



Antti Pirjetä – Seppo Ikäheimo – Vesa Puttonen

# SEMIPARAMETRIC RISK PREFERENCES IMPLIED BY EXECUTIVE STOCK OPTIONS

Antti Pirjetä – Seppo Ikäheimo – Vesa Puttonen

# SEMIPARAMETRIC RISK PREFERENCES IMPLIED BY EXECUTIVE STOCK OPTIONS

A Financial Markets and Services Research Paper  
For more finance related research at HSE, see [www.hse.fi/fms/](http://www.hse.fi/fms/)

May  
2007

HELSINGIN KAUPPAKORKEAKOULU  
HELSINKI SCHOOL OF ECONOMICS  
PL 1210  
FI-00101 HELSINKI  
FINLAND

© Antti Pirjetä, Seppo Ikkäheimo, Vesa Puttonen and  
Helsinki School of Economics

ISSN 1235-5674  
(Electronic working paper)  
ISBN 978-952-488-147-0

Helsinki School of Economics -  
HSE Print 2007

# Semiparametric risk preferences implied by executive stock options

JEL codes: M52, G13

Keywords: risk preferences, executive stock options (ESOs), risk aversion

November 30, 2006

## **Antti Pirjetä**

Dept. of Business Technology  
Helsinki School of Economics  
E-mail: [pirjeta@hse.fi](mailto:pirjeta@hse.fi)

## **Seppo Ikäheimo**

Dept. of Accounting and Finance  
Helsinki School of Economics  
[seppo.ikaheimo@hse.fi](mailto:seppo.ikaheimo@hse.fi)

## **Vesa Puttonen\***

Dept. of Accounting and Finance  
Helsinki School of Economics  
P.O.Box 1210, FIN-00101  
E-mail: [vesa.puttonen@hse.fi](mailto:vesa.puttonen@hse.fi)  
Tel: +358-9-43138405  
Fax: +358-9-43138678

\* Corresponding author

# Semiparametric risk preferences implied by executive stock options

## **Abstract**

We examine the pricing of executive stock options (ESOs) and managerial preferences using unique market data on ESOs. First we look at the pricing issue and calibrate the extended Black-Scholes model of Ingersoll (2006). We find that ESO prices contain a subjective risk premium (SRP) with median value of 8.1%, and the SRP decreases with moneyness of the option. We then estimate managerial risk preferences using the semiparametric method of Aït-Sahalia and Lo (2000). Our results suggest that relative risk aversion is just above 1 for a certain stock price interval. This level of risk aversion is low but reasonable, and it may be explained by the fact that the typical manager is wealthy and his marginal utility is low. Further, marginal rate of substitution increases considerably in states with low stock prices.

---

## 1. Introduction

The risk preferences of company management play an important role while they decide on capital investments and capital structure, and when various types of executive compensation effect on managerial behavior (Lambert, Larcker & Verrecchia, 1991; Ross, 2004). In most studies, the risk preferences have either fully neglected by implicitly defining managers as risk neutral agents or the models have varied managerial relative risk aversion (RRA) between 1 and 50 without really knowing the magnitude of it. Although this magnitude of RRA is of great importance both for analytical purposes and practical decision making, no empirically based estimates have yet been presented. This study fills the gap by estimating RRA based on the executive stock option (ESO) trades with the use of Ingersoll (2006) incentive option pricing model, Ait-Sahalia and Lo (1988, 2000) semiparametric estimator of relative risk aversion and a smooth volatility function estimated in delta-sigma space as in Bliss and Panigirtzoglou (2002, 2004). We use the word semiparametric, because our method inserts a nonparametric volatility function in parametric risk-neutral density function, and adds a nonparametric market density function to estimate risk preferences.

Such an estimation of RRA would provide guidance for deriving agency theory based models with managerial risk aversion similar to Parrino, Poteshman & Weisbach (2002) where they made a model on managerial investment behavior, with positive RRA. Their model resulted in contradictory results compared with the Myers (1977) model without such an assumption.

This result is also useful for evaluating executive compensation plans. The major question, how much compensation for top executives should increase compared with the shareholder value gains, would have different answer depending on the level of RRA. The agency theory results presented by Jensen and Murphy (1990) indicate that managers are not properly paid when the pay-performance sensitivity is only 0.003, whereas the results by Haubrich (1994) and Haubrich and Popova (1998) where RRA varied from about 1.1 upwards suggest need for rather low pay-performance sensitivity of close to 0.01 for wealthy manager.

In addition, the efficiency of stock option plans would be different depending on the level of RRA. While on the other hand higher RRA increases the dead-weight loss between the value of executive stock options between portfolio investors and undiversified managers (Hall and Murphy (2002); Ingersoll (2006)), higher RRA may also influence on the incentive effect of stock options to undertake risky projects either by increasing (Parrino, Poteshman & Weisbach (2002)) or by decreasing it (Ross (2004)), and the convexity of compensation contract similar to stock options would be optimal with moderate levels (close to one) of managerial RRA (Hemmer, Kim & Verrecchia, 2000). The positive

values of RRA may also motivate managers with incentive options to control the risk level (volatility) within the firm to increase the subjective values of these incentive options (Hodder and Jackwerth (2005)).

To attain an empirical estimate of RRA, we use the subjective valuation of incentive options of Ingersoll (2006), which is an extended Black-Scholes model offering a closed-form solution to measure *subjective risk premium*. The adjustment to risk free rate is calculated by calibrating the model to actual ESO prices. Using this model we show, how much the manager's empirical pricing kernel differs from the risk-neutral pricing kernel. In addition, we apply the semiparametric method of Aït-Sahalia and Lo (1998, 2000) to estimate the option-implied probability density function (PDF) in two steps. First, we fit a smooth volatility function in data. We model volatility as a function of delta, following Bliss and Panigirtzoglou (2004) and Kang and Kim (2006), and then estimate implied strike prices from the deltas. Second, we insert the volatility function in a lognormal density function, implied by the asset dynamics of the Black-Scholes model. The Aït-Sahalia and Lo method produces estimates for relative risk aversion and marginal rate of substitution, both as functions of underlying share price.

Our empirical estimation of managerial RRA is based on the data on the market quotations of global technology firm Nokia executive stock option (ESO) trades. Our data consists of 7,610 trades in the Nokia ESOs issued in 1999. According to Nokia's annual report of 1999, these options were issued in March 1999 to about 5000 Nokia managers and other key employees around the world. The options had a vesting period of two years, after which they were listed on the Helsinki Stock Exchange until they expired at the end of 2004. After the vesting period, these employees are free to sell their ESOs, and in most cases the trades were made via Helsinki Stock Exchange. Given these facts our results on implied preferences characterize the representative manager working for a multinational technology firm, instead of reflecting the tastes of a single nationality.

Our results regarding the risk aversion coefficient shows a stable RRA estimate of just above 1 for a wide price range of ESOs and a downward sloping marginal rate of substitution (MRS). This serves the intuition that low levels of consumption are associated with low stock prices. In these states marginal utility is high and the willingness to save is low. Our results on the RRA have a major implication on the agency theory based modeling as well as on structuring compensation packages for company managers and the incentive effects of stock options.

---

## 2. Framework for subjective option pricing

### 2.1 The subjective pricing kernel

Ingersoll (2006) solves the consumption-investment problem of a constrained manager and derives closed-form solutions for equity and option values. He adds two complications to the standard Black-Scholes model. The first one is that until the manager retires, he is required to hold a positive proportion of his wealth in the company stock. The second complication is idiosyncratic (stock-specific) risk that cannot be hedged in the market, even if the manager is allowed to trade in the market portfolio. An important advantage of this framework is that it satisfies the no-arbitrage condition for the manager who prices all assets subjectively. This is equivalent to assuming the existence of a subjective pricing kernel, and because it has a risk-neutral drift that is lower than the risk-free rate, subjective asset values will be lower than objective asset values.

We make two simplifying assumptions to the Ingersoll (2006) framework. The first one is that there is only one factor of systematic risk, being the market risk. The second assumption is that the portfolio constraint  $\theta$  equals the stock weight. In theory,  $\theta$  would be the excess weight over a benchmark index. In practice, the weight of any single stock is close to zero in a global equity index, so we set  $\theta$  equal to stock weight.

To get started, assume that the manager has power utility function, or  $U(C_t) = (C_t^{1-\gamma})/1-\gamma$ , where  $C_t$  is current period consumption and  $\gamma$  is the risk aversion parameter. The model has two risky assets, market portfolio  $M$  and company stock  $S$  with expected returns  $\mu_m$  and  $\mu$ . The manager is allowed to trade in the market portfolio. In addition there is a risk-free bond yielding  $r$ . It is assumed that the continuous-time CAPM holds, so the market portfolio is efficient. The model also accounts for dividend yields on market portfolio and the employer stock, denoted  $q_m$  and  $q$ . Both risky assets follow geometric Brownian motion, and they are correlated as shown by equations (1).

$$(1) \quad \begin{aligned} \frac{dM}{M} &= (\mu_m - q_m)dt + \sigma_m dB_m \\ \frac{dS}{S} &= (\mu - q)dt + \beta\sigma_m dB_m + \nu dB. \end{aligned}$$

Equations (1) use standard notation.  $B_m$  and  $B$  denote Brownian motions that generate the distributions of market and idiosyncratic risk factors. It is assumed that the market and idiosyncratic

risks are uncorrelated. Under this assumption the variance of company stock is the sum of market and idiosyncratic variances, or  $\sigma^2 = \beta^2 \sigma_m^2 + v^2$ .

Ingersoll (2006) solves the manager's problem in two stages to find out the effect of a positive holding constraint in the company stock. The first stage solves a standard Merton investment problem, where the resulting portfolio is unconstrained. The second stage involves solving the constrained problem. Comparing of the solutions gives the managerial hedging demand, which is the difference between unconstrained and constrained market portfolio weights. It follows from the Merton (1971) mutual fund theorem that demand for the company stock is zero in the unconstrained problem<sup>1</sup>, and if we force it be positive, the constrained manager will invest less in the market portfolio than he otherwise would.

Solution of this consumption-investment problem, derived in Ingersoll (2006), yields a subjective pricing kernel, where the risk-free rate is adjusted to compensate for the portfolio constraint  $\theta$  and idiosyncratic risk. The adjusted risk-free rate  $\hat{r}$  is given by equation (2), where  $\gamma$  is the relative risk aversion. Note that the adjusted risk-free rate is decreasing in both the portfolio constraint and idiosyncratic risk. For the same reasons the dividend rate  $q$  is also adjusted. The subjective dividend rate  $\hat{q}$  is given by equation (3).

$$(2) \quad \hat{r} \equiv r - (1 - \gamma)\theta^2 v^2$$

$$(3) \quad \hat{q} \equiv q + (1 - \gamma)(1 - \theta)\theta v^2$$

## 2.2 The subjective option pricing model

We derive the partial differential equation (PDE) for the subjective valuation to stress that arbitrage pricing in this framework means valuing all assets using the same subjective pricing kernel. We employ a simple arbitrage argument that is based on chapter 6 of Björk (1998). The resulting PDE gives rise to a closed-form option pricing formula, which we call the *Black-Scholes-Ingersoll* formula. Ingersoll (2006) derives the same PDE using martingale methods, but we believe that the arbitrage pricing approach is more intuitive. Arbitrage pricing is based on the idea that the option can be replicated. Since the underlying stock involves idiosyncratic risk, it is clear that the replicating portfolio consists of underlying stock and riskfree bond. The dynamics of these assets are given by

---

<sup>1</sup> Here 'demand' refers to demand of the company stock in excess of its share in the market portfolio.

---


$$(4) \quad \begin{aligned} dS/S &= (\mu - q)dS + \beta\sigma_m dB_m + vdB \\ dB/B &= rdt. \end{aligned}$$

In eq. (4)  $\mu$  and  $q$  are the expected stock return and volatility. If we denote the price process of the option as  $\Pi(t) = F(S(t), t)$ , an application of Itô's lemma gives the following dynamics for  $\Pi(t)$ . The notation needs some clarification:  $\sigma_s = (\beta\sigma_m, v)'$  and  $\mathbf{1} = (1, 1)'$ .

$$(5) \quad \begin{aligned} d\Pi/\Pi &= \alpha_\Pi dt + \sigma_\Pi dB \\ \alpha_\Pi &= \frac{F_S S(\mu - q) + F_t + 1/2 F_{SS} S^2 (\sigma'_s \mathbf{1})}{F} \\ \sigma_\Pi &= \frac{F_S S (\sigma'_s \mathbf{1})}{F}. \end{aligned}$$

The next step is to form a risk-free portfolio from the underlying stock and its derivative. The arbitrage portfolio has the properties given by equation (6). Note that we do not restrict the stock weight here, because the goal is to find the arbitrage price for an incremental contingent claim. The portfolio restriction is accounted in later stage by the adjusted risk-free rate  $\hat{r}$ .

$$(6) \quad \begin{aligned} u_s + u_\Pi &= 1 \\ u_s \sigma'_s \mathbf{1} + u_\Pi &= 0. \end{aligned}$$

We can solve the arbitrage portfolio weights  $(u_s, u_\Pi)$  by combining (5) and (6). Note that the replicating strategy is the same whether the derivative is priced using subjective or market measure. This is why we haven't assumed anything yet about the return on portfolio  $(u_s, u_\Pi)$ .

$$(7) \quad \begin{aligned} u_\Pi &= -F/(F_S S - F) \\ u_s &= F_S S/(F_S S - F). \end{aligned}$$

Now we will put the subjective risk-neutral measure to work in order to derive the pricing PDE. Under this measure the expected return on any asset<sup>2</sup> must be the subjective risk-free rate  $\hat{r}$ . Note that the subjective manager's expected return on underlying stock is  $\hat{\mu}$ . Hence we can write the return on arbitrage portfolio  $(u_s, u_\Pi)$  as

$$(8) \quad u_s \hat{\mu}_s + u_\Pi \alpha_\Pi = \hat{r}.$$

---

<sup>2</sup> More precisely,  $\hat{r}$  is the expected return under the subjective pricing rule  $E[d(MZ)] = 0$ , where M is the subjective stochastic discount factor.

In equation (8) the subjective expected return on stock is given by  $\hat{\mu}_s = \mu_s + (1-\gamma)\theta(1-\theta)v^2$ . Using the arbitrage relation (8) and portfolio weights (7), we get the modified Black-Scholes PDE, which we call the Black-Scholes-Ingersoll PDE. The risk-neutral drift in (9) is not simply  $\hat{r}$ , because we want to express the option price as a function of the objective stock price. The diffusion term has a clear interpretation; it is the variance of underlying stock.

$$(9) \quad F_S S(\mu - \hat{\mu} + \hat{r} - q) + F_t + 1/2 F_{SS} S^2 (\beta^2 \sigma_m^2 + v^2) = \hat{r}F.$$

In order to write the option pricing formula more conveniently, we simplify the notation in (9) using the formulae for subjective dividend yield and stock return. The final form for the Black-Scholes-Ingersoll PDE is given by eq. (10).

$$(10) \quad F_S S(\hat{r} - \hat{q}) + F_t + 1/2 F_{SS} S^2 (\beta^2 \sigma_m^2 + v^2) = \hat{r}F$$

where  $\hat{q} = q + (1-\gamma)(1-\theta)\theta v^2$

As shown by equation (10), the risk-neutral drift in options pricing is  $\hat{r} - \hat{q}$ , because the underlying equity is priced using the market rate. The Black-Scholes-Ingersoll pricing formula for subjectively valued call options is given in eq. (11).

$$(11) \quad \hat{C} = S e^{-\hat{q}(T-t)} N(d_1) - K e^{-\hat{r}(T-t)} N(d_2)$$

$$d_1 = \frac{\ln(S/K) + (\hat{r} - \hat{q} + 1/2 \sigma^2)(T-t)}{\sigma \sqrt{T-t}}; \quad d_2 = d_1 - \sigma \sqrt{T-t}.$$

When the B-S-I formula is applied, the resulting employee stock option values are in general lower compared to Black-Scholes values. From a mathematical perspective this follows, because the subjective risk-neutral drift is lower than the objective one. Hence there exists a valuation gap, also known as the *deadweight loss*, measuring how much higher the employer's cost of the option is compared to the employee's perceived value of the option.

### 3. Methodology for estimating the option-implied PDF

#### 3.1 State price density as a valuation tool

We estimate the risk preferences of option-endowed managers. Measuring risk preferences requires the knowledge of managerial expectations, which are spanned by the probability density function implied by option prices. Further, using option prices enables us to characterize the development of

---

preferences in time. As argued by Bliss and Panigirtzoglou (2004), option prices are useful in investigating market expectations, because options are risky assets with fixed expiration dates. This implies that prices of options with different maturities reflect the variation in expected returns over different time periods. In contrast, because stocks are basically perpetual claims, it is very difficult to associate stock price fluctuations with changes in investment opportunities during a certain period.

Our approach to calculating the risk-neutral probability density function (PDF) from option prices is based on estimating implied volatility as a smooth function of option delta. The idea of estimating smooth volatility function was introduced by Malz (1997) and advanced by Bliss and Panigirtzoglou (2002, 2004), among others. In contrast to these papers, we assume that the observed PDF for executive stock options (ESOs) reflects the *subjective risk-free rate*  $\hat{r}$ . This assumption is empirically motivated by the results of Ikäheimo et al. (2006), who find that actual trading prices for ESOs are considerably below Black-Scholes values. Theoretical motivation follows from the Ingersoll (2006) model, where constrained managers adjust risk-free rate downwards. This allows us to determine the preference parameters using option and cash market PDFs. We stress that this approach requires empirical data on ESO prices.

The connection between the stochastic discount factor and the probability density function is well-known. We follow the exposition of Ait-Sahalia and Lo (2000), but the ideas can be traced back to Harrison and Kreps (1979). To see how the stochastic discount factor is connected to the probability density function, we write the first-order condition (Euler equation) for the representative consumer as

$$(12) \quad \begin{aligned} U'(W_T) &= e^{-r(T-t)} U'(W_t) \zeta_T \\ &\Rightarrow \zeta_T = e^{r(T-t)} M_{t,T} \\ \text{where } M_{t,T} &\equiv U'(W_T)/U'(W_t). \end{aligned}$$

The third line of eq. (12) gives the stochastic discount factor  $M_{t,T}$ , also known as marginal rate of substitution (MRS) in microeconomics. Interpretation for the first-order condition is that a certain unit of future consumption is as good as one discounted unit now, times the probability of current state. Note that the idea of risk-neutral pricing is incorporated here by the probability transformation  $\zeta$ . Because option pricing is risk-neutral in general, option prices reflect the risk-free density  $f^*$  instead of the cash market density  $f$ , which is generated by the underlying asset's geometric Brownian motion. In order to relate  $f^*$  and  $f$  we need the Girsanov transformation  $\zeta_T$ . From a structural point of view,  $\zeta_T$  is equal to the intertemporal marginal rate of substitution (up to a constant). From a statistical point of view it is equal to the ratio of risk-neutral and cash market densities, where the cash market density is generated by the underlying asset's stochastic process. The usefulness of (12) becomes

evident below, since it is used to derive the Ait-Sahalia and Lo (2000) nonparametric estimator for relative risk aversion.

As shown in Theorem 3 of Harrison and Kreps (1979), we can calculate the risk-free density, called the *state-price density* (SPD), as

$$(13) \quad \begin{aligned} f_t^*(S_T) &\equiv f_t(S_T)\zeta_T \\ \zeta_T &= \exp\left[-\int_t^T \phi dB - \frac{1}{2}\int_t^T \phi^2 ds\right] \\ \text{where } \phi &= (\mu - r)/\sigma. \end{aligned}$$

The SPD, as defined in (13), is essentially a function of the market price of risk, denoted  $\phi$ . In our view the Black-Scholes-Ingerson model can be seen as a modification of the market price of risk in standard B-S setting. Because the representative manager has to hold an asset that does not lie on the CAPM efficient frontier, his market price of risk is inferior compared to unconstrained investors. From a mathematical point of view, this means applying Girsanov's theorem for a change of measure, where the drift of the underlying process changes from expected risky return to the risk-free rate (for textbook treatment, see section 5.7 of Bingham and Kiesel (2004)). The variance of underlying process remains unchanged. Equation (14) shows change of measure in the B-S-I model with adjusted market price of risk, denoted  $\hat{\phi}$ .  $W^*$  is the subjective risk-neutral Brownian motion and  $W$  is the market Brownian motion.

$$(14) \quad \begin{aligned} W^*(t) &= W(t) + \int_0^t \hat{\phi} ds \\ \hat{\phi} &\equiv (\hat{\mu} - \hat{r})/\sigma \end{aligned}$$

Next we will calculate the expected value of a contingent claim  $Z(S_T)$  to highlight the role of the risk-neutral density  $f^*$ . Equation (15) shows that any asset price can be calculated as the MRS-weighted expectation of outcomes, discounted with the risk-free rate.

$$(15) \quad \begin{aligned} E_t[Z(S_T)\theta_{t,T}] &= \int_0^\infty Z(S_T)M_{t,T} f_t(S_T) dS_T \\ &= e^{-\hat{r}(T-t)} \int_0^\infty Z(S_T) f_t^*(S_T) dS_T \\ &= e^{-\hat{r}(T-t)} E^{\mathcal{Q}^*}[Z(S_T)]. \end{aligned}$$

---

In calculating the integral we can eliminate the stochastic discount factor by inserting equations (12) and (13) in the first line of (15). However, the outcome of this derivation is that observed executive option prices are priced using the subjective risk-free rate  $\hat{r}$ . Ait-Sahalia and Lo (2000) recommend using the state-price density (SPD), because it gives economically meaningful estimates for asset prices. Specifically, the SPD accounts for variation in the value of consumption in different economic states. This is the reason why the SPD should be used in calculating confidence intervals for risk management purposes.

### 3.2 The probability density function implied by option prices

This section shows how we extract the (risk-neutral) probability density function of the underlying asset from option prices. By assuming that security price dynamics follow geometric Brownian motion, Ait-Sahalia and Lo (1998, 2000) find that the PDF is given by eq. (16). Note that  $\tau = T - t$  denotes time to expiration, other notation is as before.

$$(16) \quad f_{BS}^* = e^{r\tau} \left( \frac{\partial^2 C_{BS}}{\partial K^2} \right)_{K=S_T} \\ = \frac{1}{S_T (2\pi\sigma^2\tau)^{1/2}} \exp \left[ \frac{-\left(\log(S_T/S_t) - (r - q - \sigma^2/2)\tau\right)^2}{2\sigma^2\tau} \right]$$

For later use we remark that the derivative of the risk-neutral density function (17) is given by

$$(17) \quad (f_{BS}^*)' = -\frac{1}{S_T^2 \sqrt{2\pi\sigma^2\tau}} \left( 1 + \frac{\log(S_T/S_t) - (r - q - \sigma^2/2)\tau}{\sigma^2\tau} \right) \\ \times \exp \left[ \frac{\left[ \log(S_T/S_t) - (r - q - \sigma^2/2)\tau \right]^2}{2\sigma^2\tau} \right]$$

Equation (16) is the lognormal PDF implied by underlying stock dynamics. Under the Black-Scholes-Ingersoll model the risk-free rate ( $r$ ) and dividend yield ( $q$ ) are replaced by their B-S-I counterparts  $\hat{r}$  and  $\hat{q}$ . The B-S-I density function is a convenient tool for calculating *both* the theoretical and estimated densities. For example, Ait-Sahalia and Lo (1998, 2000) use the Black-Scholes PDF to calculate theoretical density using at-the-money volatility and ‘actual’ density by plugging in a nonparametric volatility function. In standard B-S economy the volatility function is a straight line, whereas in practical markets volatility is explained by a number of factors, for instance by time to expiration and moneyness.

### 3.3 Estimation of the smooth volatility function

We apply the *volatility function method* introduced by Malz (1997) and refined by Bliss and Panigirtzoglou (2002, 2004) to estimate the PDF by mapping option prices in delta-sigma space. The process starts by estimating deltas from option prices using (at-the-money) implied volatility. The next step is to estimate the smooth volatility function, which explains variation in volatility by variation in delta.

Malz (1997) has a rich data on three combinations of forex calls and puts, and he derives an arbitrage relation for their prices. As a result Malz gets a natural smoothing function, which is a second order polynomial of delta. We cannot proceed like Malz and derive similar arbitrage relations, because our data concerns executive stock options with only call features. Therefore we smooth the data in delta-implied volatility space using *nonparametric regression*, using *penalized splines* to smooth the response variable (see chapter 3 of Ruppert, Wand and Carroll (2003)). In nonparametric estimation the objective is to estimate implied volatility as a smooth function of delta. By ‘nonparametric’ we mean that our interest is in the modeling of the response variable. The functional form of response, as well as parameter values are of lesser importance.

Our smooth volatility function  $h_s$  is given by eq. (18). It is estimated using restricted maximum likelihood. The value of  $h_s$  is the estimated volatility at a given level of option delta. The ‘cubic thin plate’ splines, or the third-order polynomials in eq. (18), use radial basis functions<sup>3</sup> of third order. Cubic thin plate splines belong to the class of *linear smoothers*.

$$(18) \quad \hat{\sigma}_s = h_s(x) + \varepsilon$$

$$h_s(x) = E(\hat{\sigma}|x) = \beta_0 + \beta_1 x + \sum_{j=1}^J u_j |x - \kappa_j|^3.$$

In eq. (18)  $\varepsilon$  is the error term with zero mean and constant variance, and  $\{\kappa_j : 1 \leq j \leq J\}$  are the knots located on the  $x$ -axis. In practice we estimate the smooth volatility function by fitting equation (18) in option deltas calculated using the standard B-S formula<sup>4</sup>. For software we use the SemiPar 1.0 package for R language (Wand et al. (2005)).

---

<sup>3</sup> To elaborate, the basis functions used are  $1, x, |x - \kappa_1|^3, \dots, |x - \kappa_J|^3$ . This is called a radial basis, because the splines are radially symmetric about  $\kappa_j$ .

<sup>4</sup> Note that we use the standard Black-Scholes model only for data transformation, i.e. we do not assume that the model were true. Bliss and Panigirtzoglou (2004) perform a similar transformation.

---

We have fitted the volatility function in delta-sigma space, but delta is not an input to the risk-neutral PDF formula introduced by Breeden and Litzenberger (1978) and developed by Ait-Sahalia and Lo (1998). Recall that the PDF is calculated using the second derivative of option price with respect to strike price. Hence we need to calculate strike prices as function of delta. This is done using at-the-money volatility obtained from Datastream. We calculate the implied strike price  $K$  numerically:

$$(19) \quad \Delta = g(S, K, \tau, r, \sigma_{atm}) \Rightarrow K = g^{-1}(\Delta; S, \tau, r, \sigma_{atm})$$

Now that we have a one-on-one correspondence of delta, sigma and strike price, we can calculate the cumulative density function from eq. (16), plugging in volatilities from the smooth volatility function (18) and strike prices from (19). We think that our smooth volatility function strikes a balance between flexibility and data requirements. Compared to the results of Ait-Sahalia and Lo (1998), specifically the ‘implied volatility surface’ of their Figure 3, our type of volatility function captures the moneyness effect, since delta is a good proxy for moneyness.

When the (implied) volatility function is estimated, a key issue is the extent of smoothing. Intuition says that the fitted curve should be ‘smooth’, however there is no common norm as to *how* smooth it should be. The aim of smoothing is to decrease the amount of noise and clarify underlying trends. But the trade-off is that more and more information is lost as the amount of smoothing increases. To see the importance of choosing the smoothing parameters, note that volatility function estimation is not the only instance where we use smoothing. Below we will calculate the Ait-Sahalia–Lo risk aversion estimator, which requires us to compute derivatives of density functions. Unfortunately the derivatives of PDFs tend to be ‘wiggly’, and without smoothing the risk aversion estimates can be quite volatile.

Ruppert, Wand & Carroll (2003) recommend measuring the amount of smoothing with *degrees of freedom* of the fitted linear smoother, instead of using the smoothing parameter  $\lambda$  (for technical details see appendix 2). Nevertheless,  $\lambda$  is used as smoothing measure in a number of studies, e.g. in Bliss and Panigirtzoglou (2002, 2004). The problem with  $\lambda$  is “it does not have a direct interpretation as to how much structure is being imposed in the fit” (see p. 81 of Ruppert, Wand & Carroll (2003)). Further, applying different values  $\lambda$  of can result in quite similar fit, and because of this insensitivity it is difficult to choose an ‘optimal’ value of  $\lambda$ . For example, Bliss and Panigirtzoglou (2004, p.416) try values in the range [0.99, 0.9999] and report that within this range, their results are insensitive to the choice of  $\lambda$ .

Instead, degrees of freedom ( $df_{fit}$ ) has a clear interpretation; it is analogous to the number of parameters in a linear model. To see this, note that the generic spline model can be written as

$\hat{y} = S_\lambda y$ , where  $S_\lambda$  is called the hat matrix (or smoother matrix). Technical details and properties of the estimators are presented in Appendix 1. When applying linear regression, degrees of freedom is equal to number of parameters, but it can also be calculated as the trace of the hat matrix. However, in the case of nonparametric models degrees of freedom cannot be calculated from the number of parameters, because all the knots lie on the same axis. Hence we need to calculate degrees of freedom using the trace of the hat matrix:

$$(20) \quad df_{fit} \equiv tr(S_\lambda).$$

If we choose  $df_{fit}$  as the smoothing measure, the next question is: what is the optimal value for it? Theoretical answer is given by Ruppert, Wand & Carroll (2003), who recommend using the *Generalized Cross-Validation GCV*, which can be calculated as

$$(21) \quad GCV = \frac{RSS(\lambda)}{\left(1 - n^{-1}tr(S_\lambda)\right)^2} = \frac{RSS(\lambda)}{\left(1 - n^{-1}df_{fit}\right)^2}.$$

In equation (21)  $RSS(\lambda)$  is the residual sum of squares of the spline model and  $n$  is the number of observations. In theory, one should choose the model that minimizes  $GCV$ . However, in practice the amount of smoothing must be determined by the data at hand. In particular, estimated densities need to be smooth enough to have a smooth derivative; otherwise the risk aversion estimates may become unreasonably volatile.

### 3.4 Semiparametric estimator for relative risk aversion

This section presents the Ait-Sahalia and Lo (2000) estimator for relative risk aversion. It has the main benefit of producing the relative risk aversion (RRA) as a function of the observed option and cash market PDFs. Further, using the densities produces a mapping of RRA as a function of the state price and not just a point estimate. In fact the shape of RRA function reveals important properties of implicit preferences. If the RRA function is flat, preferences are mapped by power utility, which is consistent with the Black-Scholes model. Ait-Sahalia and Lo (2000) take this argument further and show that all utility functions implied by the Black-Scholes model correspond to constant relative risk aversion<sup>5</sup>. Hence the shape of RRA function is indicative on of how well the applied pricing model fits to the empirical data.

---

<sup>5</sup> This result applies to an economy with no intermediate consumption, hence the set-up is not exactly the same as in the Black-Scholes-Ingersoll model.

---

Here we need to elaborate our notation. Probability density function is denoted by  $f$ , a hat (^) denotes an estimate and a star (\*) denotes risk neutrality. In practice we estimate the risk-neutral PDF  $\hat{f}^*$  from eq. (16) and the cash market PDF  $\hat{f}$  using a standard kernel regression method<sup>6</sup>. With this information we're able to compute the Ait-Sahalia RRA estimator, denoted  $\hat{\rho}(S_T)$ . Recall from equation (13) that  $\zeta(S_T) = \hat{f}^*/\hat{f} = e^{r(T-t)}\theta_{t,T}$ . Using  $\zeta(S_T)$  and its first derivative  $\zeta'(S_T)$  we can relative risk aversion measure is calculated as

$$(22) \quad \rho(S_T) \equiv -S_T \frac{U''(S_T)}{U'(S_T)} = -S_T \frac{\zeta'(S_T)}{\zeta(S_T)}.$$

This expression leads to equation (23), which is the final form of RRA estimator. It very useful since it does not require any (unobserved) utility parameters. In fact this RRA estimator does not assume any specific form of utility function. While this does not directly imply that empirical estimates are independent of the presumed utility function, Table 5 of Kang and Kim (2006) shows that the presumed utility function has only a minor effect on risk aversion estimates.

$$(23) \quad \hat{\rho}(S_T) = S_T \left( \frac{\hat{f}'}{\hat{f}} - \frac{(\hat{f}^*)'}{\hat{f}^*} \right).$$

Ziegler (2002) shows that risk aversion estimates are quite sensitive to errors in density estimation. In his example a small error in estimated standard deviation leads to a large perturbation in risk aversion estimates. Further, our empirical analysis shows a key requirement in estimating density functions is that they are smooth. The reason is that a small bump in the estimated density turns into a large hole in the derivative, which is needed to calculate risk aversion. Further, in the current setup, smoothness of the density function depends on the smoothness of implied volatility function.

---

<sup>6</sup> We estimate the gaussian kernel using the 'density' function of the R system with default bandwidth.

## 4. Data description, empirical fit of the B-S-I model and risk aversion estimates

### 4.1 Data description

A unique feature of the Finnish market is that executive stock options are publicly traded on the HEX, the Helsinki Stock Exchange. Our data consists of Nokia executive stock options traded on the HEX during the period 2.4.2000 – 30.12.2002. The data spans 7611 trades in Nokia 1999 Stock option plan that expired on 31.12.2004. Average time to expiration is 3.10 years. We will not go into the details of these ESOs, but refer the interested reader to Ikäheimo et al. (2006). Since the underlying Nokia stock is the most traded stock on HEX and listed in NYSE, and ordinary Nokia options have excellent liquidity in Eurex, holders of the ESOs have parallel quotes for calculating appropriate pricing parameters. A key property of this data is that most trades (80%) are done in-the-money, as shown by Figure 1. The median and average values of moneyness ( $S/K$ ) are 1.54 and 1.49. In order to fit the volatility function it is important to have a certain amount of variation in option deltas, which requires that the data spans a wide range of moneyness. Figure 1 shows that our data has large variation in moneyness, even if it is weighted towards in-the-money trades.

We investigate the pricing of Nokia 1999 (issued during respective year) ESO plan, even though the Nokia 1997 ESO plan traded at the same time and the trade data is available. However, during our research period the Nokia 1997 ESO plan was very deep in the money, which is shown by the strike price of 3.227 euros, whereas the Nokia 1999 ESO plan had a strike of 16.89 euros. These strikes are adjusted for subsequent stock splits and euro conversion. Under these market conditions Nokia 1997 ESOs behaved very much like ordinary stock, with deltas converging to unity. In a Black-Scholes-type model, stock-like behavior is caused by the asymptotically decreasing time value as the option becomes deep in the money<sup>7</sup>. The interest rates are estimated using the Euribor interest rate and the Finnish zero-coupon bond yield curve. Using linear interpolation, we obtain all interest rates needed for discounting.

### 4.2 Empirical fit of the Black-Scholes-Ingersoll model

Now we will turn to the task of calibrating the B-S-I model (11) to empirical data. We have no prior knowledge of three model parameters: risk aversion ( $\gamma$ ), required equity holding ( $\theta$ ) and idiosyncratic

---

<sup>7</sup> More formally,  $C/(S - K)$  converges rapidly to 1 as  $S/K$  increases, even at high levels of volatility.

---

risk ( $v$ ). We assume that the manager holds an undiversified portfolio. In terms of  $\theta$ , our first assumption is that the value of company stock is one half of the net portfolio. This is our *case 1* with  $\theta = 0.5$ . The alternatives are *case 2*, where  $\theta = 1$  and *case 3*, where  $\theta = 0.25$ . Below we concentrate on the results of case 1 to show how lack of diversification affects ESO valuation. Note that the case  $\theta = 1$  is less extreme than it first appears. It compares to the case of an entrepreneur whose full financial wealth is tied in his own business.

Before reporting the results we point out a technical detail. From equations (12) and (13) we know how to calculate subjective risk-free rate  $\hat{r}$  and subjective dividend yield  $\hat{q}$ . In order to simplify notation we define the *subjective risk premium* as  $\xi \equiv (1-\gamma)v^2$ . Using the subjective risk premium allows us to fit the B-S-I model without *a priori* knowing the idiosyncratic risk and risk aversion parameters. In practice we fit the following B-S-I formulas in ESO price data. Recall that  $\theta$  gives the portfolio restriction, i.e. the minimum holding of the employer's stock.

$$(25) \quad \begin{aligned} \hat{C}_\theta &= \hat{C}(S, K, \tau, \hat{r}, \hat{q}, \sigma_{atm}) \\ \hat{r} &= r - \theta^2 \xi \\ \hat{q} &= q + \theta(1-\theta)\xi \end{aligned}$$

We calibrate the B-S-I formula using at-the-money volatility  $\sigma_{atm}$  obtained from Datastream. The average (unadjusted) risk-free rate is 4.5%. Dividend yield varies between 0.7% and 2.7%, the average being 1.2%.

Our empirical results show that a considerable the subjective risk premium (SRP), caused by risk aversion and idiosyncratic risk, can be found in ESO prices. Table 1 reports descriptive statistics in cases 1–3. In reporting the results we will concentrate in case 1, since cases 2 and 3 yield qualitatively similar results, which can be checked from Table 1. Median value of the SRP in our sample is 8.4%. Using median pricing parameters, it corresponds to a 13% discount to Black-Scholes price. This is in line with the findings of Ikäheimo et al. (2006) suggesting a major underpricing of Finnish ESOs relative to their standard Black-Scholes values.

The subjective risk premium (SRP) explains the difference between the unconstrained market values and constrained subjective values. As shown in Table 1, in some cases the SRP may be higher than the risk-free rate. However, we're not saying that *ex ante* expected return under the subjective risk-neutral measure is negative. We argue only that the *ex post* adjusted risk-free rate may be negative in some cases. Moreover, it is entirely possible that the SRP decreases in time, as managers became more familiar with option pricing, especially with the concept of time value.

Next we investigate how the SRP depends on moneyness of the option and sort the data in S/K quartiles. Looking at case 1, pairwise t-tests (assuming unequal variances) reject the equality of means for all six pairs. We also tested the hypothesis that the SRP comes from the same continuous distribution in all quartiles using a non-parametric test. The Wilcoxon rank-sum test rejected the null for five of six pairs (with the exception of Q2 and Q4 in cases 2 and 3), adding to the evidence that the SRP varies with moneyness. Thus the main conclusion from Table 1 is that the subjective risk premium clearly decreases as moneyness of the ESO increases.

The box-and-whisker plots of Figure 2 show how the SRP is distributed in S/K quartiles in case 1. The notches around median give approximate 95 % confidence bounds. In Q1 and Q2 the interquartile range is almost symmetric, but in Q3 and Q4 the second quartile (i.e. distance between 25% and 50% breakpoints) is quite narrow. This suggests that the majority of observations yield 'reasonable' values for the SRP, but there are a few trades done at unusual prices, and in our case these outliers create upward bias in the mean SRP estimates. In summary, Figure 2 confirms that the SRP decreases with moneyness.

### *4.3 Volatility skew, risk aversion estimates and model fit*

Since our data concerns ESO prices and most trades are done in-the-money, the shape of volatility function may be different from the familiar volatility 'smile' or 'skew'. Derman (1999) gives a practitioner's view of S&P 500 index options' volatility smile and concludes that 'at any time, implied volatility increases monotonically as the strike level decreases'. This would indicate a positive relation of implied volatility and delta. However, the market for executive stock options is different than the index option market.

Figure 3A shows that the implied volatility decreases with option delta. A potential explanation is that in any Black-Scholes-type model, including the B-S-I model, the sensitivity of option price to volatility becomes quite small for high values of moneyness. A similar remark applies to adjusting the risk-free rate in the B-S-I model: for high values of moneyness, the adjustment does not have a material impact in option price. Hence, if prices of in-the-money options (with delta close to 1) exhibit little time value, their implied volatilities are likely to be low. Then the implied volatility function will be decreasing in delta. Figure 3B shows how the moneyness of an ESO is associated with time value and implied volatility. On the upper right corner one can find low moneyness quartile (Q1) observations, to be associated with high time value and high implied volatility. In contrast, from the lower left corner of Fig. 3B one can infer that high moneyness (Q4) observations associate with low time value and low

---

implied volatility. Clearly, the clustering of observations shows that deep-in-the-money options are treated like common stock in subjective valuation.

We estimate the risk aversion of Nokia ESO holders using the Ait-Sahalia and Lo estimator (24). Note that we use the standard Black-Scholes model in this section using market parameters. As described in section 4.3, we estimate the smooth volatility function in delta-sigma space using cubic thin plate splines. Optimal amount of smoothing is determined by choosing the degrees of freedom ( $df_{fit}$ ) of the smoother matrix  $S_\lambda$  defined by eq. (21). Moreover, we choose  $df_{fit}$  so as to minimize the generalized cross-validation (GCV) criterion (22). In practice this means choosing  $df_{fit}=20$ , which results at GCV value of 6.92 and smoothing parameter  $\lambda$  of 0.7605. We tried  $df_{fit}$  values between 15 and 30, however it seems that here the choice of  $df_{fit}$  does not have a big impact on GCV. The smooth IV function is plotted in Figure 3A over the scatter plot of observations. The estimated curve seems to pick up nuances of our data, yet it is smooth enough. Figure 3B shows that implied volatility increases with time value of the option, a property that arises from the B-S-I option pricing model.

Our risk aversion estimator requires the calculation of risk-neutral and cash market density functions, as well as their derivatives. We estimate the risk-neutral density by plugging in the implied volatility function in equation (20). Moreover, we estimate the cash market density from Nokia stock returns using ten year period that ends at the ESO expiration date. The mean daily return for the 10-year period is +0.070%, a figure that represents typical stock market conditions. We estimate the annual cash market density using the stationary bootstrap<sup>8</sup> of Politis and Romano (1994). This method does not assume any parametric distribution for the underlying asset. We simulate with resampling 10,000 return sequences, each containing 250 daily observations. Each sequence produces a single observation for annual return, calculated as the sum of 250 daily log returns. The bootstrap yields a histogram for annual returns. Next, we estimate the cash market PDF is using a Gaussian kernel and bandwidth  $h = 0.785(\hat{q}_{0.75} - \hat{q}_{0.25})N^{-1/5}$ . This formula is obtained from Davidson and MacKinnon (2004, eq. 15.64).

Figure 4 plots the estimated risk-neutral and cash market densities. As expected, the cash market density has a flatter profile. The shape of risk-neutral density is close to lognormal. If the volatility function of Figure 3 was flat, the shape of risk-neutral PDF would be exactly lognormal. In this case relative risk aversion would be constant. Both estimated densities assume a time to expiration of one year, and they are calculated using average dividend yield. The upper panel of Figure 5 plots the

---

<sup>8</sup> We tested the stationarity of daily returns using Augmented Dickey-Fuller regression. Technically the bootstrap procedure was carried out using the `tsbootstrap` function of `tseries` package of R language.

marginal rate of substitution (MRS). Note that the MRS is higher for low stock prices, implying that in those states the manager is unwilling to transfer consumption. The lower panel of Figure 6 shows the relative risk aversion depending on terminal stock price. It is quite interesting that implied risk aversion seems to be slightly higher than one for most of the stock price range. We think this indicates two things. Firstly, constant relative risk aversion (CRRA) may be a reasonable assumption. This is comforting, since the Black-Scholes model is consistent with CRRA preferences, as pointed out by Aït-Sahalia and Lo (2000). Secondly, the data suggests that log utility approximates the average Nokia manager's preferences, and RRA is on average about 1.1.

The validity of our results rests to an extent on the assumption of a *representative manager*. There is some support that one exists. Ziegler (2002) shows that if traders have homogenous beliefs, implied risk aversion at market level is the weighted harmonic mean of individual preferences, and the weights are determined by individual consumption shares. Since our data concerns the managers of a single firm, the assumption of homogenous beliefs, leading to the existence of a representative manager, is not completely unrealistic. In fact, Ziegler's (2002) analysis may help to explain why our risk aversion estimates are relatively low. The level of relative risk aversion decreases, when the proportion of wealthy consumers increases (holding other things constant). Our sample consists of well-off people, so their risk aversion may well be lower than the average investor's. The level of RRA of company managers may be underestimated, since in the markets also other investors may trade with ESOs reducing the influence of company managers on the pricing.

Finally, we comment briefly on the fit of the Black-Scholes model with implied volatility function. Our option pricing model fits actual prices fairly well, and the quality of fit is due to the flexibility of volatility function. Average pricing error is 0.004 euros, with standard deviation of 0.29 euros. In 90 % of cases the pricing error is between  $-0.44$  and  $0.47$  euros (those are the 5<sup>th</sup> and 95<sup>th</sup> percentiles). We present the pricing error graphically in Figure 6. In the left panel, estimated prices are plotted as dots and actual prices drawn by lines, both as function of delta. The conclusion is that while there are some outliers, in general the modified B-S model fits quite well in ESO data. The right panel is the q-q plot that examines the normality of pricing errors. If the pricing error was perfectly normal, q-q plot would be a straight line. The figure shows that the error distribution has heavier tails than normality implies. The tails are associated with the outliers in the left panel.

#### 4.4 Discussion

The risk aversion level of management has implications on corporate financing decisions and optimal compensation schemes. Parrino, Poteshman and Weisbach (2002) use a dynamic model for firm value to show how risk aversion affects the manager's choice of risky projects, assuming that the manager

---

holds some equity and options of the firm. The model uses power utility, so we're able to reflect on their results. Parrino et al. find that as risk aversion increases, the manager's indifference NPV becomes a convex function of firm volatility. In other words, highly risk-averse managers will reject positive NPV projects that a risk-neutral manager would accept, if those projects increase firm volatility.

The empirical evidence showing low level of pay-performance gets a new interpretation based on Haubrich (1994), where RRA of 1.1 would indicate that 1 % profit share the manager would be fair for wealthy manager. This pay-performance relationship is larger than observed by Jensen and Murphy (1990), but far less than the 1.0 predicted by the risk-neutral version of agency theory.

In terms of optimal compensation, there is some theoretical support for using options if we accept that managerial preferences are approximated by log utility. Hemmer, Kim and Verrecchia (2000) use a principal-agent model to show that in the case of log utility, the optimal contract is linear in stock value, and if relative risk aversion is less than one, the optimal contract is convex in stock value. Also Aseff and Santos (2006) work with a principal-agent model and show that a the optimal contract involves a fixed salary and stock options, if one assumes log utility and some regularity conditions for the problem. Based on our data and the theoretical model, we argue that our data supports the use of options in managerial compensation.

## 5. Conclusions

Using actual price data on executive stock options produces a number of insights that have not been presented in the empirical literature. We calibrate the Black-Scholes-Ingersoll model in actual data and find that ESOs incorporate a subjective risk premium. This is equal to saying that constrained managers use a subjective pricing kernel, i.e. they settle for lower return since they are unable to hedge the option position. This follows because it is unlikely that the manager could take a short position in the stock issued by his employer. The existence of subjective pricing kernel opens up limited arbitrage opportunities for unconstrained market makers, who buy the ESOs from managers and hedge their positions in the open market. Limits to arbitrage are set by the limited number of ESOs and competition among market makers.

Our data gives rise to a downward sloping implied volatility function. This unusual shape occurs because prices of in-the-money options incorporate little time value. Semiparametric estimation of managerial preferences, or the marginal rate of substitution, suggests that the value of consumption is higher in states with low stock prices, consistent with concave utility functions. In fact, for a certain interval of stock prices, the relative risk aversion is estimated at just above one, indicating that the

representative manager's preferences could be approximated by logarithmic utility. Further, our results suggest that log utility function, leading to constant relative risk aversion, provides a reasonable fit in the data.

## References

- Aït-Sahalia, Y. and Lo, A.W. (1998) *Nonparametric estimation of state-price densities implicit in financial asset prices*. *Journal of Finance* 53: 499–547.
- Aït-Sahalia, Y. and Lo, A.W. (2000) *Nonparametric risk management and implied risk aversion*. *Journal of Econometrics* 94: 9– 51.
- Aseff, J.G. and Santos, M.S. (2005) *Stock options and managerial optimal contracts*. *Economic Theory* 26: 813 – 837.
- Bingham, N.H. and Kiesel, R. (2004) *Risk-neutral valuation. Pricing and hedging financial derivatives*. Second Edition, Springer.
- Björk, T. (1998) *Arbitrage theory in continuous time*. Oxford University Press.
- Bliss, R.R. and Panigirtzoglou, N. (2002) *Testing the stability of implied probability density functions*. *Journal of Banking and Finance* 26: 381 – 422.
- Bliss, R.R. and Panigirtzoglou, N. (2004) *Option-implied risk aversion estimates*. *Journal of Finance* 59: 407–446.
- Breeden, D.T. and Litzenberger, R.H. (1978) *Prices of state-contingent claims implicit in options prices*. *Journal of Business* 51: 621–651.
- Davidson, R. and MacKinnon, J.G. (2004) *Econometric theory and methods*. Oxford University Press 2004.
- Derman, E. (1999) *Regimes of volatility*. Risk magazine, April 1999. (see also [www.ederman.com](http://www.ederman.com))
- Hall, B.J. and Murphey, K.J. (2000) *Optimal exercise prices for executive stock options*. *American Economic Review* 90: 209-214.
- Hall, B.J. and Murphy, K.J. (2002) *Stock options for undiversified executives*. *Journal of Accounting and Economics*, 33: 3–42.
- Harrison, J.M. and Kreps, D.M. (1979) *Martingales and arbitrage in multiperiod securities markets*. *Journal of Economic Theory* 20: 381–408.
- Haubrich, J.G. (1994) *Risk aversion, performance pay, and the principal-agent problem*. *Journal of Political Economy* 102: 258-276.

- 
- Haubrick, J.G. and Popova, I. (1998) *Executive compensation: a calibration approach*. Economic Theory 12: 561-581.
- Hemmer, T., Kim, O. and Verrecchia, R.E. (2000) *Introducing convexity into optimal compensation contracts*. Journal of Accounting and Economics 28: 307-327.
- Hodder, J.E. and Jackwerth, J.C. (2005) Employee stock options: Much more valuable than you thought. CoFE Discussion paper series, 05.01.
- Ikäheimo, S., Kuosa, N. and Puttonen, V. (2006) *'The true and fair view' of executive stock option valuation*. European Accounting Review 15:3, 351-366.
- Ingersoll, J.E. (2006) *The subjective and objective valuation of incentive stock options*. Journal of Business 79: 453 – 487.
- Kang, B.J. and Kim, T.S. (2006) *Option implied risk preferences: An extension to wider classes of utility functions*. Journal of Financial Markets 9: 180 – 198.
- Malz, A.M. (1997) *Estimating the probability distribution of the future exchange rate from option prices*. Journal of Derivatives, Winter 1997: 18 – 36.
- Merton, R.C. (1969) *Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case*. Review of Economics and Statistics 51.
- Merton, R.C. (1971) *Optimum Consumption and Portfolio Rules in a Continuous-Time Model*. Journal of Economic Theory 3: 373 – 413.
- Merton, R.C. (1990) *Continuous-time finance*. Blackwell Publishers 1990, revised edition 1992.
- Parrino, R., Poteshman, A.M. and Weisbach, M.S. (2005) Measuring investment distortions when risk-averse managers decide whether to undertake risky projects. NBER Working Paper Series, Working Paper 8763.
- Ross, S. (2004) *Compensation, incentives, and the duality of risk aversion and riskiness*. Journal of Finance 59: 207-225.
- Ruppert, D., Wand, M.P. and Carroll, R.J. (2003) *Semiparametric regression*, Cambridge University Press.
- Wand, M.P., Coull, B.A., French, J.L., Ganguli, B., Kammann, E.E., Staudenmayer, J. and Zanobetti, A. (2005). SemiPar 1.0. R package. <http://cran.rproject.org>
- Ziegler, A. (2002) *Why does implied risk aversion smile?* FAME Research Paper 47, available at SSRN. Forthcoming in the Review of Financial Studies.

### Appendix 1: Linear mixed model representation of the penalized splines

In order to show how the cubic thin plate spline parameters are estimated, we write the spline equation in the form of a linear mixed model (A1). In short, the idea of a mixed model is to estimate the response variable as a sum of fixed effects and random effects. In (A7)  $\boldsymbol{\beta}$  is a 2-vector of fixed parameters and  $\mathbf{u}$  is a  $K$ -vector of random parameters. Mixed models are often used in longitudinal studies. In the context of volatility function modeling one could think that fixed effect is caused by delta, and the random effect is caused by repeated measurements. In other words, if we took only one cross section of deltas and volatilities, the random effect would disappear. The idea of presenting splines as linear mixed models is proposed by Ruppert, Wand and Carroll (2003) and we use their notation, in which  $n$  is the number of observations and  $K$  is the number of knots.

$$(A1) \quad \hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{Z}\hat{\mathbf{u}} + \hat{\boldsymbol{\varepsilon}}; \text{ where}$$

$$\boldsymbol{\beta}_{2 \times 1} = (\beta_0, \beta_1)'; \mathbf{u}_{K \times 1} = (u_1, \dots, u_K)';$$

$$\mathbf{X}_{n \times 2} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \text{ and } \mathbf{Z}_{n \times K} = \begin{bmatrix} |x_1 - \kappa_1|^3 & \cdots & |x_1 - \kappa_K|^3 \\ \vdots & \ddots & \vdots \\ |x_n - \kappa_1|^3 & \cdots & |x_n - \kappa_K|^3 \end{bmatrix}.$$

The BLUP (Best linear unbiased predictor) for model parameters can be derived by minimizing the following criterion, which consists of squared errors and a penalty increasing in the smoothing parameter  $\lambda$ :

$$(A2) \quad \min \left[ \frac{1}{\sigma_\varepsilon^2} \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta} - \mathbf{Z}\mathbf{u}\|^2 + \frac{\lambda^2}{\sigma_\varepsilon^2} \|\mathbf{u}\|^2 \right].$$

The fitting criterion shows that without the random component the linear mixed model would reduce to an ordinary linear model leading to the OLS estimator. Minimizing (A3) yields the following estimator  $\mathbf{S}_\lambda$ , which is known as the hat matrix (or smoother matrix):

$$(A3) \quad \hat{\mathbf{y}} = \mathbf{S}_\lambda \mathbf{y}$$

$$\mathbf{S}_\lambda = \mathbf{C}(\mathbf{C}'\mathbf{C} + \lambda^3 \mathbf{D})^{-1} \mathbf{C}'$$

$$\mathbf{C}_{n \times n} = [\mathbf{X} \mathbf{Z}]; \quad \mathbf{D}_{(K+2) \times 1} = \text{diag}(0, 0, 1, \dots, 1); \quad \lambda = \sigma_\varepsilon^2 / \sigma_u^2.$$

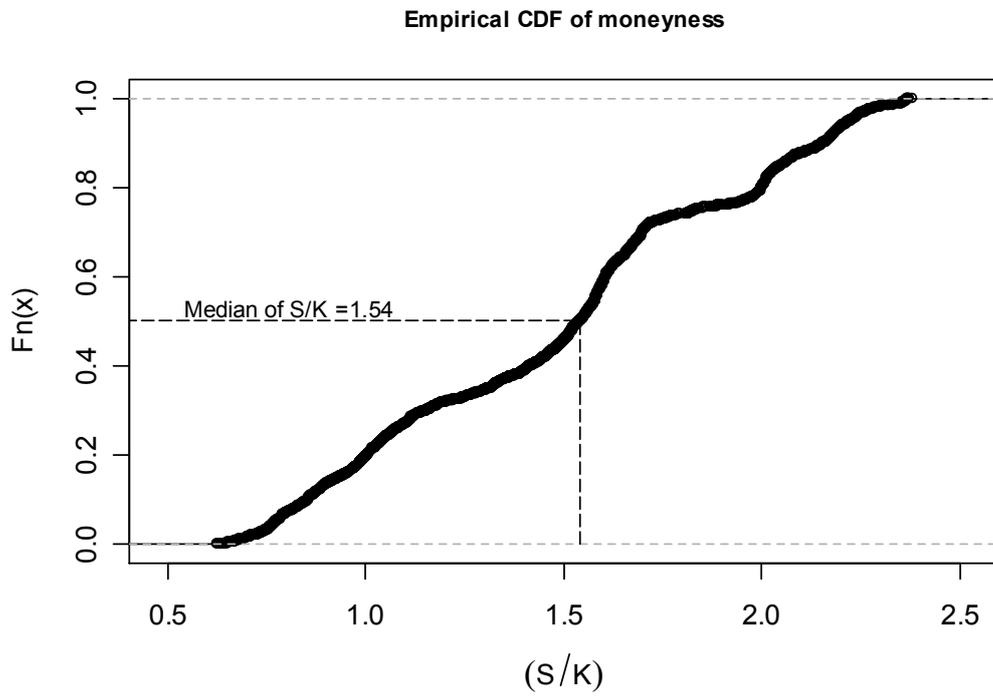
Some authors measure the amount of smoothing by  $\lambda$ . It can be interpreted as the variance ratio of model residuals and random effect. However, Ruppert et al. (2003) strongly argue for measuring the

---

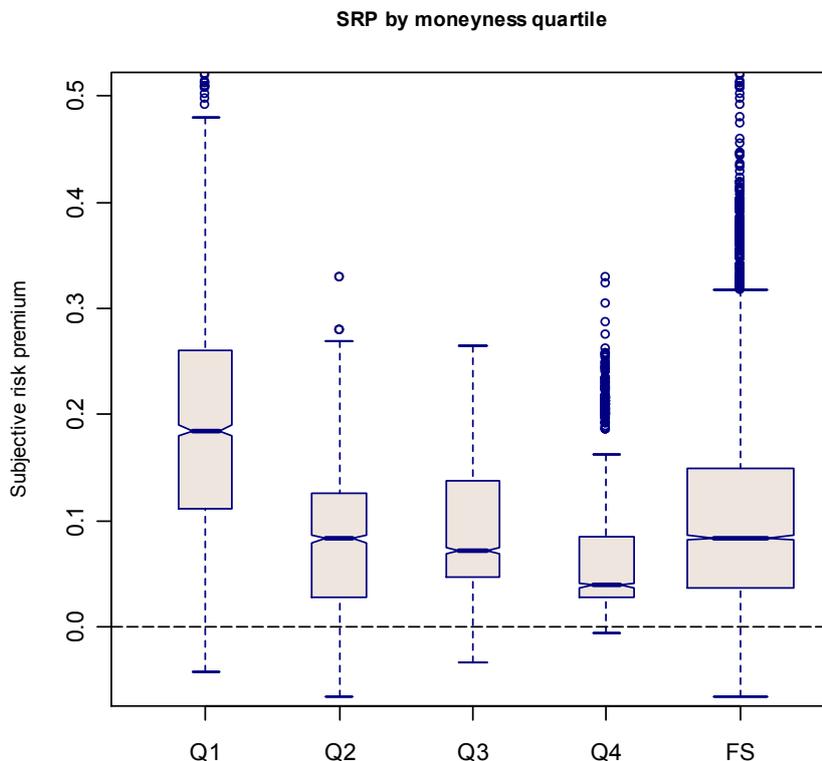
amount of smoothing by the degrees of freedom of the hat matrix (or smoother matrix) defined as  $df_{fit} \equiv tr(S_\lambda)$ .

**Table 1.** Descriptive statistics for the subjective risk premium (SRP). Indices of *t*-statistics stand for pairwise comparison, for example  $t_{12}$  tests the equality of means in moneyness quartiles Q1 and Q2. A star (\*) denotes rejection of the null at 1% significance level. The SRP is defined as  $\xi = (1-\gamma)v^2$  and it is calculated by calibrating the B-S-I option pricing formula (25) to actual prices.

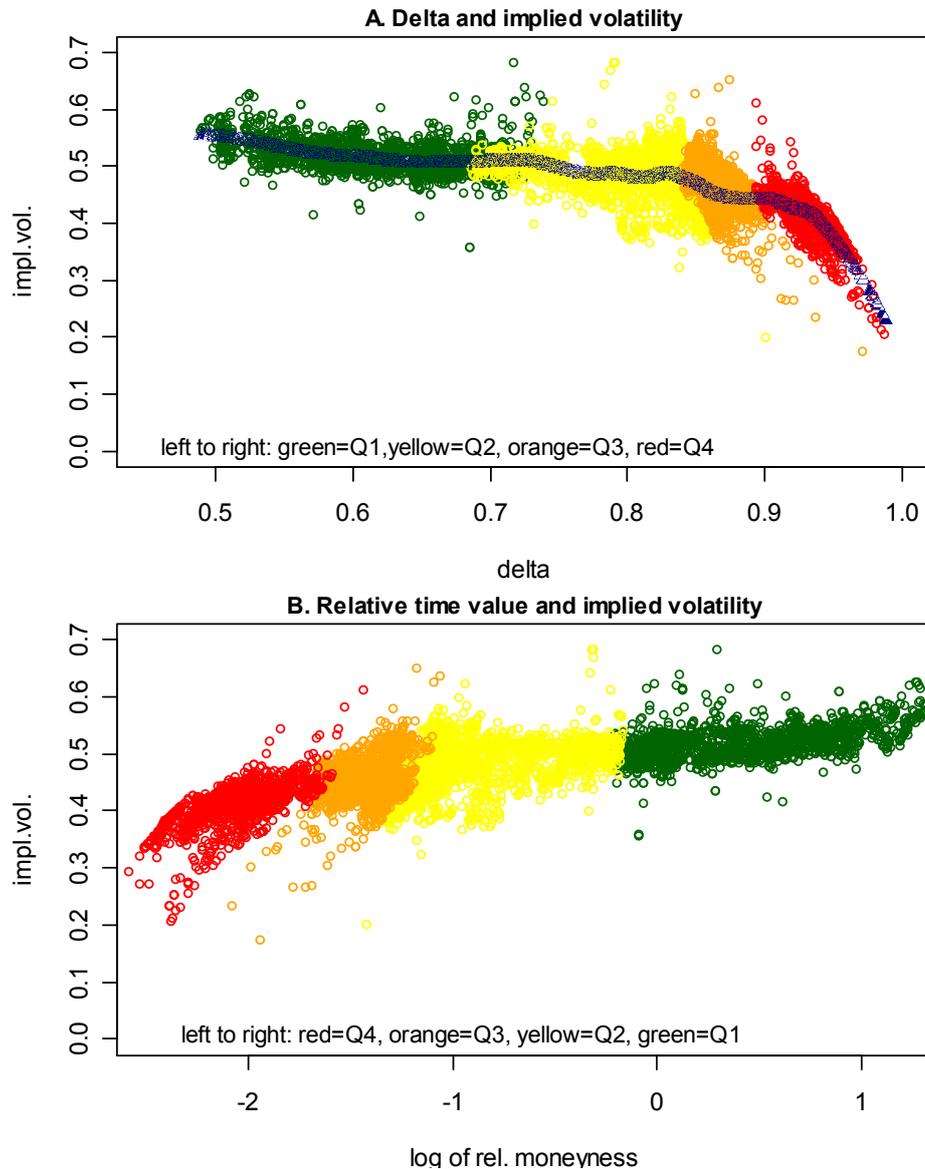
| Moneyness quartile: (range)               | Median | Mean   | St.dev. | t-statistics  |
|---|--------|--------|---------|---|
| <b>Case 1: <math>\theta = 0.5</math></b>  |        |        |         |   |
| Q1: (0.63, 1.06)                          | 0.1852 | 0.1928 | 0.1141  | $t_{12}=36.38^*$ , $t_{13}=36.05^*$ ,<br>$t_{14}=43.62^*$ |
| Q2: (1.06, 1.54)                          | 0.0829 | 0.0827 | 0.0665  | $t_{23}= -2.82^*$ , $t_{24}=8.65^*$                       |
| Q3: (1.54, 1.83)                          | 0.0719 | 0.0882 | 0.0546  | $t_{34}=12.67^*$  |
| Q4: (1.83, 2.38)                          | 0.0391 | 0.0654 | 0.0567  |   |
| Full sample                               | 0.0839 | 0.1072 | 0.0917  |   |
| <b>Case 2: <math>\theta = 0.25</math></b> |        |        |         |   |
| Q1: (0.63, 1.06)                          | 0.3134 | 0.3234 | 0.1882  | $t_{12}=38.54^*$ , $t_{13}=39.40^*$ ,<br>$t_{14}=47.31^*$ |
| Q2: (1.06, 1.54)                          | 0.1336 | 0.1327 | 0.1056  | $t_{23}= -1.62$ , $t_{24}=10.67^*$                        |
| Q3: (1.54, 1.83)                          | 0.1149 | 0.1377 | 0.0824  | $t_{34}=14.06^*$  |
| Q4: (1.83, 2.38)                          | 0.0609 | 0.0998 | 0.0838  |   |
| Full sample                               | 0.1341 | 0.1733 | 0.1510  |   |
| <b>Case 3: <math>\theta = 1</math></b>    |        |        |         |   |
| Q1: (0.63, 1.06)                          | 0.1436 | 0.1542 | 0.0972  | $t_{12}=27.65^*$ , $t_{13}=20.58^*$ ,<br>$t_{14}=24.92^*$ |
| Q2: (1.06, 1.54)                          | 0.0774 | 0.0794 | 0.0669  | $t_{23}= -8.08^*$ , $t_{24}= -0.91$                       |
| Q3: (1.54, 1.83)                          | 0.0703 | 0.0974 | 0.0707  | $t_{34}=6.39^*$   |
| Q4: (1.83, 2.38)                          | 0.0431 | 0.0816 | 0.0817  |   |
| Full sample                               | 0.0807 | 0.1031 | 0.0855  |   |



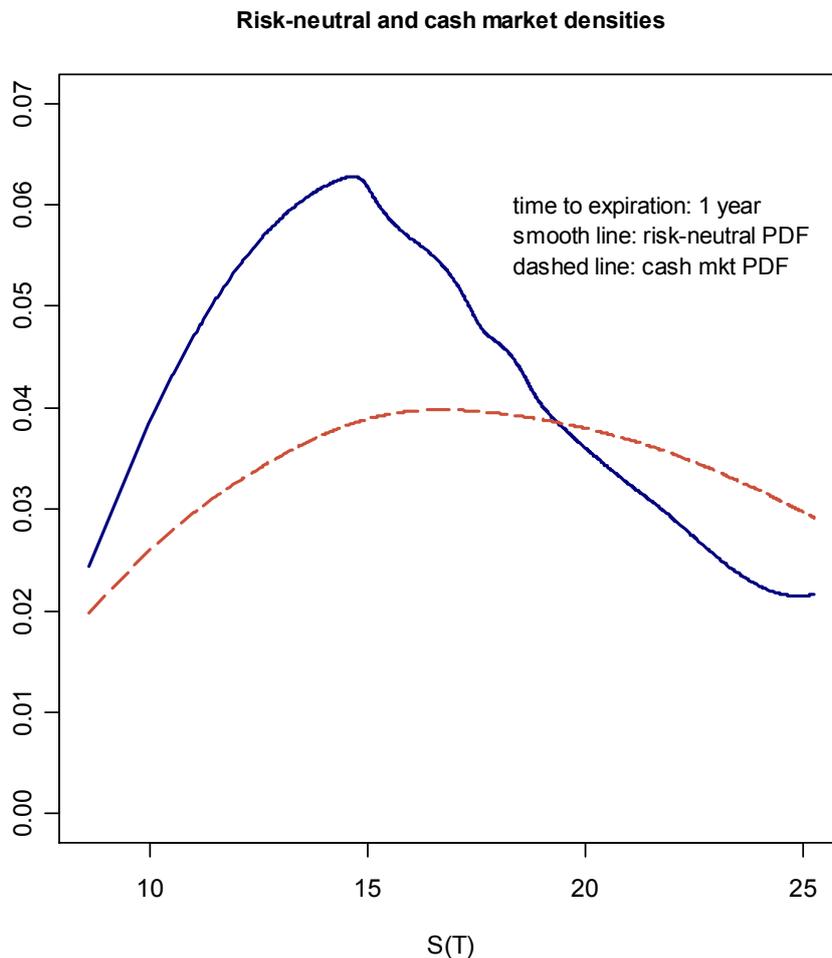
**Figure 1.** Moneyness of ESO trades in our data, plotted using the empirical CDF. The data includes 7610 trades in Nokia's executive stock options, traded on the Helsinki Stock Exchange.



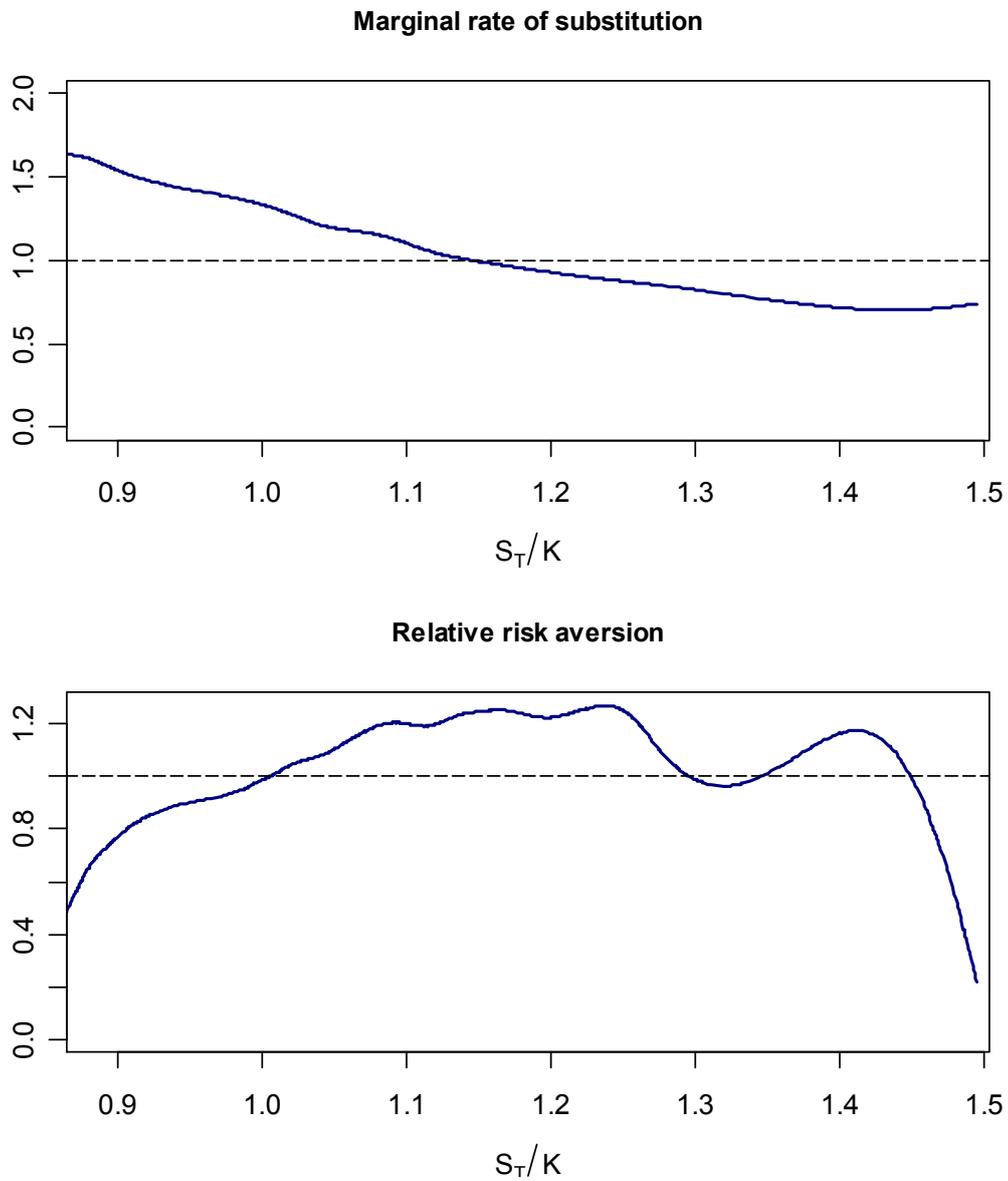
**Figure 2.** Box-and-whisker plots of the subjective risk premium in different moneyness quartiles and full sample (FS), in the case  $\theta = 0.5$ . Thick line in the box plots the median, and the top and bottom of the box are the 25% and 75% breakpoints. Whiskers extend to 1.5 times interquartile range. The “notches” around the median plot approximate 95% confidence bounds. Each quartile has around 1900 observations and full sample size is 7610.



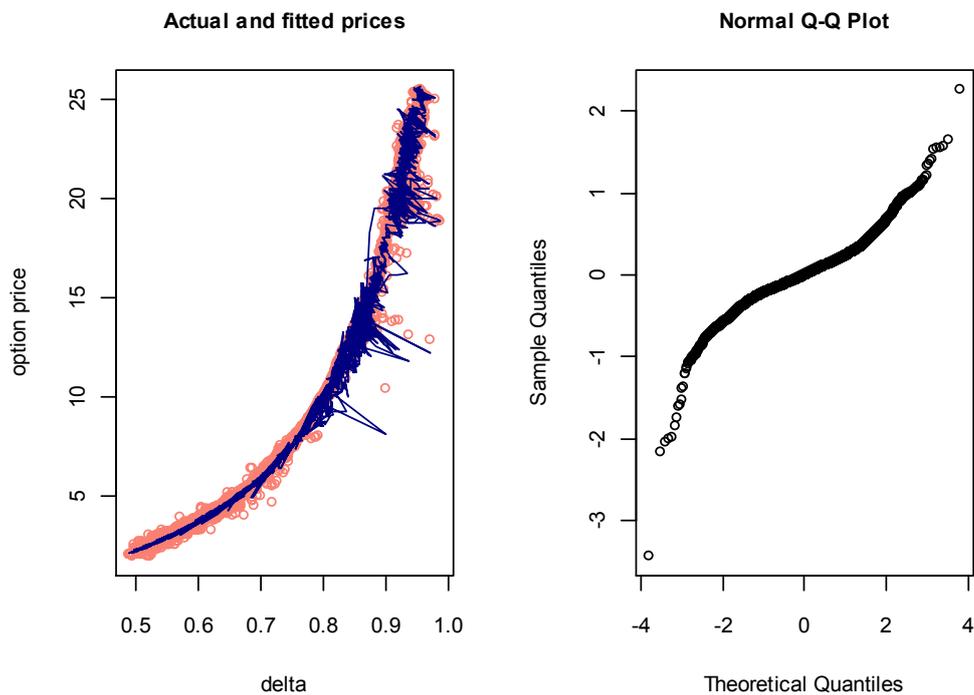
**Figure 3.** Panel A is the scatter plot of delta and implied volatility. Going from left to right, the data is grouped by moneyiness quartiles in increasing order. Clearly, the conclusion is that implied volatility decreases as delta increases. The black line in panel A draws the smooth volatility function. Panel B features the scatter plot of relative time value and implied volatility. Relative time value is calculated as log of time value divided by option price, that is  $\log [(C - (S - K))/C]$ . We use a logarithmic scale in order to make the increasing pattern more visible. The data are plotted in decreasing order of moneyiness.



**Figure 4.** Risk-neutral (smooth line) and cash market (dashed line) density functions. The risk-neutral density is estimated from Nokia ESO trade data which contains 7610 trades in Nokia series 99 ESOs that vested in April 2001 and expired at the end of 2004. The cash market density is estimated using the bootstrap from daily returns of Nokia stock. Both densities have a time frame of one year.



**Figure 5.** The upper panel plots marginal rate of substitution and the lower panel plots relative risk aversion, plotted as a function of moneyiness. The dashed line is drawn at  $RRA=1$ , equivalent to logarithmic utility function.



**Figure 6.** The left panel plots actual ESO prices as dots and estimated option prices as lines. Estimated prices are based on Black-Scholes formula using the smooth volatility function. The right panel is a quantile-quantile plot of (absolute) pricing errors vs. the normal distribution. If the pricing error distribution was exactly normal, the Q-Q plot would be a straight line.