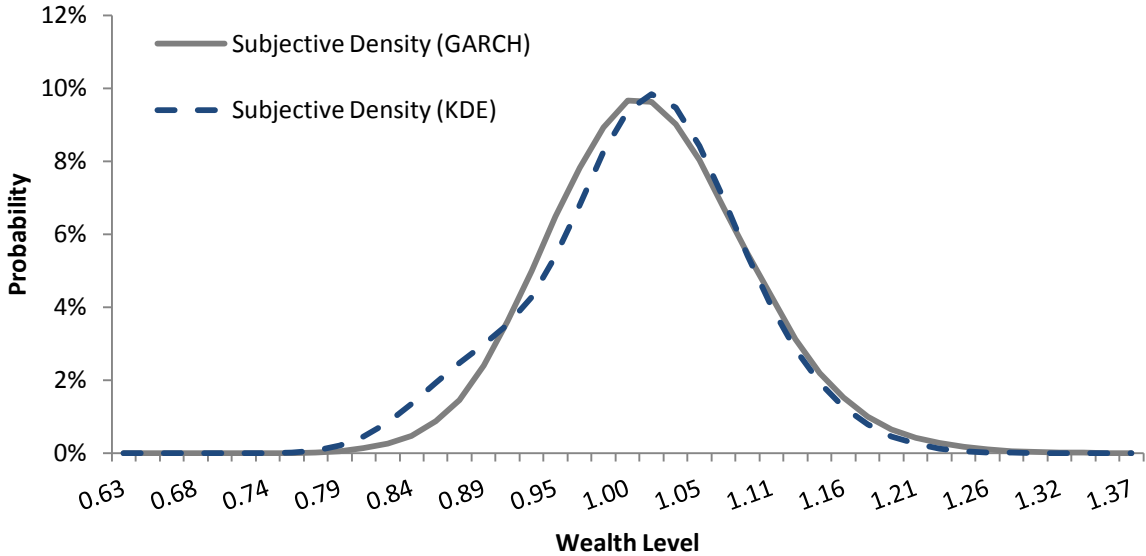


**Figure 6 Subjective Density Estimated by the GARCH Monte Carlo Method**
Subjective densities are estimated by the GARCH Monte Carlo method and the KDE method respectively. The estimation period is from 16.Apr.2009 to 16.Apr.2012. For the KDE method, Gaussian kernel is adopted, and the bandwidth is $h = 1.4\,\sigma/n^{0.2}$. The subjective densities depict the representative investor's wealth distribution on 15.Jun.2012.

Figure 6 depicts the empirical subjective densities on 15.Jun.2012. The subjective density estimated directly by the KDE method seems to have a fatter left tail, indicating that there is a larger probability of wealth decrease from 16.Apr.2012 to 15.Jun.2012. Just as criticized by Rosenberg and Engle (2002), the KDE method assumes that the probability beliefs are formed by equally weighting events over the estimation period, but previous events are disregarded. By equally weighting the events over the estimation period, the KDE method might give more weights to the left tail of the distribution due to two sudden market declines during the estimation period. On the other side, the GARCH model has the long memory property. It admits that events before the estimation period can also influence the formation of the subjective density. Therefore, I will only use the GARCH Monte Carlo method to estimate the implied risk aversion.

### 4.2.2 Risk-Neutral Density Estimation

When estimating the RND function, risk-free rates and dividend yields are needed. I obtain the dividend yields directly from Thomson Reuters Datastream. Risk-free interest rates, however, are more difficult to find. The 3-month risk-free rate is available from Datastream and the webpage of the Bank of England. Yet there is no information regarding 2-month risk-free interest rate. Unless the 2-month risk-free interest rate is equal to the 3-month rate, I am unable to accurately estimate the RND function on 15.Jun.2012 using the information by 16.Apr.2012. London Interbank Offer Rate (LIBOR) can be used to calibrate the interest rate required. But a spread should be subtracted from the LIBOR since the LIBOR is not a risk-free rate. Another approach suggested by Jackwerth and Rubinstein (1996) is to extract the risk-free interest rate from put-call parity and cost-of-carry relation

$$C_t + e^{-r(T-t)}K = e^{-d(T-t)}S_t + P_s$$
$$F_t = e^{(r-d)(T-t)}S_t$$

Since all the variables in the two equations above are observable except the interest rate, the interest rate can be backed out from the two equations

$$r = \frac{ln\dfrac{P_s + e^{-d(T-t)}S_t - C_t}{K}}{-(T-t)}$$

$$r = \frac{ln\frac{F_t}{S_t}}{T-t} + d$$

This method assumes that there is no arbitrage opportunity in the market. Moreover, factors like the bid-ask spread of the options and transaction costs are overlooked. Therefore, the derived interest rate might subject to different types of errors. The empirically estimated interest rate can even be negative. Bliss and Panigirtzoglou (2004) use different rates as proxy of risk-free rate to calculate the RNDs, e.g. the 3-month LIBOR, Federal funds rate, or the 3-month EuroDollar rate. They conclude that the choices of interest rate will have little impact on the RND estimation, since a 100 basis point change in the assumed interest rate will produce approximately a 2 basis point change in the measured at-the-money implied volatility for a 1-month contract, increasing to 5 basis points at the 6-month horizon. In this example, I will use the interest rate extracted from the put-call parity relation. When implementing the robustness test, I will estimate the RNDs with different interest rates.
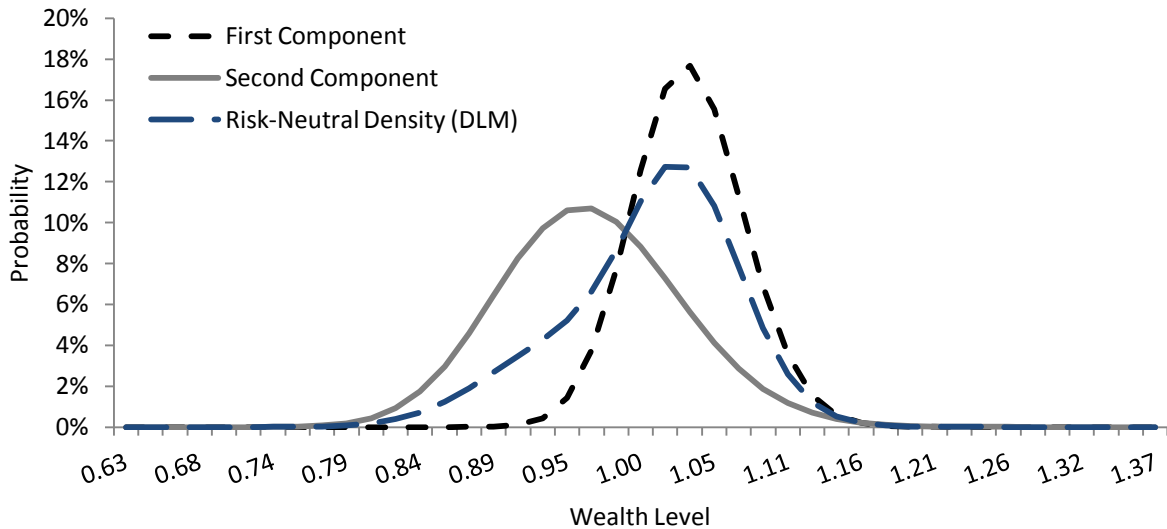


**Figure 7 Risk-Neutral Density estimated by Bahra's (1997) Double Lognormal Method**
The risk-neutral density is estimated by Bahra's double lognormal method. 7 options on 16.Apr.2012 are used. The risk-neutral density depicts the wealth distribution of the representative investor on 15.Jun.2012.

Although there are 56 different strikes for the FTSE 100 index options, some options are infrequently traded, especially the deep in-the-money options and the deep-out-of-the-money ones. These infrequently traded options may not fully incorporate investors' beliefs. Thus, for the

double lognormal method I will only use 7 options with strike prices that are closest to the FTSE 100 index level on 16.Apr.2012. For the fast and stable method, 9 such options will be selected.

When implementing the double lognormal method, we estimate five parameters of the objective function so that sum of squared differences between the theoretical prices and the actual market prices of the options is minimized. 7 options on 16.Apr.2012 are used to estimate the RND function on 15.Jun.2012. Figure 7 shows the first and second component of the RND function, and the combination of the two. The mean of the second component is about 0.95 while that of the first component is about 1.05. The kurtosis of the second component is larger. The RND function, as the combination of the two, exhibits a moderate mean and kurtosis.
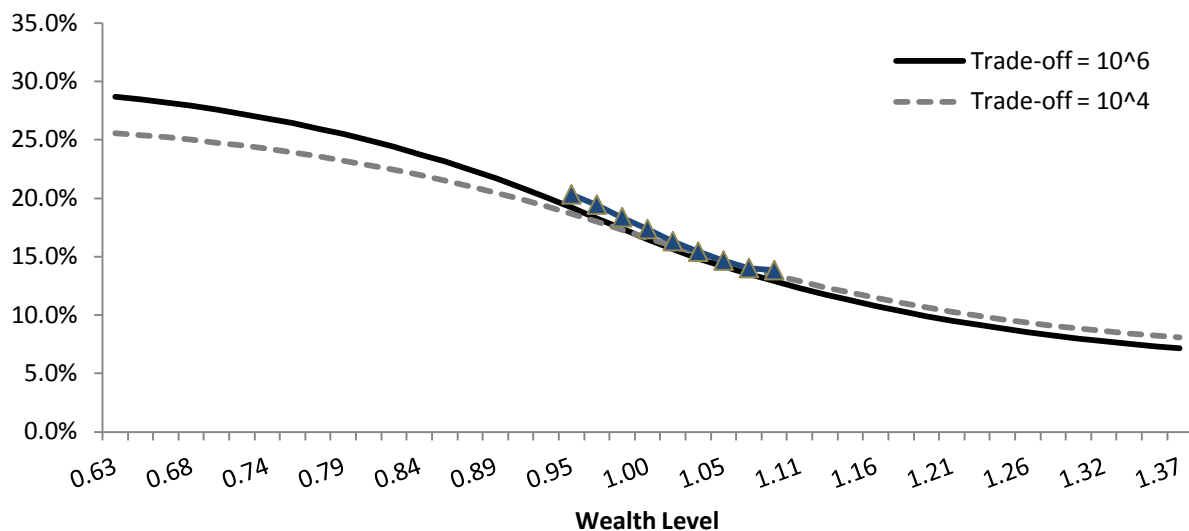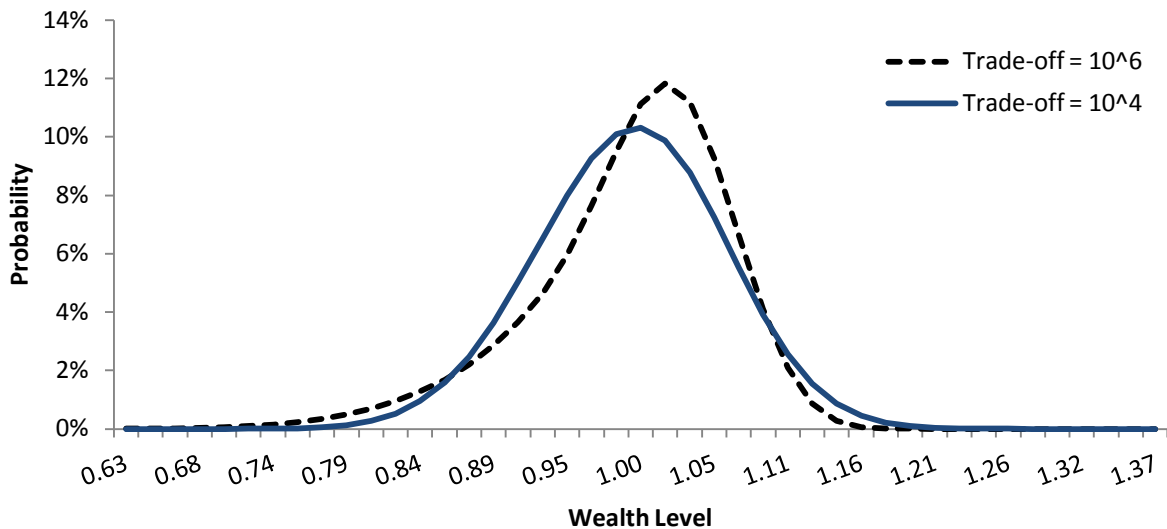


**Figure 8 Implied Volatility Curve estimated by Jackwerth's (2000) Fast and Stable Method**
The implied volatility curve is fitted by Jackwerth's (2000) fast and stable method with 9 options on 16.Apr.2012. Implied volatilities backed out from actual option prices are denoted by nine blue triangles.

Next we use the Fast and Stable method to estimate the RND function. The implied volatilities are backed out from the at-the-money (ATM) call option and eight adjacent options using Newton-Raphson interation method. We fit the the implied volatility function $\sigma_{Implied}(K)$ by the fast and stable method. When selecting the trade-off parameter, I put priotrity on the fittness of the implied volatility curve to the actual implied volatilities, but also consider the smoothness of the estimated density. The logic behind this criteria is that if the implied volatility curve does not fit the actual data well, the estimated density will not fully incorporate investors' beliefs. Figure

8 below shows two fitted implied volatility curves. It seems that a trade-off parameter of $10^6$ works well. A trade-off parameter higher than $10^6$ will lead to a under-smoothed RND.

After deducting the implied volaitlity curves, we use the Black-Scholes formula to transform the volatility curve to the option prices function. Then the RND function can be derived by taking the second-order derivative with respect to strike price. As expected, the RND function derived with a trade-off parameter of $10^4$ is smoother. However, the RND function with a trade-off parameter of $10^6$ better reflects market beliefts.
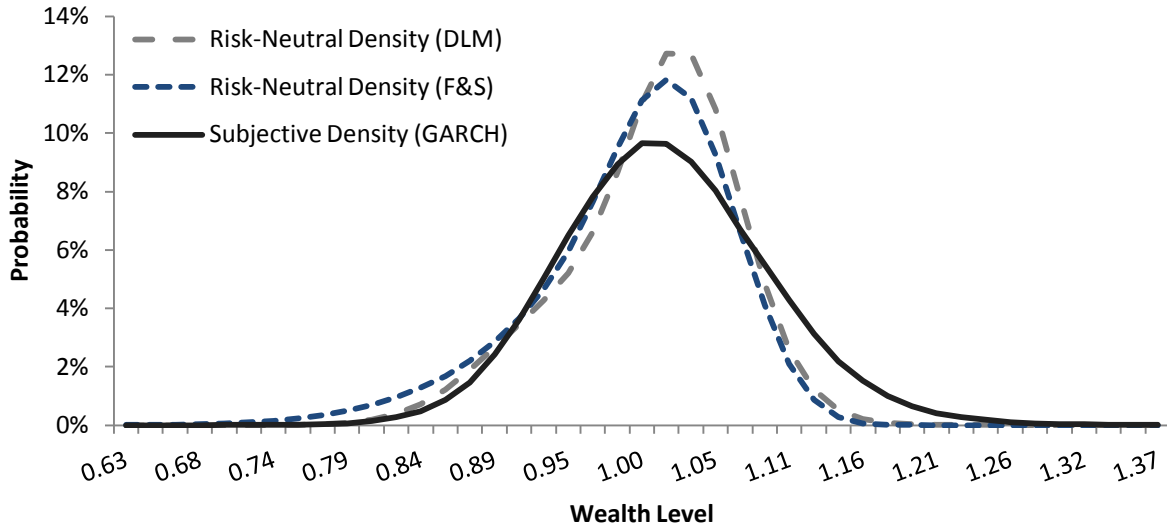


**Figure 9 Risk-Neutral Density estimated by Jackwerth's (2000) Fast and Stable Method**
The risk-neutral density is estimated by Jackwerth's (2000) fast and stable method with 9 options on 16.Apr.2012.

### 4.2.3 Implied Risk Aversion Estimation

Figure 10 displays the RND functions estimated by the double lognormal method as well as the fast and stable method, and the subjective density estimated by GARCH Monte Carlo method. The risk-neutral densities exhibit leptokurtic property. RNDs have more acute peaks, and the left tails are fatter than that of the subjective density. On the other hand, the subjective density is platykurtic. Further, subjective density indicates that there is a higher probability of obtaining a wealth level above 1.10. Jackwerth (2004) estimates the risk-neutral and actual probability distributions for S&P 500, German DAX 30 and UK FTSE, and concludes that the risk-neutral

distributions are leptokurtic and left-skewed (i.e. they have fatter left tail). Taylor (2005) points out that the RND contains more information than the subjective density. Just like implied volatility is a better predictor of future volatility than the historical volatility [Mayhew (1995)], RND function incorporates investors' beliefs while the subjective density is only a reflection of historical information.
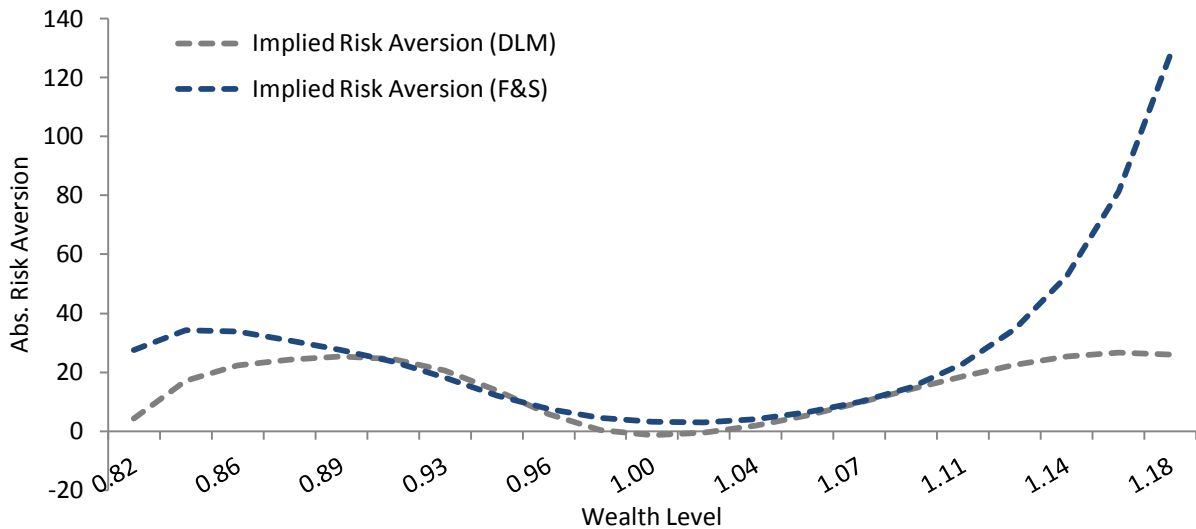


**Figure 10 Risk-Neutral Densities and Subjective Density**
The RNDs are estimated by double lognormal method and fast and stable method respectively. The subjective density is estimated by the GARCH Monte Carlo method. The densities depict the wealth distribution of the representative investor on 16.Jun.2012.

It seems that the RNDs estimated by the two different methods are in similar shape. We implement the Kolmogorov-Smirnov test (K-S test) to examine whether the RNDs estimated by the two approaches are statistically different. The null hypothesis is that the two distributions are the same, i.e. $H_0: f_{Q,DLM}(W_T) = f_{Q,F\&S}(W_T)$. Firstly of all, 10,000 observations are generated from each risk-neutral probability distribution. Next, two cumulative distribution functions with respective to the generated observations are formulated. Then we calculate the supremum of the set of absolute different between the two cumulative distributions, i.e. $D = sup|\hat{F}_{Q,DLM}(W_T) = \hat{F}_{Q,F\&S}(W_T)|$. Lastly, the critical value and p-value can be computed based on the maximum distance $D$. The p-value of the K-S test is 0.000, indicating that the null hypothesis that the RNDs estimated by the two different methods are the same is rejected at any conventional significance level. Thus, the RND calculated by the double lognormal method is statistically different from the one calculated by the fast and stable method.

Given the RND and the subjective density, we can calculate the implied risk aversion function. Implied risk aversion functions are derived using the GARCH estimated subjective density, and the RND estimated by either the double lognormal method or the fast and stable method. However, the risk aversion functions are inconsistent with the orthodox financial theory. The risk aversion functions are U-shaped. When the wealth level is higher than 1, the risk aversion increases dramatically. This is more prominent for the function estimated by the fast and stable method.



**Figure 11 Implied Risk Aversion Functions across Wealth**
The RNDs are estimated by the double lognormal method as well as the fast and stable method. The subjective density is estimated by the GARCH Monte Carlo method. The implied risk aversion describes the absolute risk aversion of the representative investor across wealth.

The implied risk aversion smile in figure 11 is also documented by other papers. Jackwerth (2000) estimates the absolute risk aversion function with options on the S&P 500 index from April 2, 1986 to December 29, 1995, while Ait-Sahalia and Lo (2000) calculate the relative risk aversion function using option prices in 1993. The RND function is estimated by the fast and stable method in Jackwerth's paper and the kernel regression method in the paper of Ait-Sahalia and Lo. Both of the papers estimate the subjective density function by the KDE method. Jackwerth indicates that after the 1987 stock market crash the risk aversion becomes an increasing function of wealth when the wealth level is higher than 0.99. On the other side, the implied risk aversion obtained by Ait-Sahalia and Lo (2000) is also an increasing function of wealth when the S&P 500 index level is higher than 500 at expiration. One exception is Perignon

and Villa (2002). They investigate the risk aversion function with the CAC 40 index options. Perignon and Villa replicate the paper of Jackwerth (2000). They find that the risk aversion function is a decreasing function of CAC 40 index, but the risk aversion becomes negative when the index is higher than a certain level. Nevertheless, in their paper the subjective density estimated by the KDE method does not capture the volatility clustering effect which exists in most financial time series.

## 4.3 Robustness Tests

Before exploring possible explanations for the oddly behaved risk aversion function, I will implement robustness test to examine the stability of the function. Empirically, the selection of estimation methods, estimation periods and dataset may influence the shape of the risk aversion function. For example, the RNDs estimated by the double lognormal method and the fast and stable method in figure 10 are statistically different, and can further influence the shape of the risk aversion functions.

The robustness test is divided into three steps. Firstly, the stability of the subjective density will be examined. I will use different GARCH-family models to derive the subjective density. Secondly, I will select different risk-free rates to estimate the RNDs, and for the fast and stable method different trade-off parameters will be chosen. Lastly, I will investigate whether the implied risk aversion function is still U-shaped under these different choices. Note that the robustness test here is a comparative statics analysis rather than a scenario analysis, i.e. in the robustness test only one element is changed at a time.

Glosten, Jagannathan and Runkle (1993) proposed an alternative ARCH-family model. For some financial time series, it could be the case that former positive and negative shocks have different impact on current shocks. Accordingly, the GJR-GARCH model introduces a dummy variable that equals one when the past shock $\varepsilon_t$ is positive. If the dummy variable is statistically significant at conventional level, there exists leverage effect. Moreover, most probability distributions of high-frequency financial data are leptokurtic, so I will employ the student t-distribution rather than the normal distribution when estimating the GJR-GARCH model with maximum likelihood method (MLM).
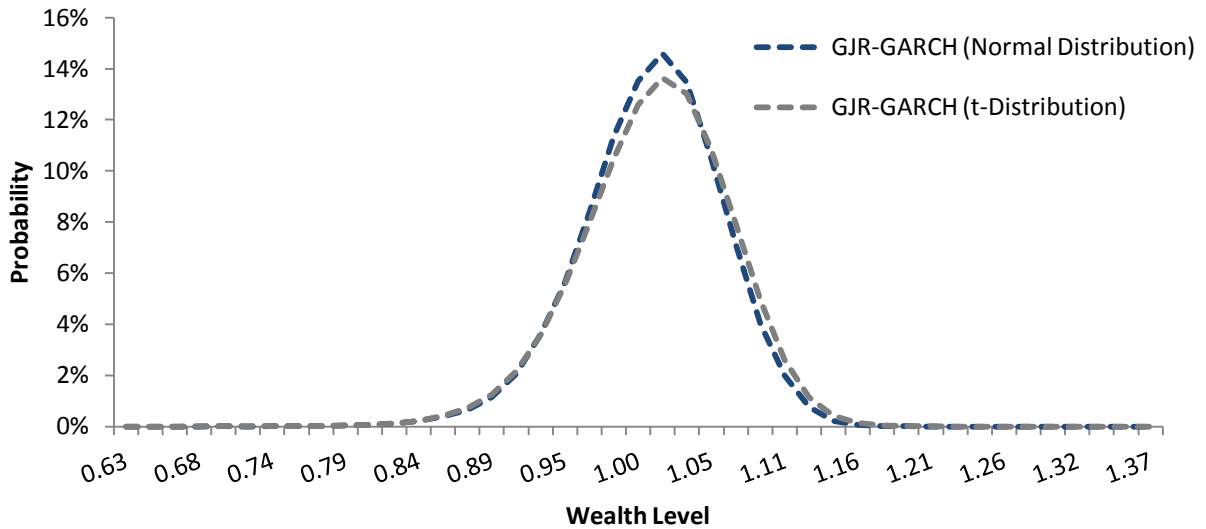
**Figure 12 Subjective Density estimated by the GJR-GARCH model**
The subjective densities are generated by the GJR-GARCH model. The parameters of the GJR-GARCH model are estimated by the maximum likelihood method (MLM). When implementing the MLM, the normal distribution and the t-distribution are used.
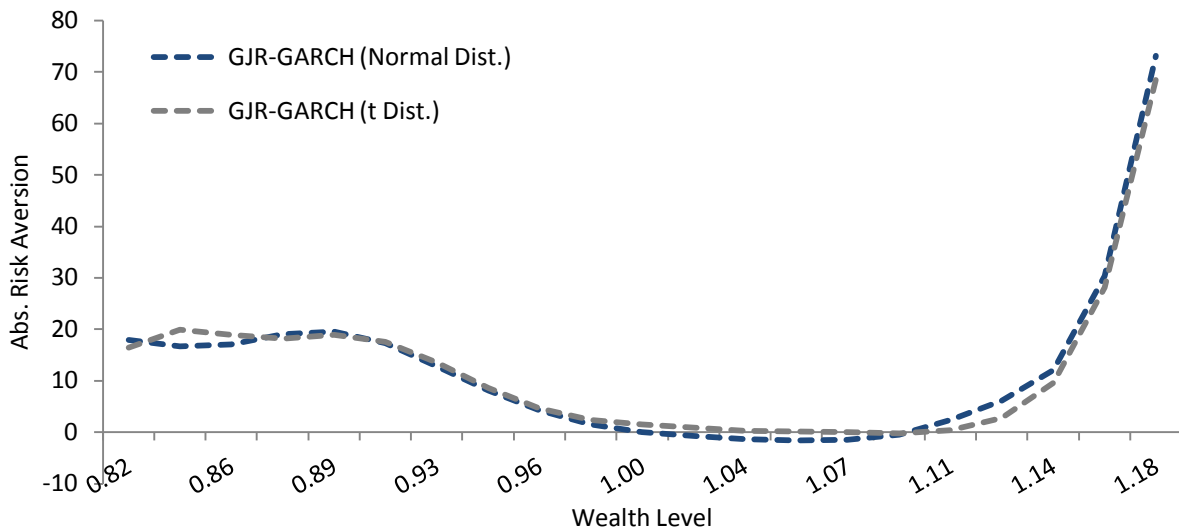


**Figure 13 Implied Risk Aversion Functions across Wealth**
The RNDs are estimated by the double lognormal method and the fast and stable method respectively. The subjective densities are generated by the GJR-GARCH model. The parameters of the GJR-GARCH model are estimated by the maximum likelihood method (MLM). When implementing the MLM, the normal distribution and the t-distribution are used.

Figure 12 displays the subjective densities generated by the GJR-GARCH model. Subjective densities derived from GJR-GARCH model are more leptokurtic than the density estimated by GARCH(1, 1) model. The implied risk aversion functions are U-shaped, and the risk aversion

increases sharply when the wealth level is above 1.14. Even if we change the model to generate the subjective density, it seems that the risk aversion function is still U-shaped.
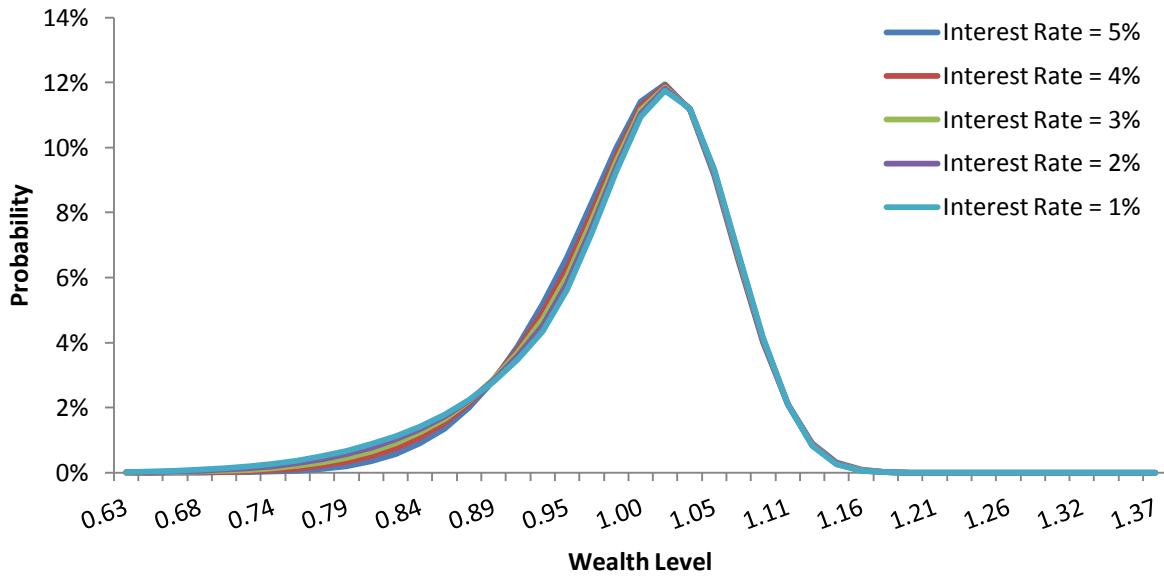


**Figure 14 RNDs Estimated by the Fast and Stable Method with Different Interest Rates**
The RNDs are estimated by the fast and stable method. The interest rates used vary from 1% to 5%.
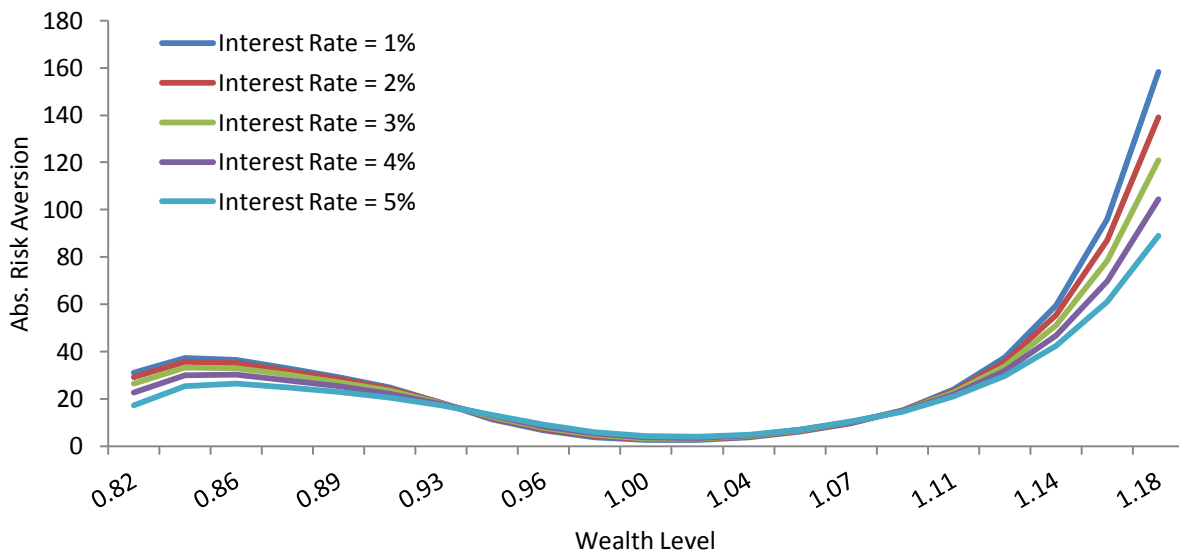


**Figure 15 Implied Risk Aversion Functions Across Wealth**
The implied risk aversion is estimated by RNDs and subjective density. RNDs are estimated by the fast and stable method with interest rates ranging from 1% to 5%, and the subjective density is estimated by the GARCH Monte Carlo Method.

The next thing is to estimate the shape of RNDs using different interest rates. Figure 14 demonstrates RNDs estimated by the fast and stable method using various interest rates. It seems that when different interest rates are employed, the shapes of the RNDs do not change significantly, except that the left tails of the distributions are fatter when interest rate decreases. The risk aversion functions in figure 15 preserve the features of those in figure 11. That is, the risk aversion is an increasing function of wealth when the wealth level is above a certain level. We also use the double lognormal method for this sensitivity analysis, and the unreported results are consistent with those obtained here. As indicated by Bliss and Panigirtzoglou (2004), interest rate does not play an important role in RND estimation.

Then I select a different set of trade-off parameters when using the fast and stable method. If the trade-off parameter gets larger, the implied volatility curves fit the observations better. Figure 16 demonstrates five fitted implied volatility curves that employ different trade-off parameters. It is interesting to point out that the fitted curves converge to straight lines when the trade-off parameter is getting smaller. Since we use the Black-Scholes formula as a transformation mechanism, the associated risk-neutral distribution will be closer to lognormal. A lognormal risk-neutral distribution will then result in a decreasing absolute risk aversion function. In figure 16, 17 and 18, we can see that if the volatility curve is close to a straight line, the risk-neutral density is close to lognormal, and the associated risk aversion decreases when the wealth level exceeds 0.93.

The reason is as follows. In the Black-Scholes economy, the asset price follows the geometric Brownian motion, and the implied volatility is constant. We obtain the risk-neutral density by taking the second-order derivative of the option function $C(S_t, T, K, \sigma, r)$. The subjective density can be generated directly using the physical asset price dynamics. Using equation (2.5) in section 2.3, we calculate the relative risk aversion function in the Black-Scholes economy. It turns out that the relative risk aversion is a constant across wealth, i.e. the absolute risk aversion is a decreasing function of wealth. See He and Leland (1993) and Ait-Sahalia and Lo (2000) for a more detailed derivation.

However, when the trade-off parameter is equivalent to or smaller than $10^3$, the fitter volatility curves no longer reflect investors' beliefs since the curves do not fit the empirical implied volatilities at all. Therefore, it is meaningless to interpret the associated risk aversion functions.

On the other side, the well-fitted implied volatility curves still lead to U-shaped risk aversion functions.
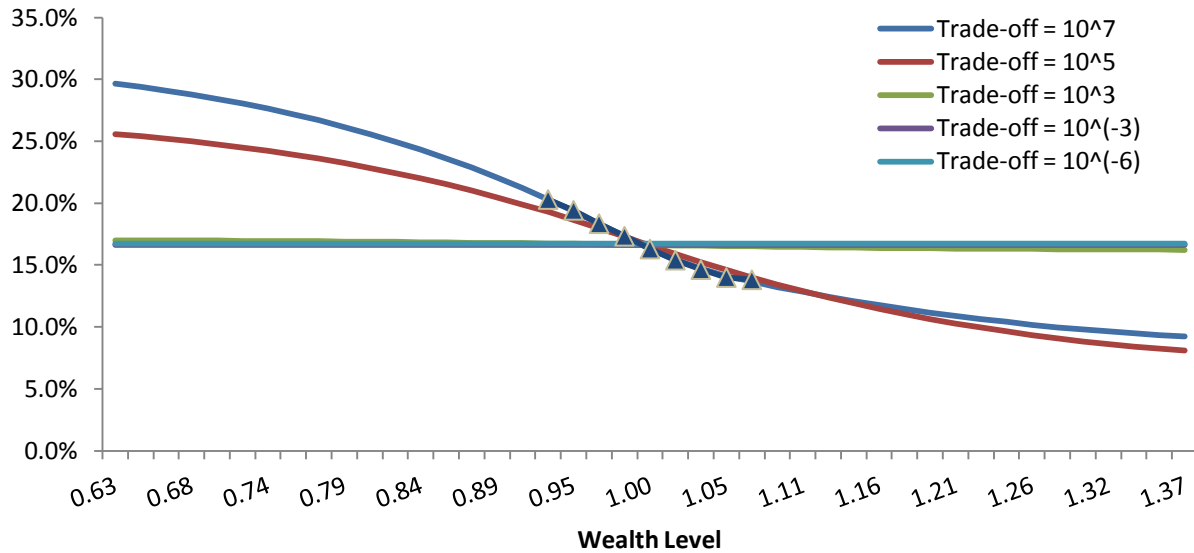


**Figure 16 Fitted Implied Volatility Curves**
The implied volatility curves are estimated by Jackwerth's (2000) fast and stable method using 9 options on 16.Apr.2012. Five different trade-off parameters are selected. Implied volatilities backed out from actual option prices are denoted by nine blue triangles.
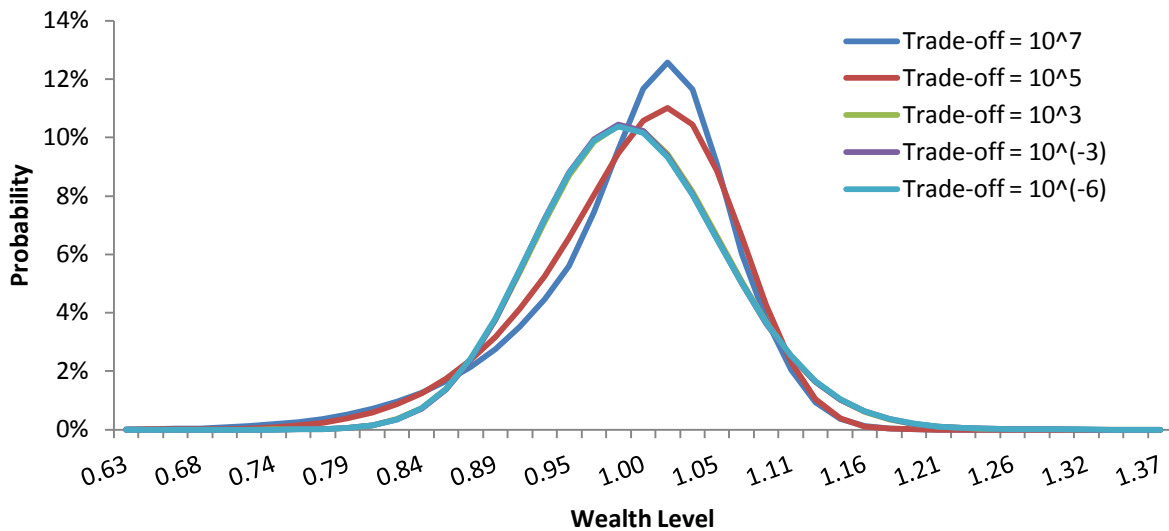


**Figure 17 RNDs Estimated by Fast and Stable Method**
The RNDs are estimated by the fast and stable method using 9 options on 16.Apr.2012. Five different trade-off parameters are used.
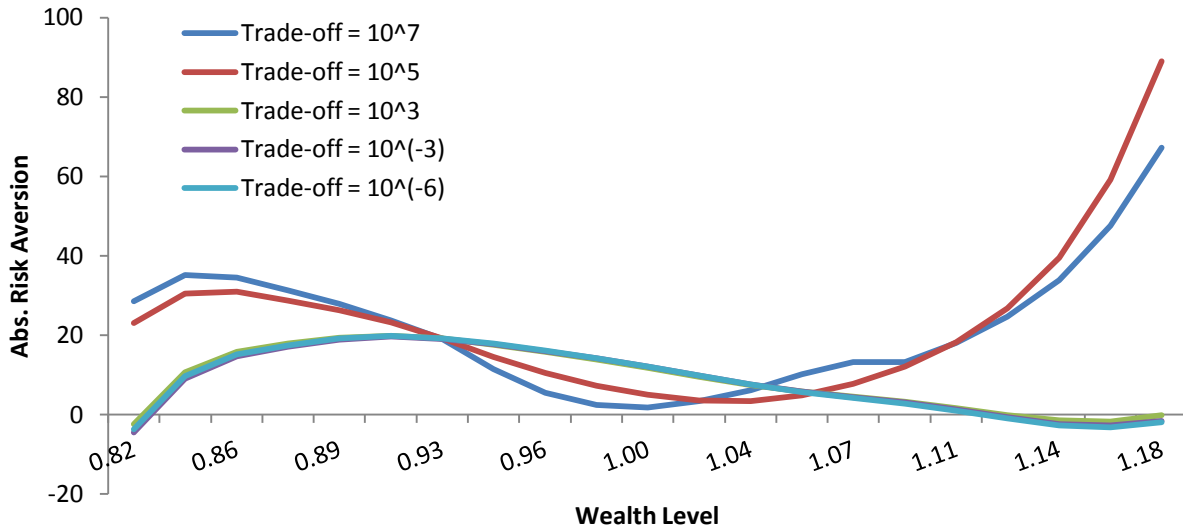
**Figure 18 Implied Risk Aversion Functions across Wealth**
The implied risk aversion functions are estimated by RNDs and subjective densities. RNDs are estimated by the fast and stable method with 9 options on 16.Apr.2012, and the subjective density is estimated by the GARCH Monte Carlo Method.

So far all the densities are derived using the option prices on 16.Apr.2012. It is possible that the implied risk aversion is unstable across different days. Therefore, we derive RNDs and subjective densities on a set of trading days, ranging from 4.Jan.2012 to 30.Apr.2012. A total number of 82 RNDs and subjective densities are extracted respectively. RNDs are estimated by Jackwerth's fast and stable method, and subjective densities are derived by the GARCH Monte Carlo method. Subseque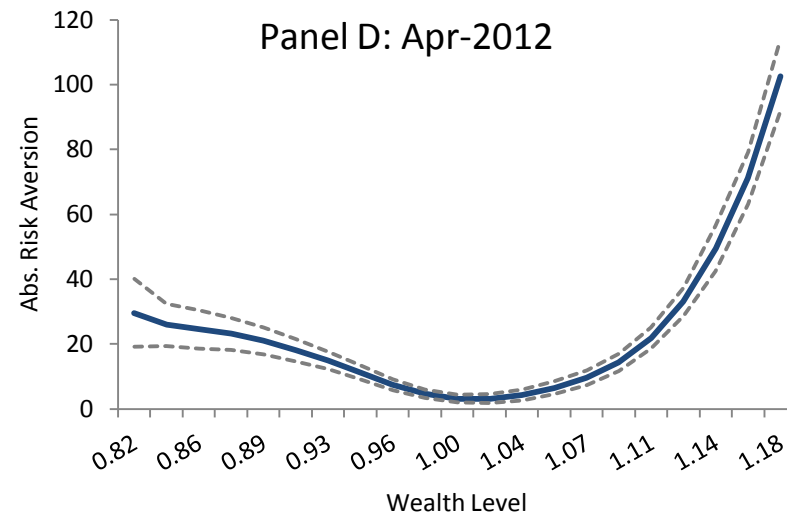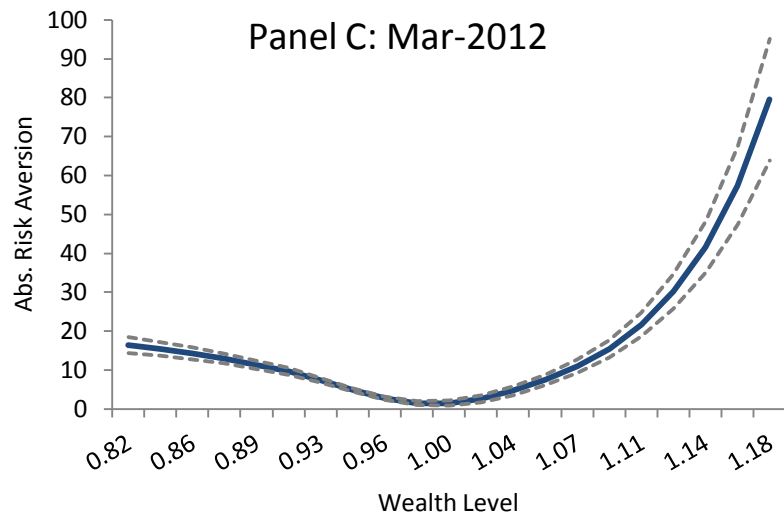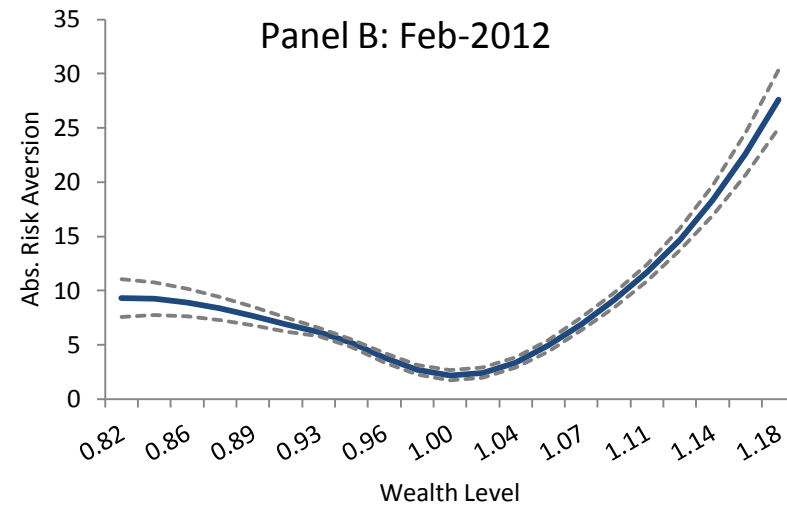ntly, 82 implied risk aversions are calculated with respective to each trading day. The figures are given in Appendix C. For each month, we compute the mean and standard deviation of risk aversions at each wealth level. Then we plot the average risk aversion across wealth for each month. It turns out that the absolute risk aversion functions are still humped for each month. Although the risk aversion is about 1~2 around the wealth level 1.00 in each month, it increases to approximately 18, 28, 80 and 110 at the wealth level 1.18 for January, February, March and April respectively. Therefore, the absolute risk aversion is time-varying. This is consistent with the finding of Rosenberg and Engle (2002). They study the relation between empirical risk aversion and the business cycle, and find that the empirical risk aversion is counter cyclical. Due to data limitation, I do not estimate the risk aversion over a longer period of time.
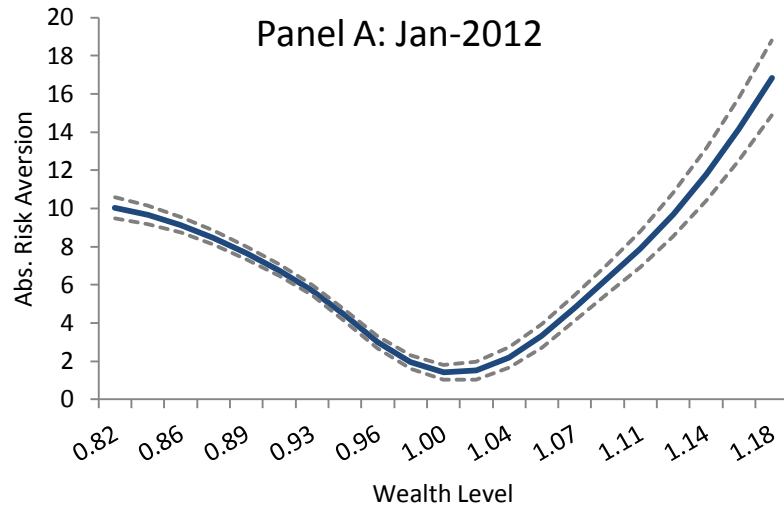
**Figure 19 Implied Risk Aversion Function across Wealth (Jan-2012, Feb-2012, Mar-2012 and Apr-2012)**
We calculate the absolute risk aversion functions across level from 4.Jan.2012 to 30.Apr.2012. For each month, we calculate the mean of risk aversion at each wealth level. The dashed line is the 0.5 times the standard of the risk aversion at each wealth level.

## 4.4 U-Shaped Implied Risk Aversion: Possible Explanations

The U-shaped implied risk aversion function is stable under different conditions. The shape of the implied risk aversion function changes only when we select a trade-off parameter that seriously deteriorates investors' beliefs. Furthermore, since options with extremely low and high strike prices are illiquid, they might not fully reflect investors' future expectations. So the tails of the RNDs might be inaccurately estimated. However, the risk aversion functions are U-shaped even around the wealth level 1, where options are actively traded. Illiquid trading is therefore not an excuse for the U-shaped risk aversion function. It seems that the U-shaped implied risk aversion might not be caused by empirical issues, i.e. data selection, methods used etc.

Theoretically, several issues might influence the validity and accuracy of the estimated risk aversion function. The equilibrium model discussed in section 2.3, so to speak, is based on some strict assumptions: in the exchange economy, there is a single consumption good, the market is a frictionless and complete, and the stock index is a proxy of aggregate endowment and aggregate wealth. However, in reality, there are obviously more consumption goods and the market is not frictionless. If the market is incomplete, we are unable to use the representative investor framework, and the RND function is no longer unique, so is the risk aversion function. Another controversial assumption is that in equilibrium the representative investor optimally holds the market portfolio so that the stock index is a proxy of his aggregate wealth and aggregate consumption. However, Lochstoer (2006) shows that the stock index returns are heteroskedastic while that of the aggregate consumption is not. Ziegler (2007) exploits the potential solutions to the implied risk aversion smile puzzle and studies the shape of the risk aversion function under different frameworks, i.e. agents with heterogeneous preferences and beliefs, asset price dynamics with stochastic volatility or jumps. He concludes that none of these can account for the risk aversion smile.

# 5. Concluding Remarks

In this paper, we derive the implied risk aversion function under a specific economy, where the market is complete so that there is a representative investor who endows with one unit of risky asset and optimally follows the path-independent strategy [Ait-Sahalia and Lo (2000), Jackwerth (2000)]. In this economy, the absolute risk aversion can be written as a function of the subjective density and the risk-neutral density.

Then we use the GARCH Monte Carlo method to estimate the subjective density, and use the double lognormal method as well as the fast and stable method to estimate the risk-neutral density. Compared with the subjective densities, the risk-neutral densities are leptokurtic with fatter left tails. Although the RNDs estimated by the two methods are in similar shape, the results of the Kolmogorov-Smirnov test indicate that the RNDs estimated by these two methods are statistically different with a p-value of 0.000.

Next, the implied risk aversion function is calculated. However, we find that the absolute risk aversion function is U-shaped. It is inconsistent with the theory of finance. Intuitively, when an agent's wealth increases, his risk aversion decreases. The U-shaped risk aversion function has also documented in other related papers, such as Jackwerth (2000, 2004), Ait-Sahalia and Lo (2000), Rosenberg and Engle (2002), and Ziegler (2007), to name a few. We implement some robustness tests to examine the stability of the U-shaped risk aversion function by comparative static analysis. The subjective densities are estimated using a different GARCH-family model, and the RNDs are estimated with options on various trading days. The results indicate that the U-shaped risk aversion function is robust under different empirical settings. Therefore, it seems that the U-shaped function is less likely to be caused by the empirical issues. As stressed by Ziegler (2007), the standardized consumption-based model is too idealistic. Future research can study the implied risk aversion function under a more sophisticated economy, which for example, admits market incompleteness.

# Appendix A: Asset Pricing Theory

Asset pricing can be categorized by equilibrium pricing and arbitrage pricing. The former analyzes the demand and supply of assets while the latter is based on the ideal that assets with the same expected payoffs should have the same price. However, these two types of models are interrelated. In section 1 of this appendix, we will start with Pareto optima, and end with the multi-period dynamic equilibrium model. The connection between arbitrage pricing and equilibrium pricing is the Arrow-Debreu security. In section A2, the risk-neutral pricing and the martingale measure will be discussed. The whole framework is the foundation of modern finance theory. In order to better understand the RND estimation methods and implied risk aversion, it is necessary to study the theory behind.

## A1 Equilibrium Pricing

Equilibrium models analyze the supply and demand of assets. From the first fundamental theorem of welfare economics, the competitive economy market equilibrium must be Pareto optimum, that is, resources reallocation cannot improve one agent's utility without deteriorate other agents' utilities. I will start with the simplest but classic one-period equilibrium model, and will introduce time and uncertainty gradually. Note that in all the equilibrium models here technology and production are not considered.

### A1.1 Static Competitive Exchange Model

In a static model, all the transactions occur at one time point and there is no uncertainty. There are N agents and G tradable goods. Each individual initially owns tradable goods as endowment, and consumes all the goods after trading. Agents' preferences are represented by their utility function. They act to maximize their own utilities. Using mathematical notations, we have

    *(1) Agents $N = [1, \dots, n, \dots N]$;*

    *(2) Tradable goods $G = [1, \dots, g, \dots G]$;*

    *(3) Price vector of the goods $P = [P_1, \dots, P_g, \dots P_G]$;*

    *(4) Endowment for agent $n, \varphi^n = [\varphi_1^n, \dots, \varphi_g^n, \dots \varphi_G^n]$;*

(5) $Consumption\ for\ agent\ n,\ C^n = [C_1^n, \ldots, C_g^n, \ldots C_G^n];$

(6) $Utility\ (for\ agent\ n)\ U^n(C^n),\ where\ U(.) \in \mathbb{C}.\ If\ U(.)\ is\ continuous,\ concave\ and$
$strictly\ monotonic,\ then\ we\ say\ U(.) \in \mathbb{C}.$

An equilibrium allocation $C = [C^1, \ldots, C^n, \ldots C^N]$ and price vector $P = [P_1, \ldots, P_g, \ldots P_G]$ satisfy the following conditions

$$Max\ [U^n(C^n)]$$

$$s.t.\ P \cdot C^n = P \cdot \varphi^n$$

$$C^n \geq 0$$

If there is no other feasible allocation $\bar{C} = [\bar{C}^1, \ldots, \bar{C}^n, \ldots \bar{C}^N]$ that makes $U^n(\bar{C}^n) > U^n(C^n), \forall\ n \in N$, then the equilibrium allocation $C = [C^1, \ldots, C^n, \ldots C^N]$ is Pareto optimum.

This is the fundamental framework for the equilibrium analysis. The existence of the equilibrium in this economy is proved by Arrow and Debreu (1954), and further elaborated by Varian (1992) and Mas-Colell et al. (1995).

## A1.2 Two-period Competitive Exchange Model

In this framework, we extend the one-period model into a two-period model, and introduce uncertainty. There are two time points $t_0$ and $t_1$ when agents consume, but the endowments at period $t_1$ are uncertain, i.e. there are different states of world at $t_1$. The inputs of the model are as follow

(1) $States\ of\ world\ at\ time\ t_1, S = [1, \ldots, s, \ldots S];$

(2) $Agents, N = [1, \ldots, n, \ldots N];$

(3) $Tradable\ goods, G = [1, \ldots, g, \ldots G];$

(4) $Price\ vector\ of\ goods\ at\ time\ t_0, \boldsymbol{P_0} = [P_{01}, \ldots, P_{0g}, \ldots P_{0G}];$

(5) $Price\ vector\ of\ goods\ at\ time\ t_1\ in\ state\ s, \boldsymbol{P_s} = [P_{s1}, \ldots, P_{sg}, \ldots P_{sG}];$

(6) $Endowment\ for\ agent\ n\ at\ time\ t_0, \boldsymbol{\varphi_0^n} = [\varphi_{01}^n, \ldots, \varphi_{0g}^n, \ldots \varphi_{0G}^n];$

(7) $Endowment\ for\ agent\ n\ at\ time\ t_1\ in\ state\ s, \boldsymbol{\varphi_s^n} = [\varphi_{s1}^n, \ldots, \varphi_{sg}^n, \ldots \varphi_{sG}^n];$

(8) $Consumption\ for\ agent\ n\ at\ time\ t_0, \boldsymbol{C_0^n} = [C_{01}^n, \ldots, C_{0g}^n, \ldots C_{0G}^n];$

(9) $Consumption\ for\ agent\ n\ at\ time\ t_1\ in\ state\ s, \boldsymbol{C_s^n} = [C_{s1}^n, \ldots, C_{sg}^n, \ldots C_{sG}^n];$

(10) $Prices\ of\ contingent\ claims, \boldsymbol{\acute{P}_s} = [\acute{P}_{s1}, \ldots, \acute{P}_{sg}, \ldots \acute{P}_{sG}];$

(11) $Trading\ strategy\ of\ the\ agent, \boldsymbol{\theta_s^n} = [\theta_{s1}^n, \ldots \theta_{sg}^n, \ldots \theta_{sG}^n];$

(12) $Additive\ utility\ (for\ agent\ n)\ U^n(\boldsymbol{C^n}),\ where\ U(.) \in \mathbb{C}\ and\ satisfies\ the$

*von Neumann − Morgenstern (vNM) utility theorem*;

(13) *Agents are impatient so that utility at future period is discounted by a time consistent discount rate* $\delta^n$.

In this model, the utility function is additive. This type of utility function is frequently used and convenient for computation purpose. Yet it has been criticized for linking risk aversion coefficient to intertemporal substitution[10]. Since uncertainty is introduced, we assume that the utility at period $t_1$ is the expected utility across states, discounted by a discount rate $\delta$. The tradable goods are assumed to be perishable so that people cannot store the goods from $t_0$ to $t_1$. Here the contingent claim is defined as a commitment to pay one unit of commodity $g$ in state $s$ at time $t_1$. The price for this contingent claim is $\acute{P}_{sg}$. Contingent claim works as an important instrument for intertemporal asset allocation and risk hedging.

To demonstrate the importance of contingent claim, we first ignore it and inputs (10) and (11). The equilibrium in such case satisfies

$$Max \ \{U^n(\boldsymbol{C}_0^n) + \delta^n E[U^n(\boldsymbol{C}_s^n)]\}$$

$$s.t. \ \boldsymbol{P}_0 \cdot \boldsymbol{C}_0^n \leq \boldsymbol{P}_0 \cdot \boldsymbol{\varphi}_0^n$$

$$\boldsymbol{P}_s \cdot \boldsymbol{C}_s^n \geq \boldsymbol{P}_s \cdot \boldsymbol{\varphi}_s^n, \forall s \in S$$

$$\boldsymbol{C}^n \geq \boldsymbol{0}$$

Due to the perishability of the goods, agents are not able to make intertemporal decisions. In other words, one cannot store his goods to smooth his marginal utility. Therefore, there is no connection between either $t_0$ and $t_1$ or state $s_a$ and $s_b$. The equilibria at time $t_0$ are in fact the same as those discussed in the static model. At $t_1$, every individual's consumption decision is bounded by his own endowment at that specific state. He cannot hedge his risk in those states where he has fewer endowments. However, from a welfare point of view, it is possible for two or more agents to achieve higher utility by trading. For example, agents with high endowment at time $t_0$ but low endowment at $t_1$ can trade with agents who have high endowment at time $t_1$ but nothing at $t_0$; on the other hand, agents with high endowment at state $s_1$ can make an agreement with those with high endowment at state $s_2$ to alleviate the consumption volatility among states. As agents are assumed to be risk averse, such trades can actually increase the utility of both traders.

---

[10] For example, in the approximated log linearized Euler equation with a constant relative risk aversion (CRRA) utility function, the risk aversion coefficient is exactly the intertemporal substitution parameter.

Contingent claim market works as an instrument for intertemporal asset allocation and risk hedging. Hence, investors' consumption restriction caused by time and uncertainty can be solved. Suppose (10) and (11) are now included, and the equilibrium should satisfies

$$Max \ \{U^n(\boldsymbol{C_0^n}) + \delta^n E[U^n(\boldsymbol{C_s^n})]\}$$

$$s.t. \ \boldsymbol{P_0} \cdot \boldsymbol{C_0^n} + \sum_{s=1}^{S} \acute{\boldsymbol{P}}_s \cdot \boldsymbol{C_s^n} \leq \boldsymbol{P_0} \cdot \boldsymbol{\varphi_0^n} + \sum_{s=1}^{S} \acute{\boldsymbol{P}}_s \cdot \boldsymbol{\varphi_s^n}$$

$$\boldsymbol{C_0^n}, \boldsymbol{C_s^n} \geq \boldsymbol{0}$$

The inequality can be rewritten as $\boldsymbol{P_0} \cdot \boldsymbol{C_0^n} \leq \boldsymbol{P_0} \cdot \boldsymbol{\varphi_0^n} - \sum_{s=1}^{S} \acute{\boldsymbol{P}}_s \cdot \boldsymbol{\theta_s^n}$ since $\boldsymbol{\theta_s^n} + \boldsymbol{\varphi_s^n} = \boldsymbol{C_s^n}$. Market clearing condition is fulfilled when $\sum_{n=1}^{N} \boldsymbol{C_{sg}^n} = \sum_{n=1}^{N} \boldsymbol{\varphi_{sg}^n}$, $\forall s \in S \ \& \ g \in G$. This equilibrium is also Pareto optimum. This is the very classic Arrow-Debreu economy by Arrow (1964) and Debreu (1959). Mas-Colell et al. (1995) also elaborate the construction of the economy and the existence of equilibria. A more intuitive but less technical interpretation is given by Danthine and Donaldson (2005). Arrow-Debreu economy brings us the concept of pure security, which plays a significant role in arbitrage pricing. Pure security builds a bridge between stochastic discount factor (SDF) and risk-neutral probability, and closely relates to market completeness. These will be demonstrated in section A2.2.

## A1.3 Multi-period Competitive Exchange Model

In the two-period setting, all the decisions about contingent claims are made at $t_0$. If we extent to multi-period framework, agents can make investment and consumption decision at every time point. Before explaining the model, a complete probability space has to be setup. A complete probability space $(\Omega, P, \mathcal{F})$ consists

    (1) *A common sample belief $\Omega = (\omega_1, \dots, \omega_s, \dots, \omega_S)$ with S events*;

    (2) *The probability measure P corresponds to each event*;

    (3) *The $\sigma - algebra$[11] $\mathcal{F}$ which is a collection of events*.

The event $\omega_s \in \mathcal{F}$ is measureable if it is able to project to the real number set, i.e. $F: \Omega \to \mathcal{R}$. Then we define $\{\mathcal{F}_t\}_{t \in [0,T]}$ as an information flow which contains all the information at time $t$ and before time $t$,

---

[11] The formal definition of the $\sigma - algebra$ on $\Omega$ is:
    (1) The empty set $\emptyset \in \mathcal{F}$;
    (2) If $\omega_s \in \mathcal{F}$, then its complement $\omega_s^C \in \mathcal{F}$;
    (3) If $\omega_1, \omega_2, \dots, \omega_n \in \mathcal{F}$, then $\bigcup_{i=1}^{n} \omega_i \in \mathcal{F}$.

i.e. $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \cdots \subseteq \mathcal{F}_t$. If the consumption stochastic process $C_t(\omega)$ or the stochastic endowment process $\varphi_t(\omega)$ is $\mathcal{F}_t$ measurable, then we say that they are $\mathcal{F}_t$-adapted. The trading strategy $\boldsymbol{\theta_{t+1}}$ shown in the following equation is also $\mathcal{F}_t$-adapted.

In a dynamic equilibrium model, all the agents should have the common information structure and all the consumption, endowment and price process are $\mathcal{F}_t$-adapted. Each individual makes investment and consumption decisions according to the information available at that specific time. The setting of the dynamic model is similar to that in section A1.2 except that there is now only one good and the agents are identical. Notations $\boldsymbol{\varphi}$ and $\boldsymbol{C}$ are now in monetary forms. The consumption for each period equals the endowment minus the net investment in that period, i.e. $\boldsymbol{C_t^n = \varphi_t^n - \acute{P}_t[\theta_{t+1}^n - \theta_t^n]}$. Thus, agents will maximize their lifetime utility subject to the budget constraint

$$Max \, \mathrm{E}\left[\sum_{t=0}^{T}(\delta^n)^t \, U(\boldsymbol{C_t^n})\right]$$

$$s.t. \, \boldsymbol{C_t^n = \varphi_t^n - \acute{P}_t[\theta_{t+1}^n - \theta_t^n]}$$

A market clearing allocation $\boldsymbol{C = [C^1, \dots, C^n, \dots C^N]}$ is Pareto optimum when no other feasible allocation $\boldsymbol{\overline{C} = [\overline{C}^1, \dots, \overline{C}^n, \dots \overline{C}^N]}$ can lead to $\boldsymbol{U^n(\overline{C}^n) > U^n(C^n)}$.

This model can further be generalized if we consider K securities instead of the contingent claim which pays one unit of goods at a specific state. Then the maximization problem can be written as

$$Max \, \mathrm{E}\left[\sum_{t=0}^{T}(\delta^n)^t \, U(\boldsymbol{C_t^n})\right]$$

$$s.t. \, \boldsymbol{W_{t+1}^n = (W_t^n - C_t^n + Y_t^n)\sum_{i=1}^{K} \theta_i^n R_i}$$

This is the modern way of deriving equilibrium asset pricing models. Agents' wealth at each period is the total portfolio return of last period's wealth plus his endowment minus the consumption at that period. Merton (1969) first solves a similar portfolio selection problem. Further discussion includes Pennacchi (2007) and Back (2010). Materials about measure theory with application in finance can be found in Øksendal (2010).

## A1.4 Stochastic Discount Factor

Stochastic discount factor (SDF), or the pricing kernel, is one of the most important concepts in asset pricing. The SDF can be deduced from a standard dynamic exchange economy, with one consumption good and no exogenous income. These assumptions are made to reduce computation[12]. Moreover, financial economists concentrate more on trading strategy and asset pricing, rather than how income process influences consumption.

Here the model is restricted to two periods, but the two-period model can be extended to multi-period model without changing the fundamental setting. Also, the endowment is ignored. Without considering endowments, agent's consumption depends purely on his initial wealth and investment decisions. Then we introduce a representative investor who is treated as a weighted average of all individuals in the market. This representative investor's consumption at period $t$ is the aggregate consumption of at that period and his preference is a weighted average of all the agents' preference. This is been proved in the first part of Constantinides's (1982) paper. The model in this setting is

$$Max \, U(C_0) + \mathrm{E}_0 \left[ \delta \, U(C_1) \right]$$

$$s.t. \; C_1 = (W_0 - C_0) \sum_{i=1}^{K} \theta_i R_{i1}$$

$$\sum_{i=1}^{K} \theta_i = 1$$

The representative investor maximizes his overall discounted expected utility subject to his consumption budget. He has initial wealth of $W_0$ and invests the difference between initial wealth and consumption $W_0 - C_0$ in the capital or money market. At period $t_1$ he will spend all the wealth for consumption, and the consumption for period $t_1$ depends how much the investor save during the first period, the trading strategy, and expected return of assets. By substituting the first budget constraint into the second term of the maximization problem, we can obtain the first-order conditions (a1) and (a2)

$$\mathrm{E}_0 \left[ \delta \, U'(C_1) R_{i1} \right] = \lambda, \; \forall i \in K \quad (a1)$$

$$U'(C_0) = \lambda \quad (a2)$$

---

[12] For example, multi-period optimization problem with stochastic income is difficult to solve since the optimal consumption with respect to income at each period might not have a closed-form solution. The consumption function has to be approximated by mathematical software like Mathematica or Matlab.

From the F.O.C. above, we know that

$$U'(C_0) = E_0 [\delta\, U'(C_1) R_{i1}], \ \forall i \in K \quad (a3)$$

$$E_0 [\delta\, U'(C_1) R_{i1}] = E_0 [\delta\, U'(C_1) R_{j1}], \ \ \forall i,j \in K \quad (a4)$$

(a3) and (a4) are consistent with the arguments in section A1.2 regarding the two-period equilibrium model. When agents can make intertemporal consumption and investment decision, they will eliminate consumption uncertainty by equating the expected marginal utility between two periods weighted by asset returns. Also, to hedge the risks among different states, investors will smooth the expected marginal utility weighted by different asset returns. Asset return can be defined as

$$R_{it+1} = \frac{X_{it+1}}{P_{it}} \quad (a5)$$

For assets like contingent claims that expire at period $t + 1$, the return is the payoff at $t + 1$ divided by $P_{it}$. For assets like stocks, the return is the summation of dividend and price at $t + 1$ divided by stock price at t. Substituting (a5) into the first equation

$$U'(C_0) = E_0 \left[\delta\, U'(C_1) \frac{X_{i1}}{P_{i0}}\right]$$

As $P_{i0}$ is measurable under period 1, we can reorganize the equation by taking $P_{i0}$ out of the expectation sign

$$P_{i0} = E_0 \left[X_{i1} \frac{\delta\, U'(C_1)}{U'(C_0)}\right] = E_0 [X_{i1} M_{01}] \quad (a6)$$

Equation (a6) indicates that we can price any assets with the price kernel and next period payoffs. The intuition is that investors are willing to pay greater amount for assets that have high payoff, or when their future consumption level is low or when their current consumption level is high. On the one hand, if their future consumption at one specific future state is low, the representative agent is willing to hedge the risk and pay more for an asset that have higher payoff in that state. On the other hand, if the agent have high consumption at period $t_0$, his is willing to smooth his marginal utility intertemporally and is willing to pay a higher price for assets that have payoff in the next period $t_1$.

Materials about stochastic discount factor are covered by Back (2010) and Pennacchi (2007). The construction of a representative investor is discussed by Danthine and Donaldson (2005), Back (2010) and Constantinides (1982).

# A2 Arbitrage Pricing

Arbitrage pricing models are straight forward and receive more positive results in empirical studies. The only underlying assumption of arbitrage pricing is that assets or portfolio of assets should have the same price if their future payoffs are identical across different states. An alternative statement of no arbitrage condition[13] is that it is not possible to earn positive expected future payoffs with an asset or portfolio that has zero or negative price.

A good starting point to discuss arbitrage pricing is the Arrow-Debreu security. In section A1.2, we defined a contingent claim as the right to receive one unit of commodity $m$ at the state $s$. However, then a total number of $M \times S$ kinds of contingent claims are required for the consumers to fully hedge the risks through all states. Instead, if we use a security that pays one unit of currency when state $s$ occurs, then only $S$ securities are need. This security is called the pure security or the Arrow-Debreu security.

## A2.1 Market Completeness

Ross (1976) first illustrates the ideal of market completeness. If the number of pure security $N$ is equivalent to that of the states $S$, then the market is complete. If, however, $N < S$, then the market is incomplete. Using matrix notation, the payoffs of $N$ securities in $S$ states are

$$\begin{bmatrix} X_{11} & \cdots & X_{1S} \\ \vdots & \ddots & \vdots \\ X_{N1} & \cdots & X_{NS} \end{bmatrix}$$

In a complete market, there are $S$ pure securities, and their payoffs are

$$S_1 = \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T, \dots, S_s = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}^T, \dots, S_S = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix}^T$$

Then the payoff of any asset can be replicated by a combination of pure securities

$$P_i = X_{i1} \begin{bmatrix} 1 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix}^T + \cdots + X_{is} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}^T + \cdots + X_{iS} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 1 \end{bmatrix}^T = S_1 X_{i1} + \cdots + S_s X_{is} + \cdots + S_S X_{iS}, \forall i \in N \quad (a7)$$

---

[13] This is called the "no free lunch" condition.

Under no arbitrage condition, assets or portfolio of assets with the same expected payoffs should have the same price, and the price of any asset can be represented by equation (a7) above. If the states are continuous rather than discrete, the price of an asset equals

$$P = \int_0^\infty f_S(S_T) \, X(S_T) \, dS_T \quad (a8)$$

where $f_S(S_T)$ is the state price density and $X(S_T)$ is the end-of-period payoff that depends on the final asset price $S_T$.

In the real world, there are no pure securities traded in the market[14]. However, we can create the system of pure securities payoff matrix by selling and buying the existing assets in the market. The selling and buying process can be treated as the elementary transformation in linear algebra. After implementing the elementary transformation, the payoffs matrix should become an identity matrix. If the rank of the identity matrix is larger than or equal to the states of world, then we say the market is complete.

$$\begin{bmatrix} X_{11} & \cdots & X_{1S} \\ \vdots & \ddots & \vdots \\ X_{N1} & \cdots & X_{NS} \end{bmatrix} \xrightarrow[\text{Transformation}]{\text{Elementary}} \begin{bmatrix} 1_{11} & \cdots & 0_{1S} \\ \vdots & \ddots & \vdots \\ 0_{N1} & \cdots & 1_{NS} \end{bmatrix}$$

In a complete market, any assets can be replicated by existing assets in the market. Thus, any arbitrage profits can be captured by selling the one with high price and buying the other with low price. So any markets that have arbitrage opportunities are not in equilibrium. In an incomplete market, the rank of the payoff matrix is smaller than the number of states. Then pure securities for some specific states are not replicable, so agents cannot hedge their risks from some states.

## A2.2 Risk-Neutral Pricing

The ideal of risk-neutral pricing is proposed by Ross (1976) and Cox and Ross (1976). Risk-neutral probability and the state price are closely related. This will be demonstrated in the following paragraphs. Using pure security, we can calculate the price of different assets once their payoffs are known. Suppose there are two states, and two securities. Their payoffs at date 1 for two different states are as follow

$$X_1 = \begin{bmatrix} B_1 & B_1 \\ X_{1U} & X_{1D} \end{bmatrix}$$

---

[14] We can, however, create a portfolio of options that approximates the payoff of a pure security manually using a butterfly spread.

If the prices of pure securities are given, then prices of the risky and riskless securities equal the product of payoffs and the prices of pure securities. The payoff of the riskless security can be replicated by buying $B_1$ units of pure securities for both state, and the payoff of the risky asset can be replicated by buying $X_{1U}$ units of pure security for state $U$ and buying $X_{1D}$ units of pure security for state $D$.

$$\begin{bmatrix} P_{B,0} \\ P_{S,0} \end{bmatrix} = \begin{bmatrix} B_1 & B_1 \\ X_{1U} & X_{1D} \end{bmatrix} \begin{bmatrix} S_{U,0} \\ S_{D,0} \end{bmatrix} = \begin{bmatrix} (S_{U,0} + S_{D,0})B_1 \\ X_{1U}S_{U,0} + X_{1D}S_{D,0} \end{bmatrix}$$

This problem can be illustrated in terms of return. The payoff of the riskless asset is $P_{B,0}(1+r)$, where $r$ is the risk-free rate. Then the equation above can be rewritten as

$$\begin{bmatrix} P_{B,0} \\ P_{S,0} \end{bmatrix} = \begin{bmatrix} P_{B,0}(1+r) & P_{B,0}(1+r) \\ X_{1U} & X_{1D} \end{bmatrix} \begin{bmatrix} S_{U,0} \\ S_{D,0} \end{bmatrix} = \begin{bmatrix} (S_{U,0} + S_{D,0})P_{B,0}(1+r) \\ X_{1U}S_{U,0} + X_{1D}S_{D,0} \end{bmatrix}$$

By reorganizing the equation, we have

$$\left( S_{U,0} + S_{D,0} \right)(1+r) = 1$$

$$P_{S,0} = X_{1U}S_{U,0} + X_{1D}S_{D,0}$$

Define $S_{U,0}(1+r) = Q_U$ and $S_{D,0}(1+r) = Q_D$ such that $Q_U + Q_D = 1$. We can infer that

$$P_{S,0} = \frac{1}{1+r}\left[ (1+r)X_{1U}S_{U,0} + (1+r)X_{1D}S_{D,0} \right] = \frac{1}{1+r}\left[ Q_U S_{1U} + Q_D S_{1D} \right]$$

This formula means that the price of an asset is equivalent to the expected future payoffs discounted by risk-free rate, if we treat the $Q_U$ and $Q_D$ as probabilities. This is feasible since $Q_U + Q_D = 1$ and both $Q_U$ and $Q_D$ are positive [15]. In microeconomics, risk-neutral investors are those who require no additional compensation for risk since they only concern about expected payoff. Thus the probabilities are called risk-neutral probabilities. Generally, when the states of world are continuously, the formula above can be written as

$$P = e^{-rT} \int_0^\infty f_Q(S_T) X(S_T) \, dS_T \quad (a9)$$

Comparing (a8) and (a9), we find that the risk-neutral density is simply the state price density times the continuous-time risk-free return, i.e. $f_Q(S_T) = e^{rT} f_S(S_T)$.

---

[15] $P_{U,0}(1+r) = Q_U$ and $P_{D,0}(1+r) = Q_D$, and riskless rate is always positive. In addition, when there is no arbitrate opportunity, the price of pure security will always be positive.

## A2.3 Equivalent Martingale Measure

Martingale method is a revolution in quantitative finance. Using martingale method for asset pricing is very convenient. For example, using martingale method avoids solving the partial differential equation in the Black-Scholes model. The definition of martingale[16] $X_t$ is

(1) In a probability space $(\Omega, P, \mathcal{F})$ with the information flow $\{\mathcal{F}_t\}_{t\in[0,T]}$, the stochastic process $X_t$ is $\mathcal{F}$-adapted;

(2) $E[|X_t|] < \infty$;

(3) $E[X_{t+n}|\mathcal{F}_t] = X_t, \forall n \geq 0$.

The definition of sub-martingale[17] is the same as that of martingale except that condition (3) becomes $E[X_{t+n}|\mathcal{F}_t] > X_t$.

In financial markets, most normalized assets prices[18] are sub-martingales rather than martingales because most price series have an upward growing trend. Only normalized riskless asset prices are martingales. Therefore, martingale method is developed to change the measure, and thus transforms normalized prices that were sub-martingales to martingales. This can be done by Girsanov transformation. If there is a probability measure $Q$ such that normalized prices are martingales under measure $Q$, then the measure $Q$ is called equivalent martingale measure[19]. As a matter of fact, the equivalent martingale measure is the risk-neutral measure, under which the asset prices equal the expected future payoff discounted by risk-free rate. In other words, the normalized asset prices under risk-neutral measure are martingales.

Using equivalent martingale measure (risk-neutral measure), I will introduce important theorems in arbitrage pricing in the following section. The equivalent martingale measure and its application are illustrated by Harrison and Kreps (1979).

---

[16] For example, $Y_t = Y_{t-1} + \varepsilon_t$ and $dY_t = \sigma(t,\omega)dB_t$ are martingales in discrete time and continuous time respectively, where $\varepsilon_t$ is the white noise term, $\sigma(t,\omega)$ is a stochastic process and $B_t$ is the Brownian motion. The definition of Brownian motion can be found in Appendix B.

[17] Examples of sub-martingales in discrete and continuous time are $Y_t = \alpha + Y_{t-1} + \varepsilon_t$ and $dY_t = \alpha dt + \sigma(t,\omega)dB_t$, where $\alpha$ is a positive constant term.

[18] Normalized asset prices equal its prices divided by risk-free asset prices. For example, suppose the risk-free rate equals $r$, and the riskless asset price at time $t_0$ is 1. Then its prices from period $t_1$ to $t_n$ are $r + 1, (r + 1)^2, \ldots, (r + 1)^n$, respectively. Then the normalized price for this riskless asset is always 1 in each period. However, for risky assets, the expected returns contain risk premiums that compensate for risk bearing, and the expected return should be $r + \psi$. For an risky asset with initial price 1, the normalized prices from period $t_1$ to $t_n$ are $r + \psi + 1/r + 1, (r + \psi + 1)^2/(r + 1)^2, \ldots, (r + \psi + 1)^n/(r + 1)^n$. This is a sub-martingale since $r + \psi > r$.

[19] Measure $Q$ is equivalent to measure $P$ in the sense that $\forall \omega \in \Omega, P(\omega) = 0 \Leftrightarrow Q(\omega) = 0$.

## A2.4 Fundamental Theorem of Asset Pricing

The first fundamental theorem of asset pricing states that a market with equivalent martingale measure does not admit arbitrage opportunity. In the two-period setting in section A2.1, without the no arbitrage condition, the state prices and risk-neutral probabilities cannot be extracted from the market. This can also be proved from the perspective of equivalent martingale measure under continuous time.

Firstly, we define the price process of an asset or a portfolio as $V_t$. If $V_0 = 0$, an arbitrage satisfies $P(V_T \geq 0) = 1, P(V_T > 0) > 0$. The former equation means that the value of the portfolio will not be negative at $T$, and the latter one indicates that it is possible that the value of the portfolio is greater than zero at time $T$. Secondly, we prove by apagoge. If there is an equivalent martingale measure $Q$ such that the normalized price process under $(\Omega, Q, \mathcal{F})$ is a martingale, then $E^Q[e^{-rt}V_T|\mathcal{F}_0] = 0$. Since measure $Q$ is an equivalent measure regarding $P$, $P(V_T \geq 0) = 1$ indicates that $Q(V_T \geq 0) = 1$. Suppose $Q(V_T > 0) > 0$, then the martingale property $E^Q[e^{-rt}V_T|\mathcal{F}_0] > 0$. This is a contradiction. Therefore, the existence of equivalent martingale measure means that there is no arbitrage opportunity.

The second fundamental theorem of asset pricing indicates that there exists a unique risk-neutral measure if and only if the market is complete. In section A2.1, we know that the market is complete if the rank of the payoff matrix equals the number of states. Given the asset prices and their payoff matrix, the state price $S$ can be calculated by solving the system of equations

$$\begin{bmatrix} X_{11} & \cdots & X_{1S} \\ \vdots & \ddots & \vdots \\ X_{N1} & \cdots & X_{NS} \end{bmatrix} \cdot \begin{bmatrix} S_1 \\ \vdots \\ S_s \\ \vdots \\ S_S \end{bmatrix} = \begin{bmatrix} P_1 \\ \vdots \\ P_n \\ \vdots \\ P_N \end{bmatrix}, S = N$$

Given that the number of equations equals the number of unknown variables, i.e. $S = N$, and the rank of the matrix equals $S$, the solution of the equations is unique. Therefore, there exist unique state prices given that the market is complete. In addition, from equations (a8) and (a9), the risk-neutral probabilities are one-to-one correspond to state prices, and thus market completeness leads to unique risk-neutral probabilities. However, if the market is incomplete, then $S > N$. The number of equations is less than that of the unknown variables, and there are infinite groups of solutions for state prices. Accordingly, in an incomplete market, there are infinite groups of risk-neutral probabilities. It is easy to prove that the reverse of these two statements are true.

This important theorem can also be proved by applying Girsanov theorem under continuous time. The Girsanov theorem is given below. In a probability space $(\Omega, P, \mathcal{F})$, there is an Ito process of the form

$$dY_t = \beta(t, \omega)dt + \sigma(t, \omega)dB_t \quad (a10)$$

where $\beta(t, \omega)$ and $\sigma(t, \omega)$ are stochastic processes of dimension $n$ and $n \times m$, and $B_t$ is the Brownian motion of dimension $m$.

If there exists processes $u(t, \omega)$ and $\alpha(t, \omega)$ such that $\sigma(t, \omega)u(t, \omega) = \beta(t, \omega) - \alpha(t, \omega)$. Then we can define a Brownian under measure $Q$

$$d\hat{B}_t = u(t, \omega)dt + dB_t \quad (a11)$$

where measure $Q$ is defined as

$$dQ = M_T dP \quad (a12)$$

$$M_T = exp\left(-\frac{1}{2}\int_0^t u^2(s, \omega)ds - \int_0^t u(s, \omega)dB_s\right) \quad (a13)$$

Substitute $(a11)$ into $(a10)$, we have

$$dY_t = \beta(t, \omega)dt + \sigma(t, \omega)\left[d\hat{B}_t - u(t, \omega)dt\right] = [\beta(t, \omega) - \sigma(t, \omega)u(t, \omega)]dt + \sigma(t, \omega)d\hat{B}_t$$

Alternatively, it can be written as

$$dY_t = \alpha(t, \omega)dt + \sigma(t, \omega)d\hat{B}_t$$

The mission of Girsanov transformation is to find a measure such that $\alpha(t, \omega)$ equals 0 so that the stochastic process $Y_t$ under new measure $Q$ is a martingale. This stochastic process is a normalized asset prices process[20]. Suppose the market assets under probability space $(\Omega, P, \mathcal{F})$ are in the form

$$dY_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix} dt + \begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} dB_t, dB_t = \begin{bmatrix} dB_{1t} \\ dB_{2t} \end{bmatrix}$$

This is a complete market as there are two risk factors and two assets that are not linearly correlated. The equivalent martingale measure in this case is unique, we can find the martingale under measure $Q$ such that $\sigma(t, \omega)u(t, \omega) = \beta(t, \omega) - \alpha(t, \omega)$ and $\alpha(t, \omega) = 0$:

$$\begin{bmatrix} 1 & 2 \\ 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \xrightarrow{yields} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix}$$

---

[20] This is equivalent to find a price process that makes $\alpha(t, \omega)$ equal r(t, $\omega$), which is the risk-free interest rate process.

Accordingly, a unique equivalent martingale measure can be found in this market because the solution $u(t, \omega)$ is unique. However, if the risk factors are larger than the number of linearly uncorrelated assets, the market is incomplete. One can find different $u(t, \omega)$, substitute back to equation (a13) and (a12) and obtain infinitely numbers of equivalent martingale measures, under which the processes are martingales. For example, suppose the market assets under probability space $(\Omega, P, \mathcal{F})$ are in the form

$$dY_t = \begin{bmatrix} 1 \\ 0 \end{bmatrix} dt + \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} dB_t, dB_t = \begin{bmatrix} dB_{1t} \\ dB_{2t} \end{bmatrix}$$

Then there are infinitely numbers of solutions for the following system of equations. Therefore, there are infinitely numbers of equivalent martingale measures in the incomplete market.

$$\begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

Overall, we show the relationship between risk-neutral measure/equivalent martingale measure and market completeness. If there is a normalized martingale process, e.g. a continuous-time process with no drift term, the market is complete and admits no arbitrage opportunity.

From the proofs above, we know that the existence of a unique risk-neutral measure indicates that the market is complete and there is no arbitrage opportunity. The inverse proposition is only valid between the market completeness and uniqueness of risk-neutral measure. The no arbitrage condition will result in a unique risk neutral measure only in some special circumstances. These relations are demonstrated in figure 20. The first and second theorem are proved by Harrison and Kreps (1979) and Harrison and Pliska (1983) respectively, further interpretation is made by Øksendal (2010).
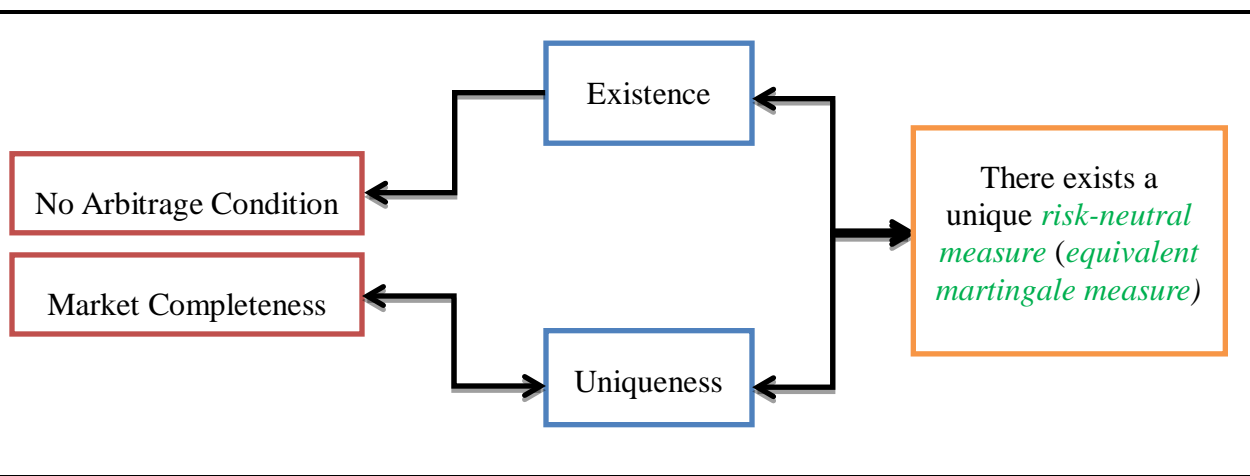


**Figure 20 The Fundamental Theorem of Asset Pricing**

## A3 Summary

In this section, both the equilibrium and arbitrage models are studied. The theoretical results presented are crucial in empirical analyses. Equations (a8) and (a9) indicate that given the payoffs at time $t$, state price density (SPD) and risk-neutral density can be used to price assets. Conversely, the SPD and the RND function for a specific date can be extracted from asset market. This is the theoretical guidance to derive RND function. Further, SPD and RND are be transformed conveniently by multiplying or dividing the risk-free total return. On the other hand, equation (a6) is applied to estimate the pricing kernel. Hansen and Singleton (1982, 1983) estimate pricing kernel using generalized method of moments (GMM) and maximum likelihood method (MLM). However, quality of aggregate consumption data is worse than that of the capital market data due to the measurement error and measurement frequency. Moreover, unlike the arbitrage models, equilibrium models are less accurate in short run because the market might deviate from equilibrium and the adjustment might take a long time. Arbitrage opportunities, on the other side, will be captured by arbitrageurs in minutes. Nevertheless, Rosenberg and Engle (2002) back out the empirical pricing kernel from option markets. This is possible since pricing kernel depends on the subjective probability density and RND. Combine equations (a6) and (a9), we have

$$e^{-rT} \int_0^\infty f_Q(S_T) \, X(S_T) \, dS_T = \int_0^\infty M_{0T} f_P(S_T) \, X(S_T) \, dS_T$$

$$M_{0T} = e^{-rT} \frac{f_Q(S_T)}{f_P(S_T)} \quad (a14)$$

Hence instead of estimating the stochastic discount factor directly, we can extract the RND function and subjective function. It is also possible to back out the representative agent's preference which is inherent in the stochastic discount factor. In section A.1.4, the stochastic discount factor is defined as

$$M_{0T} = \frac{\delta \, U'(C_T)}{U'(C_0)} \quad (a15)$$

Consider we are in a dynamic economy [for example, He and Leland (1993)], as we discussed in section 2.1. Since the representative investor optimally hold the market stock, $C_T = S_T$. Using $(a14)$ and $(a15)$, we can also obtain the implied risk aversion function

$$A(S_T) = \frac{f_P'(S_T)}{f_P(S_T)} - \frac{f_Q'(S_T)}{f_Q(S_T)}$$

This method is provided by Ait-Sahalia and Lo (2000), and the only different is that they derive it under a continuous-time equilibrium framework. This method yields the same expression as we did in section 2.1.

# Appendix B: Derivation of Double Lognormal Method

## B1 Geometric Brownian Motion and Ito's Lemma

Before going through the derivation, we should demonstrate the mathematical technique required. One of the most important mathematical techniques in quantitative finance is the Ito's lemma. However, only part of the results is provided. Øksendal (2010) gives a thorough and rigorous proof.

Suppose asset prices follow Geometric Brownian motion (GBM)

$$dS_t = \mu S_t dt + \sigma S_t dB_t \quad (b1)$$

$$Y_t = g(t, S_t) \quad (b2)$$

where $B_t$ is a stochastic process satisfying

(1) $B_0 = 0$;
(2) $B_t$ has independent increments;
(3) $B_t - B_s \sim \mathcal{N}(0, \sqrt{t-s})$.

Then

$$dY_t = \frac{\partial g(t, S_t)}{\partial t} dt + \frac{\partial g(t, S_t)}{\partial S_t} dS_t + \frac{1}{2} \frac{\partial^2 g(t, S_t)}{\partial S_t^2} (dS_t)^2 \quad (b3)$$

and

$$dt \cdot dt = 0, dt \cdot dB_t = dB_t \cdot dt = 0, dB_t \cdot dB_t = dt \quad (b4)$$

By substituting $(b1)$ into $(b2)$ and applying $(b3)$, we have

$$dY_t = \left[ \frac{\partial g(t, S_t)}{\partial t} + \frac{\partial g(t, S_t)}{\partial S_t} \mu S_t + \frac{1}{2} \frac{\partial^2 g(t, S_t)}{\partial S_t^2} \sigma^2 S_t^2 \right] dt + \frac{\partial g(t, S_t)}{\partial S_t} \sigma S_t dB_t \quad (b5)$$

If $(b2)$ becomes

$$Y_t = ln(S_t) \quad (b6)$$

Then

$$dln(S_t) = \left[ \mu S_t \cdot \frac{1}{S_t} + \frac{1}{2} \left( -\frac{1}{S_t^2} \right) \sigma^2 S_t^2 \right] dt + \frac{1}{S_t} \sigma S_t dB_t = \left( \mu - \frac{1}{2} \sigma^2 \right) dt + \sigma dB_t \quad (b7)$$

Therefore, we know that if asset prices $S_t$ follow GBM, then

$$\frac{\Delta S_t}{S_t} \sim \mathcal{N}(\mu \Delta t, \sigma \sqrt{\Delta t}) \quad (b8)$$

$$ln(S_t) - ln(S_0) \sim \mathcal{N}\left(\left(\mu - \frac{1}{2}\sigma^2\right)t, \sigma\sqrt{t}\right) \quad (b9)$$

$$ln(S_t) \sim \mathcal{N}\left(ln(S_0) + \left(\mu - \frac{1}{2}\sigma^2\right)t, \sigma\sqrt{t}\right) \quad (b10)$$

## B2 Double Lognormal Method

The derivation of Bahra's double lognormal method is based on the original Black-Scholes model and the equivalent martingale measure of Harrison and Kreps (1979). Instead of using one lognormal distribution, Bahra uses a combination of two lognormal distributions. Lognormal density function is defined as

$$L(\alpha, \beta; S_T) = \frac{1}{S_T \beta \sqrt{2\pi}} exp\left(-\frac{(ln(S_T) - \alpha)^2}{2\beta^2}\right) \quad (b11)$$

The formula to price the call option can be written as

$$C_{Call} = e^{-rt} \int_0^\infty [\omega L(\alpha_1, \beta_1; S_T) + (1 - \omega)L(\alpha_2, \beta_2; S_T)]Max(S_T - K, 0)\, dS_T$$

$$= e^{-rt} \int_K^\infty [\omega L(\alpha_1, \beta_1; S_T) + (1 - \omega)L(\alpha_2, \beta_2; S_T)](S_T - K)\, dS_T \quad (b12)$$

Substitute $(b11)$ into $(b12)$

$$C_{Call} = \frac{e^{-rt}}{\sqrt{2\pi}} \int_K^\infty \left[\frac{\omega}{\beta_1} exp\left(-\frac{(ln(S_T) - \alpha_1)^2}{2\beta_1^2}\right) + \frac{(1 - \omega)}{\beta_2} exp\left(-\frac{(ln(S_T) - \alpha_2)^2}{2\beta_2^2}\right)\right] dS_T$$

$$-\frac{Ke^{-rt}}{\sqrt{2\pi}} \int_K^\infty \left[\frac{\omega}{S_T\beta_1} exp\left(-\frac{(ln(S_T) - \alpha_1)^2}{2\beta_1^2}\right) + \frac{(1 - \omega)}{S_T\beta_2} exp\left(-\frac{(ln(S_T) - \alpha_2)^2}{2\beta_2^2}\right)\right] dS_T \quad (b13)$$

A transformation from lognormal distribution to normal distribution is made by substituting $X = ln(S_T)$ and $dS_T = exp(X)dX$ into the first integral term of $(b13)$. Since $S_T$ is lognormally distributed, X should follow normal distribution. Then the first term of $(b13)$ becomes

$$\frac{e^{-rt}}{\sqrt{2\pi}} \int_{ln(K)}^\infty \left[\frac{\omega}{\beta_1} exp\left(X - \frac{(X - \alpha_1)^2}{2\beta_1^2}\right) + \frac{(1 - \omega)}{\beta_2} exp\left(X - \frac{(X - \alpha_2)^2}{2\beta_2^2}\right)\right] dX$$

Reorganizing the terms

$$e^{-rt} exp\left(\alpha_1 + \frac{\beta_1^2}{2}\right)\frac{\omega}{\beta_1} \int_{ln(K)}^\infty \frac{1}{\sqrt{2\pi}} exp\left(\frac{-[X - (\alpha_1 + \beta_1^2)]^2}{2\beta_1^2}\right) dX$$

$$+ e^{-rt} exp\left(\alpha_2 + \frac{\beta_2^2}{2}\right)\frac{(1 - \omega)}{\beta_2} \int_{ln(K)}^\infty \frac{1}{\sqrt{2\pi}} exp\left(\frac{-[X - (\alpha_2 + \beta_2^2)]^2}{2\beta_2^2}\right) dX$$

Again, the normal distribution can be transformed to standard normal distribution when substituting
$Y_1 = [X - (\alpha_1 + \beta_1{}^2)]/\beta_1$ , $Y_2 = [X - (\alpha_2 + \beta_2{}^2)]/\beta_2$ and $dX = \beta_1 dY_1 = \beta_2 dY_2$

$$e^{-rt} exp\left(\alpha_1 + \frac{\beta_1{}^2}{2}\right)\omega \int_{ln(K)-(\alpha_1+\beta_1{}^2)/\beta_1}^{\infty} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Y_1{}^2\right) dY_1$$

$$+e^{-rt} exp\left(\alpha_2 + \frac{\beta_2{}^2}{2}\right)(1-\omega) \int_{ln(K)-(\alpha_2+\beta_2{}^2)/\beta_2}^{\infty} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Y_2{}^2\right) dY_2$$

This expression can be written in terms of cumulative normal distribution function

$$e^{-rt}\left[\omega \, exp\left(\alpha_1 + \frac{\beta_1{}^2}{2}\right) N\left(\frac{(\alpha_1 + \beta_1{}^2) - ln(K)}{\beta_1}\right)\right]$$

$$+e^{-rt}\left[(1-\omega) exp\left(\alpha_2 + \frac{\beta_2{}^2}{2}\right) N\left(\frac{(\alpha_2 + \beta_2{}^2) - ln(K)}{\beta_2}\right)\right] \quad (b14)$$

For the second integral term of $(b13)$, the same approach is used. Substitute $Y_1 = (ln(S_T) - \alpha_1)/\beta_1$, $Y_2 = (ln(S_T) - \alpha_2)/\beta_2$ and $dS_T = S_T\beta_1 dY_1 = S_T\beta_2 dY_2$

$$e^{-rt}K\left[\omega \int_{(ln(K)-\alpha_1)/\beta_1}^{\infty} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Y_1{}^2\right) dY_1 + (1-\omega) \int_{(ln(K)-\alpha_2)/\beta_2}^{\infty} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Y_2{}^2\right) dY_2\right]$$

Write the above expression in terms of cumulative distribution function

$$e^{-rt}K\left[\omega N\left(\frac{\alpha_1 - ln(K)}{\beta_1}\right) + (1-\omega) N\left(\frac{\alpha_2 - ln(K)}{\beta_2}\right)\right] \quad (b15)$$

Substitute $(b14)$ and $(b15)$ back to $(b13)$, and we obtain the closed-form solution for the call:

$$C_{Call} = e^{-rt}\left\{\omega\left[e^{\alpha_1 + \frac{\beta_1{}^2}{2}} N(d_1) - KN(d_2)\right] + (1-\omega)\left[e^{\alpha_2 + \frac{\beta_2{}^2}{2}} N(d_3) - KN(d_4)\right]\right\} \quad (b16)$$

where

$$d_1 = \frac{\alpha_1 + \beta_1{}^2 - ln(K)}{\beta_1}$$

$$d_2 = d_1 - \beta_1$$

$$d_3 = \frac{\alpha_2 + \beta_2{}^2 - ln(K)}{\beta_2}$$

$$d_4 = d_3 - \beta_2$$

The formula to price the put option can be written as

$$C_{Put} = e^{-rt} \int_0^{\infty} [\omega\, L(\alpha_1, \beta_1; S_T) + (1-\omega)L(\alpha_2, \beta_2; S_T)] Max(K - S_T, 0)\, dS_T$$

$$= e^{-rt} \int_0^{K} [\omega\, L(\alpha_1, \beta_1; S_T) + (1-\omega)L(\alpha_2, \beta_2; S_T)](K - S_T)\, dS_T \quad (b17)$$

Substitute $(b11)$ into $(b17)$

$$C_{Call} = \frac{Ke^{-rt}}{\sqrt{2\pi}} \int_0^{K} \left[ \frac{\omega}{S_T \beta_1} exp\left(-\frac{(ln(S_T) - \alpha_1)^2}{2\beta_1^2}\right) + \frac{(1-\omega)}{S_T \beta_2} exp\left(-\frac{(ln(S_T) - \alpha_2)^2}{2\beta_2^2}\right) \right] dS_T$$

$$- \frac{e^{-rt}}{\sqrt{2\pi}} \int_0^{K} \left[ \frac{\omega}{\beta_1} exp\left(-\frac{(ln(S_T) - \alpha_1)^2}{2\beta_1^2}\right) + \frac{(1-\omega)}{\beta_2} exp\left(-\frac{(ln(S_T) - \alpha_2)^2}{2\beta_2^2}\right) \right] dS_T \quad (b18)$$

A transformation from lognormal distribution to normal distribution is made by substituting $X = ln(S_T)$ and $dS_T = exp(X)dX$ into the second integral term of $(b18)$. Since $S_T$ is lognormally distributed, $X$ should follow normal distribution. Then the second term of $(b18)$ becomes

$$- \frac{e^{-rt}}{\sqrt{2\pi}} \int_0^{ln(K)} \left[ \frac{\omega}{\beta_1} exp\left(X - \frac{(X - \alpha_1)^2}{2\beta_1^2}\right) + \frac{(1-\omega)}{\beta_2} exp\left(X - \frac{(X - \alpha_2)^2}{2\beta_2^2}\right) \right] dX$$

Reorganizing the terms

$$-e^{-rt} exp\left(\alpha_1 + \frac{\beta_1^2}{2}\right) \frac{\omega}{\beta_1} \int_0^{ln(K)} \frac{1}{\sqrt{2\pi}} exp\left(\frac{-[X - (\alpha_1 + \beta_1^2)]^2}{2\beta_1^2}\right) dX$$

$$-e^{-rt} exp\left(\alpha_2 + \frac{\beta_2^2}{2}\right) \frac{(1-\omega)}{\beta_2} \int_0^{ln(K)} \frac{1}{\sqrt{2\pi}} exp\left(\frac{-[X - (\alpha_2 + \beta_2^2)]^2}{2\beta_2^2}\right) dX$$

Again, the normal distribution can be transformed to standard normal distribution when substituting $Y_1 = [X - (\alpha_1 + \beta_1^2)]/\beta_1$, $Y_2 = [X - (\alpha_2 + \beta_2^2)]/\beta_2$ and $dX = \beta_1 dY_1 = \beta_2 dY_2$

$$-e^{-rt} exp\left(\alpha_1 + \frac{\beta_1^2}{2}\right) \omega \int_0^{ln(K)-(\alpha_1+\beta_1^2)/\beta_1} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Y_1^2\right) dY_1$$

$$-e^{-rt} exp\left(\alpha_2 + \frac{\beta_2^2}{2}\right) (1-\omega) \int_0^{ln(K)-(\alpha_1+\beta_1^2)/\beta_1} \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}Y_2^2\right) dY_2$$

This expression can be written in terms of cumulative normal distribution function

$$-e^{-rt} \left[ \omega\, exp\left(\alpha_1 + \frac{\beta_1^2}{2}\right) N\left(\frac{ln(K) - (\alpha_1 + \beta_1^2)}{\beta_1}\right) \right]$$

$$-e^{-rt}\left[(1-\omega)exp\left(\alpha_2+\frac{\beta_2{}^2}{2}\right)N\left(\frac{ln(K)-\left(\alpha_2+\beta_2{}^2\right)}{\beta_2}\right)\right] \quad (b19)$$

For the first integral term of $(b18)$, the same approach is used. Substitute $Y_1=(ln(S_T)-\alpha_1)/\beta_1$, $Y_2=(ln(S_T)-\alpha_2)/\beta_2$ and $dS_T=S_T\beta_1dY_1=S_T\beta_2dY_2$

$$e^{-rt}K\left[\omega\int_0^{(ln(K)-\alpha_1)/\beta_1}\frac{1}{\sqrt{2\pi}}exp\left(-\frac{1}{2}Y_1{}^2\right)dY_1+(1-\omega)\int_0^{(ln(K)-\alpha_1)/\beta_1}\frac{1}{\sqrt{2\pi}}exp\left(-\frac{1}{2}Y_2{}^2\right)dY_2\right]$$

Write the above expression in terms of cumulative distribution function

$$e^{-rt}K\left[\omega N\left(\frac{ln(K)-\alpha_1}{\beta_1}\right)+(1-\omega)N\left(\frac{ln(K)-\alpha_2}{\beta_2}\right)\right] \quad (b20)$$

Substitute (b19) and (b20) back to (b18), and we obtain the closed-form solution for the put

$$C_{Put}=e^{-rt}\left\{\omega\left[-e^{\alpha_1+\frac{\beta_1{}^2}{2}}N(-d_1)+KN(-d_2)\right]+(1-\omega)\left[-e^{\alpha_2+\frac{\beta_2{}^2}{2}}N(-d_3)+KN(-d_4)\right]\right\} \quad (b21)$$

where

$$d_1=\frac{\alpha_1+\beta_1{}^2-ln(K)}{\beta_1}$$

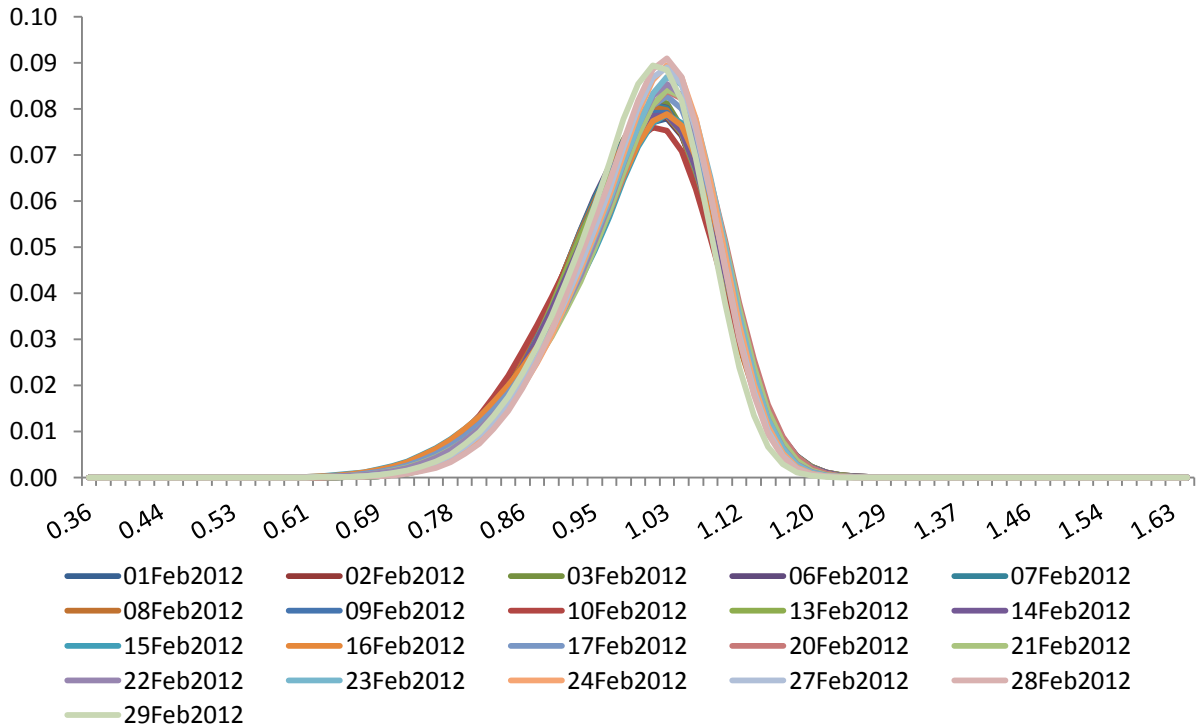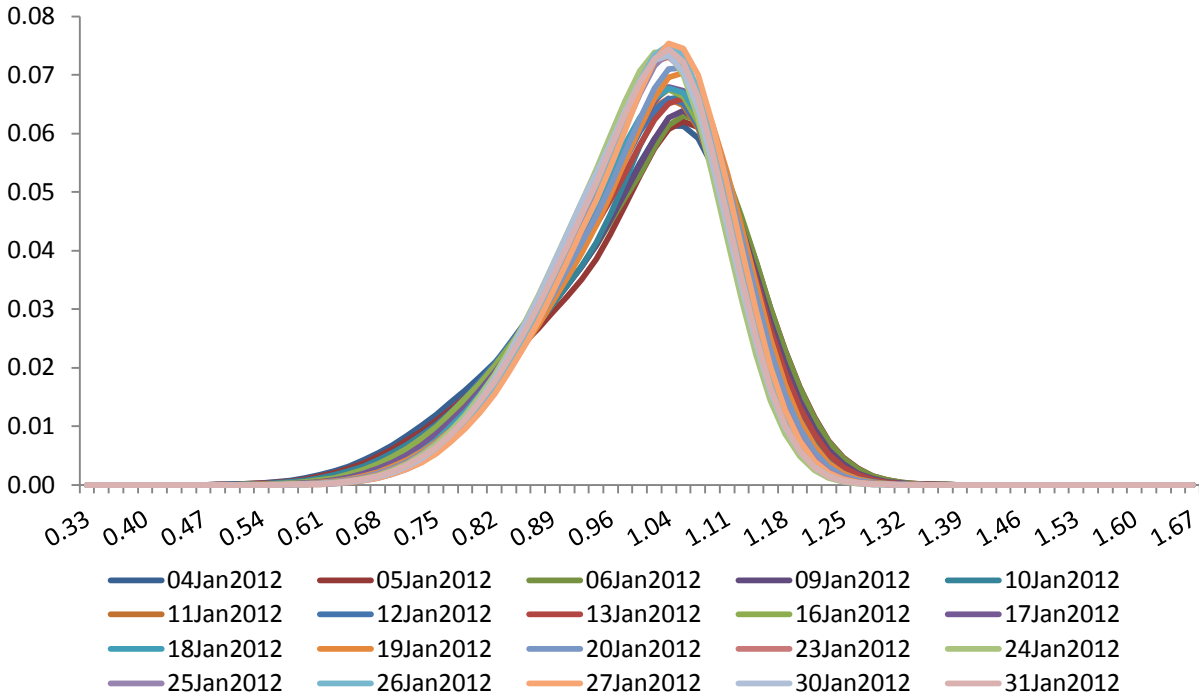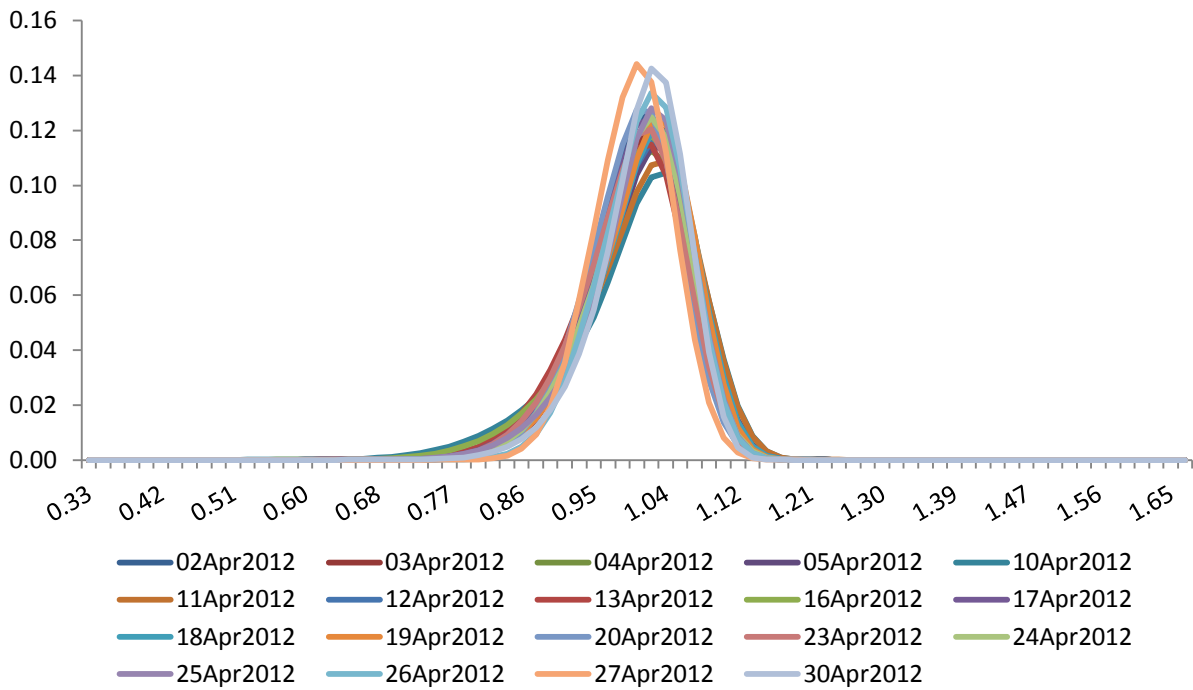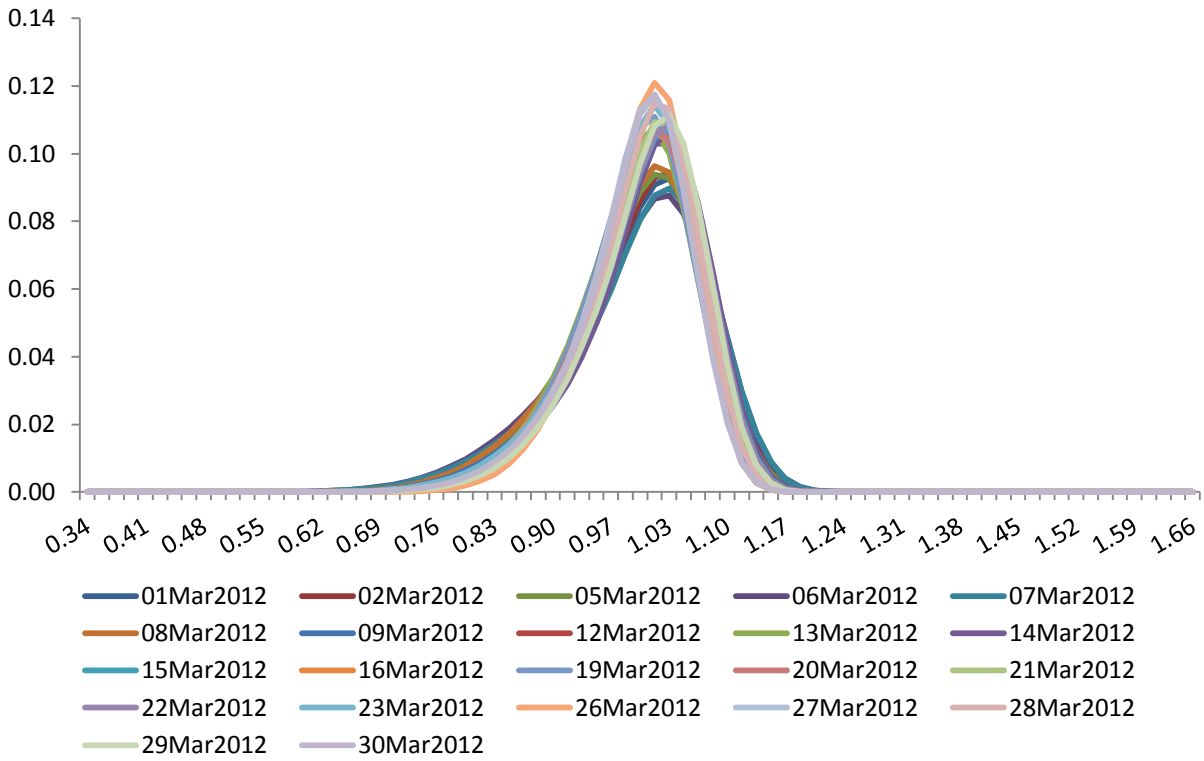$$d_2=d_1-\beta_1$$

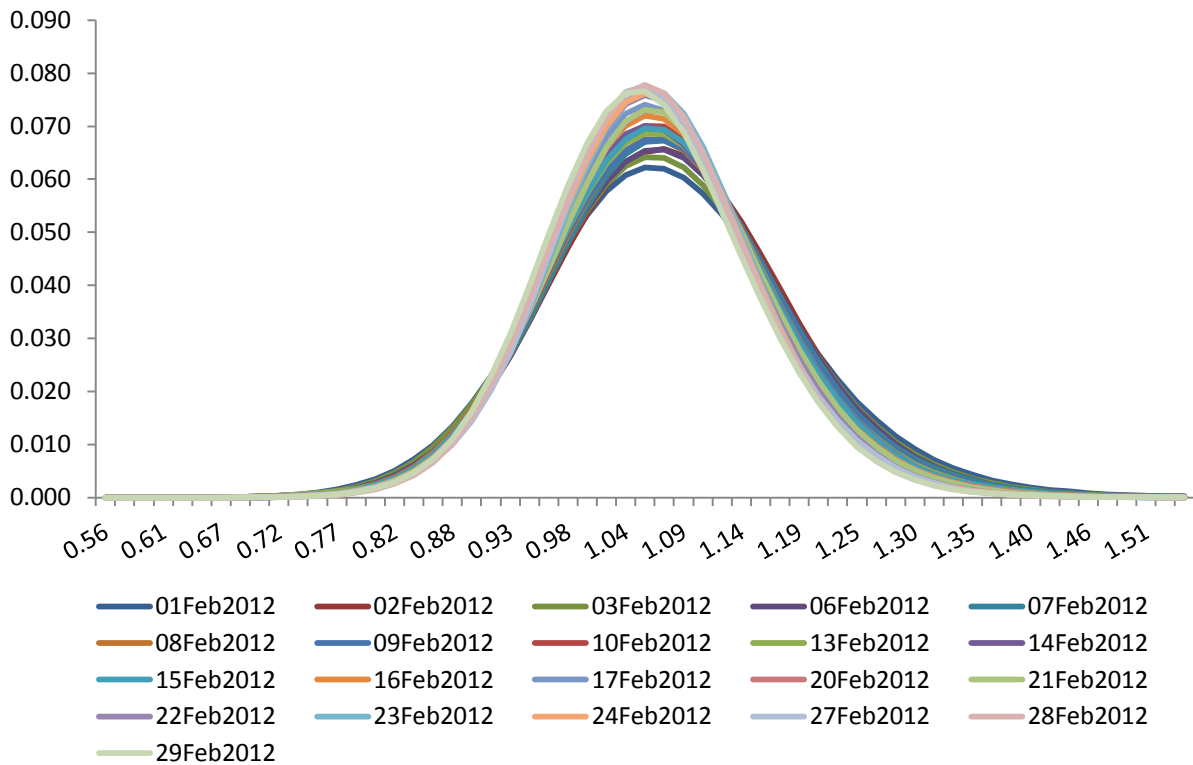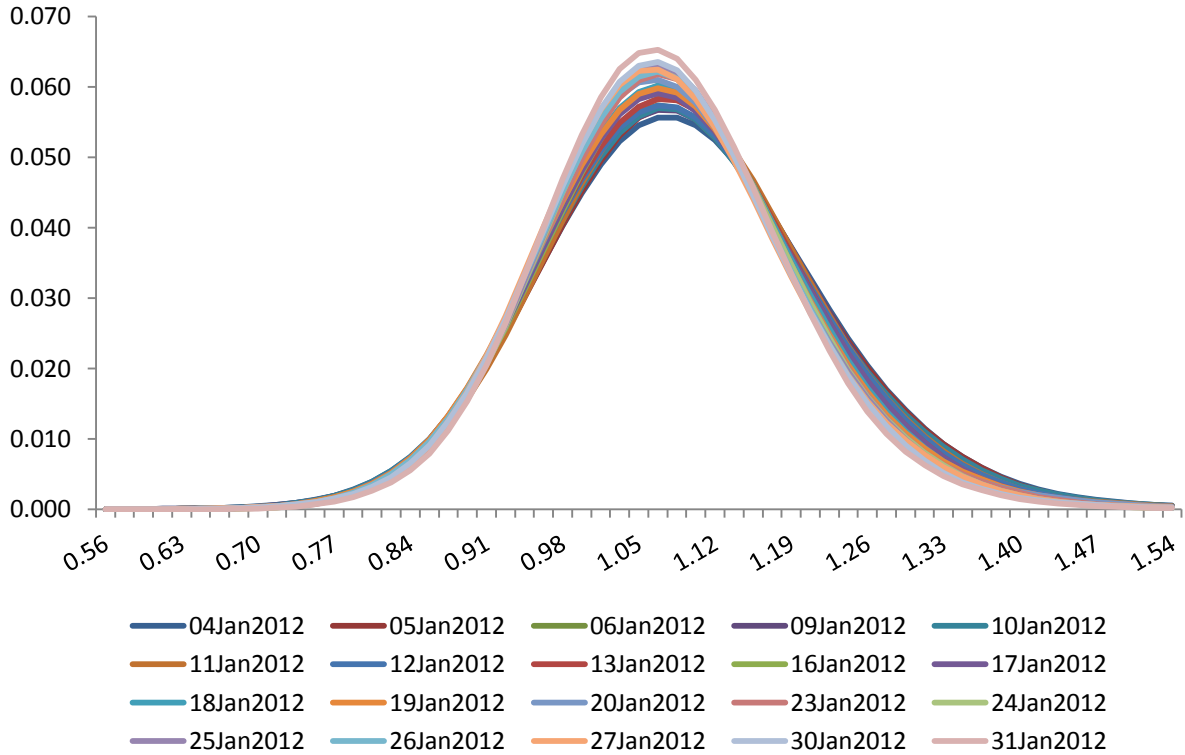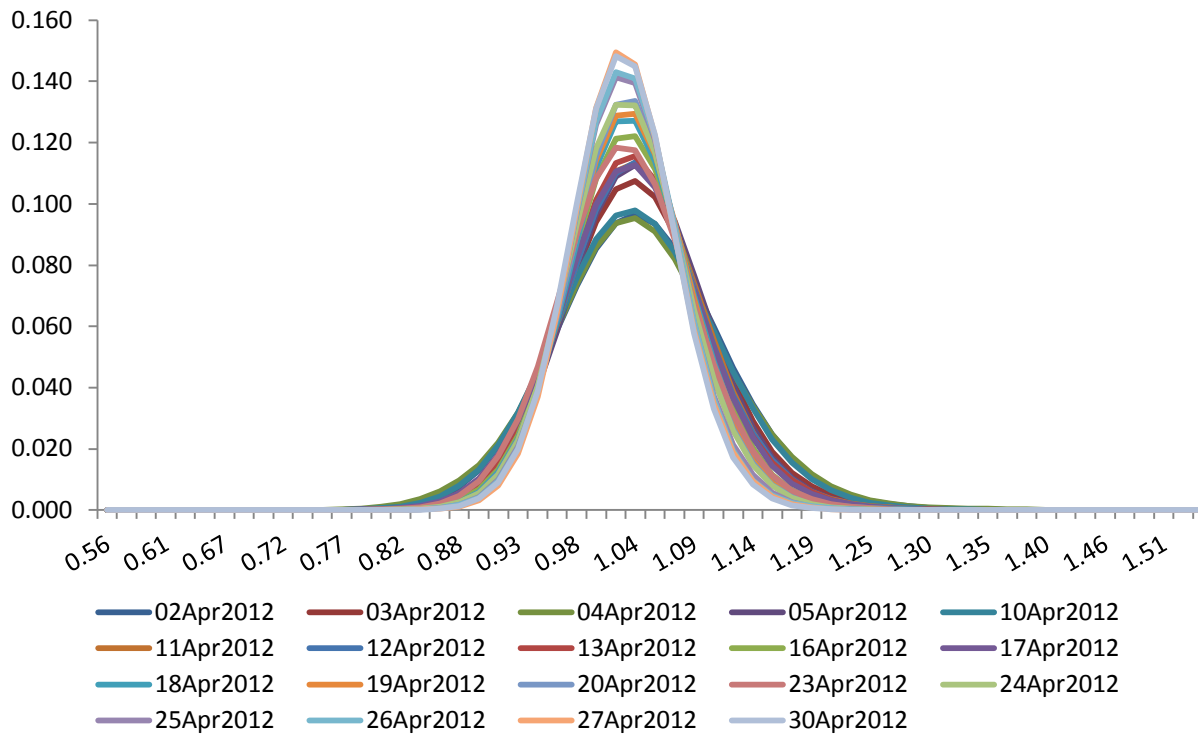$$d_3=\frac{\alpha_2+\beta_2{}^2-ln(K)}{\beta_2}$$
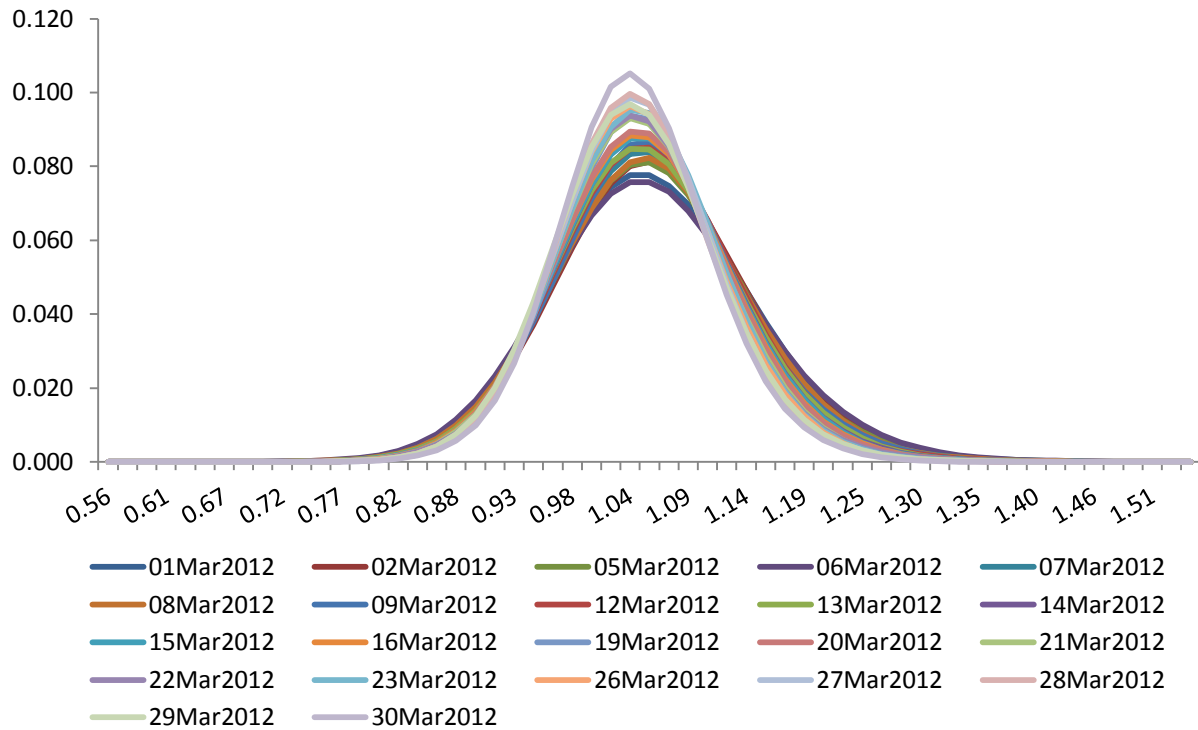
$$d_4=d_3-\beta_2$$

# Appendix C: Densities and Risk Aversion Functions
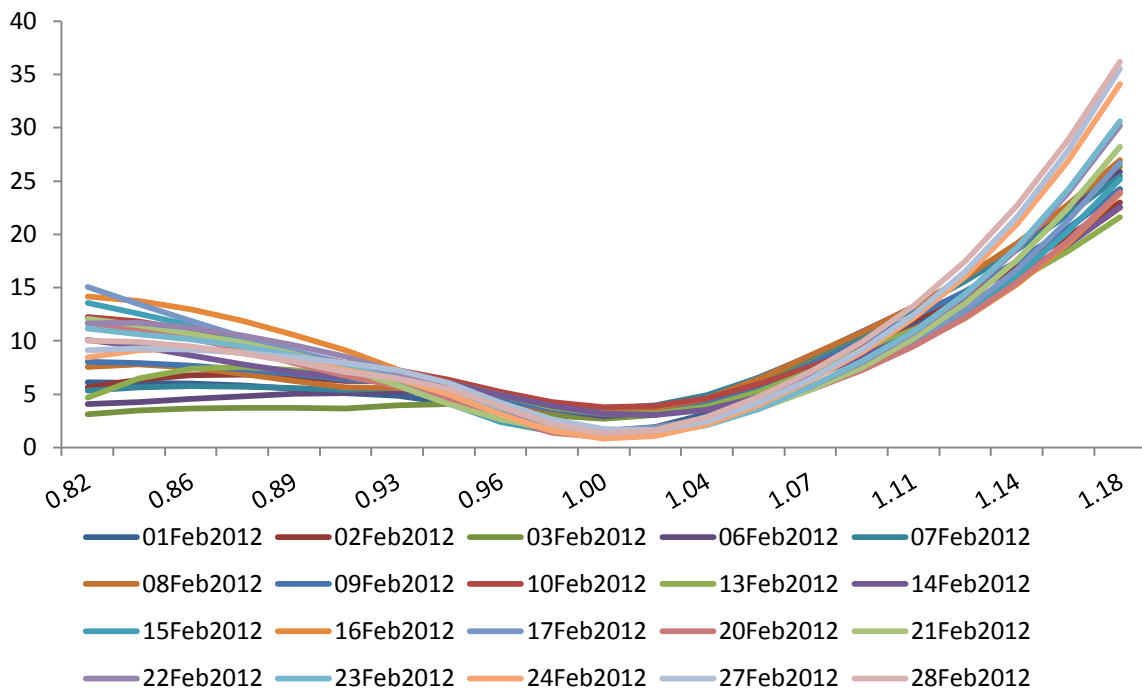
## C1 Risk Neutral Densities



04Jan2012 05Jan2012 06Jan2012 09Jan2012 10Jan2012 11Jan2012 12Jan2012 13Jan2012 16Jan2012 17Jan2012 18Jan2012 19Jan2012 20Jan2012 23Jan2012 24Jan2012 25Jan2012 26Jan2012 27Jan2012 30Jan2012 31Jan2012



01Feb2012 02Feb2012 03Feb2012 06Feb2012 07Feb2012 08Feb2012 09Feb2012 10Feb2012 13Feb2012 14Feb2012 15Feb2012 16Feb2012 17Feb2012 20Feb2012 21Feb2012 22Feb2012 23Feb2012 24Feb2012 27Feb2012 28Feb2012 29Feb2012
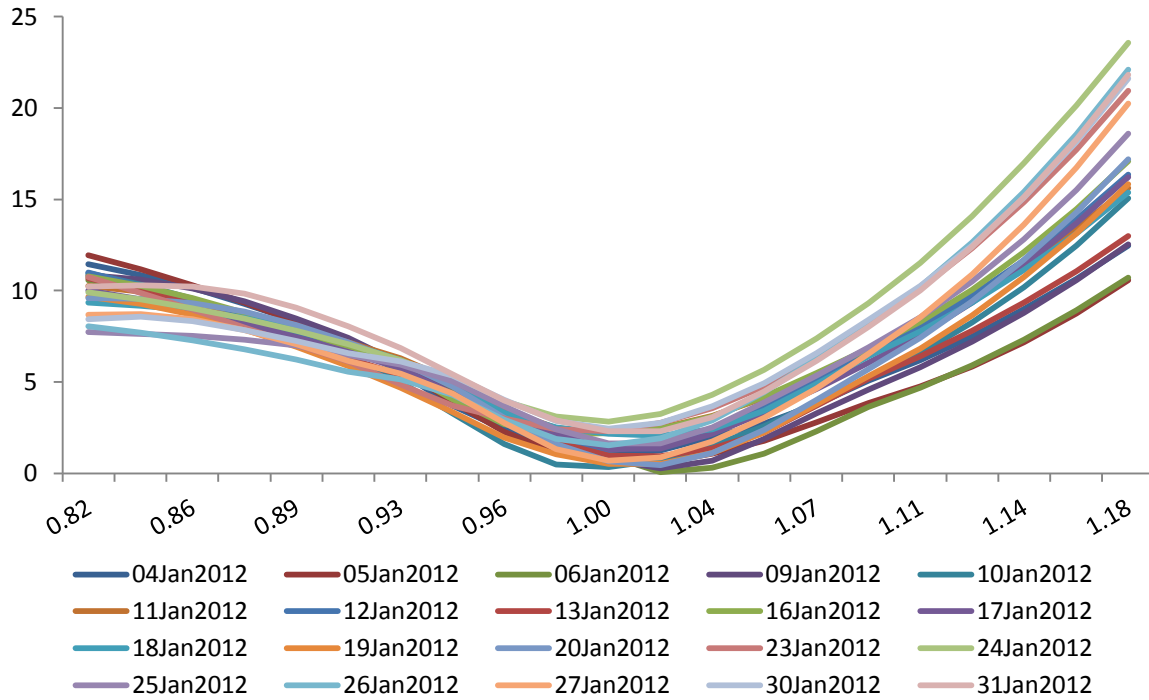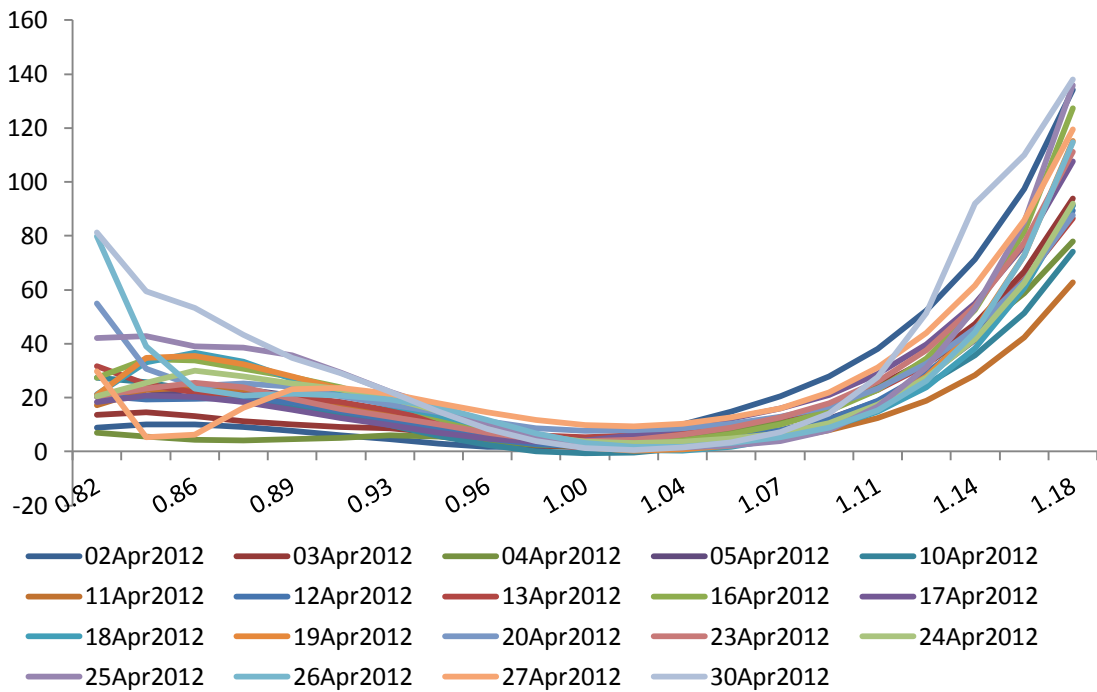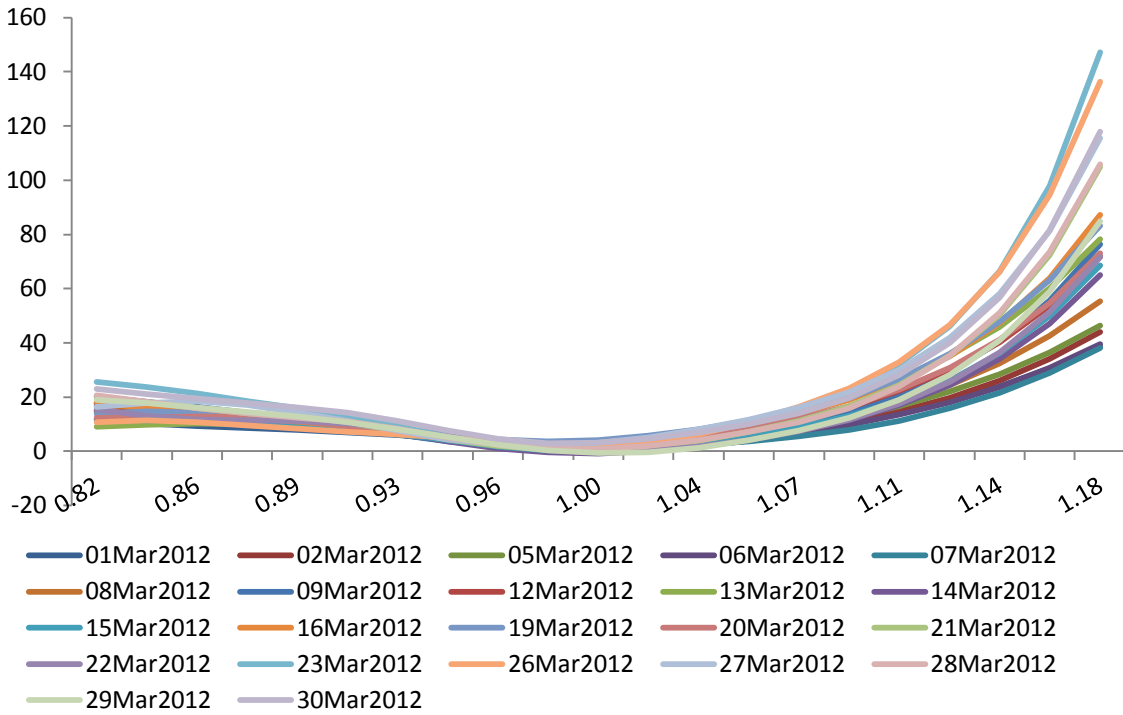
# C2 Subjective Densities

# C3 Risk Aversion Functions

# Bibliography

Ait-Sahalia, Yacine and Andrew Lo (1998). Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance*, 499-547.

Ait-Sahalia, Yacine and Andrew Lo (2000). Nonparametric Risk Management and Implied Risk Aversion. *Journal of Econometrics*, 9-51.

Arrow, Kenneth (1964). The Role of Securities in the Optimal Allocation of Risk-Bearing. *Review of Economic Studies*, 91-96.

Arrow, Kenneth and Gerard Debreu (1954). Existence of a Competitive Equilibrium for a Competitive Economy. *Econometrica*, 265-290.

Back, Kerry (2010). *Asset Pricing and Portfolio Choice Theory.* New York: Oxford University Press.

Bahra, Bhupinder (1997). Implied Risk-Neutral Probability Density Function from Option Prices: Theory and Application. *Workng Paper*, Bank of England.

Bates, David (1991). The Crash of '87: Was It Expected? The Evidence from Options Markets. *Journal of Finance*, 1009-1044.

Black, Fischer and Myron Scholes (1973). The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 637-654.

Bliss, Robert and Nikolaos Panigirtzoglou (2002). Testing the Stability of Implied Probability Density Function. *Journal of Banking and Finance*, 381-422.

Bliss, Robert and Nikolaos Panigirtzoglou (2004). Option-Implied Risk Aversion Estimates. *Journal of Finance*, 407-446.

Bookstaber, Richard and James McDonald (1987). A General Distribution for Describing Security Price Returns. *Journal of Business*, 401-424.

Breeden, Douglas (1979). An Intertemporal Asset Pricing Model with Stochastic Consumption and Investment Opportunities. *Journal of Financial Economics*, 265-296.

Breeden, Douglas and Robert Litzenberger (1978). Prices of State-contingent Claims Implicit in Option Prices. *Journal of Business*, 621-651.

Campbell, John, Andrew Lo and Craig MacKinlay (1997). *The Econometrics of Financial Markets.* New Jersey: Princeton University Press.

Castren, Olli (2005). Estimating and Aanalysing Currency Options Implied Risk-Neutral Density Functions for the Largest New EU Member States. *Working Paper*, European Central Bank.

Constantinides, George (1982). Intertemporal Asset Pricing with Heterogeneous Consumers and without Demand Aggregation. *Journal of Business*, 253-267.

Cox, John and Hayne Leland (2000). On Dynamic Investment Strategies. *Journal of Economic Dynamics and Control*, 1859-1880.

Cox, John, Jonathan Ingersoll and Stephen Ross (1985). A Theory of the Term Structure of Interest Rates. *Econometrica*, 385-407.

Cox, John and Stephen Ross (1976). The Valuation of Options for Alternative Stochastic Processes. *Journal of Financial Economics*, 145-166.

Danthine, Jean-Pierre and John Donaldson (2005). *Intermediate Financial Theory.* San Diego: Elsevier Academic Press.

Debreu, Gerard (1959). *Theory of Value.* New York: John Wiley & Sons.

Figlewski, Stephen (2008). Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio. *Working Paper*, New York University Stern School of Business.

Glosten, Lawrence, Ravi Jagannathan and David Runkle (1993). On the Relation Between the Expected Value And the Volatility of Nominal Excess Return on Stocks. *Journal of Finance*, 1779-1801.

Hansen, Lars and  Kenneth Singleton (1983). Stochastic Consumption, Risk Aversion, and the Temporal Behavior of Asset Returns. *Journal of Political Economy*, 249-265.

Hansen, Lars and Kenneth Singleton (1982). Generalized Instrumental Variables Estimation of Nonlinear Rational Expectations Models. *Econometrica*, 1269-1286.

Harrison, Michael and David Kreps (1978). Speculative Investor Behavior in a Stock Market with Heterogeneous Expectations. *Quarterly Journal of Economics*, 323-336.

Harrison, Michael and David Kreps (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory*, 381-408.

Harrison, Michael and Stanley Pliska (1983). A Stochastic Calculus Model of Continuous Trading: Complete Markets. *Stochastic Processes and their Applications*, 313-316.

He, Hua and Hayne Leland (1993). On Equilibrium Asset Price Processes. *Review of Financial Studies*, 593-617.

Heston, Steven (1993). A Close-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *Review of Financial Studies*, 327-343.

Hordahl, Peter and David Vestin (2005). Interpreting Implied Risk-Neutral Densities: The Role of Risk Premia. *Review of Finance*, 97-137.

Hull, John (2006). *Options, Futures, and Other Derivatives.* New Jersey: Prentice Hall.

Hull, John and Alan White (1987). The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance*, 281-300.

Jackwerth, Jens (2000). Recovering Risk Aversion from Option Prices and Realized Returns. *The Review of Financial Studies*, 433-451.

Jackwerth, Jens (1998). Recovering Stochastic Processes from Option Prices. *Working paper*, University of Konstanz.

Jackwerth, Jens (2004). *Option-Implied Risk-Neutral Distributions and Risk Aversion.* Charlotteville: Research Foundation of AIMR.

Jackwerth, Jens and Mark Rubinstein (1996). Recovering Probability Distributions from Option Prices. *Journal of Finance*, 1611-1631.

Lochstoer, Lars (2006). Expected Returns and the Business Cycle: Heterogeneous Agents and Heterogeneous Goods. *Working Paper*, London Business School.

Lucas, Robert (1978). Asset Pricing in an Exchange Economy. *Econometrica*, 1429-1445.

Malz, Allan (1997). Estimating the Probability Distribution of the Future Exchange Rate from Options Prices. *Journal of Derivatives*, 18-36.

Mas-Colell, Andreu, Michael Whinston and Jerry Green (1995). *Microeconomic Theory.* New York: Oxford University Press.

Mayhew, Stewart. (1995). Implied Volatility. *Financial Analysts Journal*, 8-20.

Mehra, Rajnish and Edward Prescott (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics*, 145-161.

Melick, William and Charles Thomas (1997). Recovering an Asset's Implied PDF from Option Prices: An Application to Crude Oil during the Gulf Crisis. *Journal of Financial and Quantitative Analyst*, 91-115.

Merton, Robert (1973). An Intertemporal Capital Asset Pricing Model. *Econometrica*, 867-887.

Merton, Robert (1969). Lifetime Portfolio Selection Under Uncertainty. *Review of Economics and Statistics*, 247-257.

Øksendal, Brent (2010). *Stochastic Differential Equation.* New York: Springer-Verlag.

Pennacchi, George (2007). *Theory of Asset Pricing.* New York: Pearson.

Perignon, Christophe and Christophe Villa (2002). Extracting Information from Options Markets: Smiles, State-Price Densities and Risk Aversion. *European Finainical Management*, 495-513.

Ritchey, Robert (1990). Call Option Valuation for Discrete Normal Mixtures. *Journal of Financial Research*, 285-296.

Rosenberg, Joshua and Robert Engle (2002). Empirical Pricing Kernels. *Journal of Financial Economics*, 341-372.

Rosenberg; Joshua (1998). Pricing Multivariate Contingent Claims Using Estimated Risk-Neutral Density Functions. *Journal of International Money and Finance*, 229-247.

Ross, Stephen (1976). Options and Efficiency. *Quarterly Journal of Economics*, 75-89.

Rubinstein, Mark (1994). Implied Binomial Tree. *Journal of Finance*, 771-818.

Shimko, David (1993). Bounds of Probability. *Risk*, 33-37.

Silverman, Bernard (1986). *Density Estimation for Statistics and Data Analysis.* New York: Chapman and Hall.

Taylor, Stephen (2005). *Asset Price Dynamics, Volatility, and Prediction.* NewJersey: Princeton University Press.

Varian, Hal (1992). *Microeconomic Analysis.* New York: W. Norton & Company, Inc.

Wang, Susheng (1993). The Integrability Problem of Asset Prices. *Journal of Economic Theory*, 199-213.

Yatchew, Adonis and Wolfgang Hardle (2006). Nonparametric State Price Density Estimation using Constrained Least Squared and the Bootstrap. *Journal of Econometrics*, 579-599.

Ziegler, Alexandre (2007). Why Does Implied Risk Aversion Smile? *Review of Financial Studies*, 859-904.