STOCHASTIC DOMINANCE BOUNDS ON OPTION PRICES IN THE PRESENCE OF TRANSACTION COSTS: AN EMPIRICAL APPROACH

Michal Czerwonko

A Thesis
in
the John Molson School of Business

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science in Administration at
Concordia University
Montreal, Quebec, Canada

November 2002

ABSTRACT

Stochastic Dominance Bounds on Option Prices in the Presence of Transaction Costs: An Empirical Approach

Michal Czerwonko

This paper investigates the multi-period upper bound on the European call price in the presence of transaction costs derived by Constantinides-Perrakis (2002). Numerical results verifying an assumption of the monotonictity of wealth of the call writer in the underlying asset on which the Constantinides-Perrakis (2002) model relies are derived, and it is shown that the assumption is satisfied for relatively small ratios of stock to option account. The classic second order stochastic dominance argument is applied to the dynamic trading in discrete time in the S&P 500 options under the portfolio selection criteria in the presence of transaction costs. It is shown that the improvement in expected utility does occur under the prescribed investment policy in the S&P 500 calls whose prices exceed the bound. Under the lognormality of the S&P 500 price process, the quantitative improvement in expected utility is derived.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF FREQUENTLY USED SYMBOLS

μ denotes estimated or true expected return on an index.

σ denotes estimated or true stock volatility.

R denotes the riskless return.

r denotes the riskless rate.

S, S_t denotes the current period stock price.

K denotes the exercise price of an option.

x, x_t denotes the current period dollar value of the holdings in the riskless bond.

y, y_t denotes the current period dollar value before transaction costs of the holdings in the risky asset.

1. Introduction

Constantinides and Perrakis (2002, 2002a) derived stochastic dominance bounds on the prices of European and American contingent claims in the presence of transaction costs for a multi-period economy. In this thesis we exploit empirically their stochastic dominance results with traded call options on S&P 500. Specifically, we examine the violations of the upper bound on the European call price.

To measure the improvement in expected utility, first we compute the optimality conditions of the portfolio of a representative investor before he assumes a position in the derivative. Such conditions were established in the literature of the optimal portfolio selection in the presence of transaction costs.

The model of Constantinides and Perrakis imposes a monotonicity condition involving the minimum number of shares held by a representative investor before the improvement in expected utility can occur. We verify the satisfaction of this condition using simulated data. Last, we apply classic second order stochastic dominance to measure the improvement in expected utility of wealth resulting from trading in the S&P 500 options whose prices violate bounds derived in Constantinides and Perrakis (2002).

The remainder of the thesis is organized as follows: Part 2 reviews the literature of option pricing in the presence of transaction costs, Part 3 describes the methodology, Part 4 presents empirical results, Part 5 summarizes and closes.

2. Literature Review

Transaction costs destroy the "benchmark" Black-Scholes (BS, 1973) option pricing model. The continuously adjusted riskless hedge becomes ruinously expensive no matter how small transaction costs might be because of the infinite variation of diffusion processes (Leland, 1985, Merton, 1989). In this review we present several attempts to incorporate transactions costs into the option-pricing model by considering discretized versions of that model relying on arbitrage such as Leland (1985) Boyle and Vorst (1992). In a similar vein, the notion of option super-replication will be introduced by reviewing Bensaid, Lesne, Pagès and Scheinkman, (1992) and Perrakis and Lefoll (1997). The arbitrage-based approaches suffer from a major theoretical drawback: the presence of transaction costs establishes a region of the risky to riskless asset proportion (the no transaction region) within which it is optimal to refrain from trading (Magill and Constantinides, 1976, Constantinides, 1979). The notion of super-replication was introduced to model such a region, but the solution that it provides is not empirically meaningful.

The expected utility approach introduced the optimal portfolio selection problem in the presence of transaction costs into the pricing of derivatives. Here we review contributions of Davis, Panas and Zariphopoulou (1993) and Constantinides and Zariphopoulou (1999, 2001) from the area of option pricing under the expected utility approach, and those of Constantinides (1986), Norman and Davis (1990) and Dumas and Luciano (1991) related to optimal portfolio selection in the presence of transaction costs.

As our focus is on the contribution of Constantinides and Perrakis (2002, 2002a), we relate their results to previous work of Perrakis (1986), Perrakis and Ryan (1984), Ritchken and Kuo (1988), Ritchken (1985) and Levy (1985), whose results are similar to those of Constantinides-Perrakis (2002). Finally, note that the area of derivative pricing in the presence of transaction costs also encompasses the pricing of American options, which is omitted in this review.

Leland (1985) considered the stock process over a small (but noninfinitesmal) time interval Δt. Under the assumption that $\Delta S/S$ is normally distributed with zero mean it follows that:

$$E\left|\frac{\Delta S}{S}\right| = \sigma\sqrt{2\Delta t/\pi}\,. \quad (2.1)$$

Leland (1985) defines a modified variance that accounts for transactions cost at each revision opportunity:

$$\begin{aligned}\hat{\sigma}^2 &= \sigma^2\left[1 + 2k\,E\left|\frac{\Delta S}{S}\right| \Big/ \sigma^2\Delta t\right] \quad (2.2)\\ &= \sigma^2[1 + 2k\sqrt{2/\pi}\,/\,\sigma\Delta t].\end{aligned}$$

Define $\hat{C}$ as the Black-Scholes call price with this modified volatility. The modified self-financing replicating strategy of the Black-Scholes type with the modified volatility will, for small Δt, yield an expected terminal value equal to the call payoff *inclusive* of transaction costs; furthermore, the hedging errors (including transaction costs) will almost surely approach zero as Δt tends to zero.

$\hat{C}$ is also the upper bound on the call price, as arbitrage profits can be made by writing the option and buying the portfolio replicating the long call[1]. The lower bound (the short call) can be similarly obtained by reducing the variance by the same amount of $\sigma 2k\sqrt{2/\pi}/\Delta t$. These bounds are not a satisfactory result as they create a Catch 22: if the revision interval tends to zero, the variance explodes resulting in trivial bounds equal to the current stock price and the Merton (1973) lower bound $[S - Ke^{-rT}]^+$; if the time interval is large, the hedging errors nullify the riskless arbitrage oportunities.

Boyle and Vorst (1992) incorporated transactions cost into the binomial model. They extended the results of Merton (1989), who had provided the solution for the two-period tree, into the multi-period setting. Without transaction costs, a self-financing portfolio hedging a long call in one period will become NSu + RB (NSd +RB) if the stock goes up (down). In the presence of transaction costs, the portfolio value has to include the transactions cost of restructuring $k|N-N_1|Su$ and $k|N-N_2|Sd$ (assuming $k = k_1 = k_2$). As it can be shown that $N_2 \leq N \leq N_1$, we get:

$$\begin{aligned} NSu(1+k) + BR &= N_1 Su(1+k) + B_1, \\ NSd(1-k) + BR &= N_2 Sd(1-k) + B_2. \end{aligned} \qquad (2.3)$$

This system of equations has always a feasible solution, and the portfolio at the origin of the stock process can be found recursively starting from the call payoff at the expiration. For Su > K and Sd < K we have:

$$\begin{aligned} NSu(1+k) + BR &= Su(1+k) - K, \\ NSd(1-k) + BR &= 0. \end{aligned} \qquad (2.4)$$

[1] The riskless hedge of the BS model has to be adjusted to the quantity of stock equal to the delta of $\hat{C}$.

For the short call replication the system will become:

$$\begin{aligned} NSu(1-k)+BR &= N_1Su(1-k)+B_1\,, \\ NSd(1+k)+BR &= N_2Sd(1+k)+B_2\,. \end{aligned} \quad (2.5)$$

The solution to this system is now feasible only by restricting the parameters of the binomial process:

$$R(1+k) \le u(1-k), \quad d(1+k) \le R(1-k)\,. \quad (2.6)$$

Unfortunately, for a large number of binomial steps, u, d and R all tend towards 1 whereas k remains unchanged. Eventually, for a sufficiently dense partition, the restrictions (2.6) will be violated. In addition, the terminal stock value must not fall into the interval [K/(1+k), K/(1-k)]; otherwise, the optimal action of the option holder cannot be determined.

Boyle and Vorst (1992) also found the limiting form of the value of the replicating portfolios. They found that this limit was equal to an expected option payoff under a risk-neutral distribution tending to a lognormal diffusion with an adjusted variance, as in Leland (1985).

$$\hat{\sigma}^2 = \sigma^2[1+2k/\sigma\Delta t]\,. \quad (2.7)$$

The Leland adjustment in (2.2) was, however, smaller than that of Boyle and Vorst, by a factor $\sqrt{2/\pi}$, or app. 0.8. The Boyle-Vorst adjustment results in a slightly higher upper bound than Leland's (1985). The lower bound, where the variance has to be reduced by the same amount of $2k\sigma/\Delta t$ is not reliable due to the fact that the model restrictions (2.6) become violated before the process approaches diffusion. On the other hand, the

Boyle and Vorst (1992) model has the advantage over the Leland (1985) model that it produces arbitrage profits *provided* the binomial model restrictions are satisfied.

Bensaid, Lesne, Pagès and Scheinkman (BLPS, 1992) introduced the notion of option super-replication, i.e. they formulated a dynamic programming problem with the objective of minimizing the initial cash position necessary to cover the restructurings of the portfolio at each node of the binomial tree and the terminal option payoff. They also introduced an economically important distinction between the cash-settled[2] and physical delivery options.

Let (N, B) denote a portfolio with N shares of the underlying and B$ of the riskless asset. At the expiration date T the portfolio replicating the long call position must hold (1, -K) if $S_T>K$, (0, 0) otherwise for physical delivery options, and $(0, [S_T - K]^+)$ for cash settled options[3]. The final objective is to minimize the portfolio value at the origin of the binomial tree. The BLPS program minimizes recursively the cash value including the transaction costs of restructuring at each node, potentially creating a path dependence of the final solution since the number of shares at each node depends on the inherited amount of stock. The path dependence is, however, avoided as a consequence of specifying terminal conditions depending on the delivery terms. Like Boyle and Vorst (1992), BLPS failed to provide a solution when the terminal stock price falls into the interval [K/(1+k), K/(1-k)], and for the case when conditions (2.6) are violated[4].

[2] This is the case of index options.
[3] For the short position, the signs assume the reciprocal values.
[4] They assumed that the exercise policy at expiration is a predetermined function of the terminal stock price and ignored the dependence of the optimal exercise policy on the holder's objectives.

Perrakis and Lefoll (1997) provided an algorithm deriving a portfolio that hedges the short option position perfectly when either S_T lies within the interval [K/(1+k), K/(1-k)] or the conditions (2.6) are violated. In the case of short options, Perrakis and Lefoll (1997) showed that the optimal hedging portfolio contains either 0 or (-1, K/R) positions, depending on whether (S – K/R) is less than or grater than zero. This rather trivial solution is inevitable when the size of the binomial tree becomes large.

Several papers have undertaken the problem of continuous-time super-replication of European contingent claims in incomplete markets[5]. The results are, however, not satisfactory: a portfolio containing one share of stock turns out to be, in the presence of transaction costs, the only solution dominating the call payoff. This result was first conjectured by Davis and Clark (1994) and later proved analytically in Soner, Shreve and Cvitanić (1995). Cvitanić and Karatzas (1996) derived a stochastic control representation of the problem.

We now turn to the portfolio selection problem under transaction costs. Constantinides (1986) derived a closed-form solution for an infinite horizon investment problem with proportional transaction costs for the two-asset portfolio consisting of the riskless bond and a risky asset with a natural interpretation of a mutual fund account[6]. The Constantinides (1986) result was based on earlier findings of Magill and Constantinides (1976) and Constantinides (1979). The former had conjectured, and the latter proved that the no transaction (hereafter NT) region is a cone, and the optimal investment policy is

[5] See Cvitanić, Pharm, Touzi, (1999) for a non-technical summary.
[6] The solution for a two-asset problem *without* transaction costs was found by Merton (1969, 1971).

simple in the presence of transaction costs. The NT is determined by two boundaries $\bar{\lambda}$ and $\underline{\lambda}$, $\bar{\lambda} > \underline{\lambda}$ denoting proportions of y/x such that for $\bar{\lambda} \geq y/x \geq \underline{\lambda}$ it is optimal to refrain from transacting. The simple investment policy stipulates transacting to the closest boundary of the NT region for $y/x > \bar{\lambda}$ or $y/x < \underline{\lambda}$. Constantinides (1986) made a simplifying assumption regarding the optimal consumption policy, by stipulating that consumption comes from the bound account at a constant rate β, and presented an argument that the loss of derived utility due to this assumption is small. Under his assumption, in the NT region the account dynamics follow:

$$\begin{aligned} dx &= rxdt - cdt = rxdt - \beta xdt \\ dy &= \mu dt + \sigma d\omega , \end{aligned} \qquad (2.8)$$

where ω denotes the Wiener process.

Constantinides (1986) solved a dynamic control problem of maximizing the derived (power) utility with the boundary conditions established by the simple investment policy:

$$V[x, y; \beta, \underline{\lambda}, \bar{\lambda}] \equiv \max E_0 \int_0^\infty e^{-\rho t} u(c(t))dt = \max E_0 \int_0^\infty e^{-\rho t} \gamma^{-1} c^\gamma (t)dt , \quad (2.9)$$

where c(t) denotes consumption at t, ρ is the impatience (time discount) factor, and E_0 is the current time expectation over the Wiener process ω (the time subscripts have been suppressed for brevity). The derived utility satisfies the Bellman equation[7]. Here we present several properties of the NT region derived by Constantinides (1986):

A. Transaction cost increases broaden the NT region, with a decreasing sensitivity of the bounds to the transaction cost rate as the rate increases.

[7] The detailed solution follows in Part 3.

B. Transaction cost increases shift the region towards the risk-free asset.

C. The relative width of the NT region is insensitive to the risk aversion and to the variance of the rate of return whereas the absolute width increases in the variance, and decreases in the risk aversion.

D. Transaction costs weakly decrease the consumption rate.

Norman and Davis (ND, 1990) considered transaction costs directly entering into the dynamic optimization problem[8]. Define auxiliary variables L(t) and U(t) as, respectively, the cumulative purchases in [0,t] (U(0), L(0) = 0). Setting dL(t) = l(t)dt and dU(t) = udt and suppressing time subscripts, the portfolio dynamics becomes:

$$\begin{aligned} dx &= xrdt - cdt - (1+k_1)ldt + (1-k_2)udt, \\ dy &= y\mu dt + \sigma y d\omega + ldt - udt. \end{aligned} \tag{2.10}$$

The derived utility to be maximized is $E_0\int_0^\infty e^{-\rho t}u(c(t))dt$. ND solved the dynamic optimization problem for power and logarithmic utility functions. Given the initial endowment (x, y), the triplet maximizing the value function is (c^*, L^*, U^*). The impatience factor has to satisfy $\rho > \gamma[r + (\mu - r)^2/2\sigma^2(1-\gamma)]$; otherwise unbounded growth of discounted utility is possible. ND assumed $r < \mu < r + (1-\gamma)\sigma^2$, stipulating that holding long positions in both assets is optimal rather than shortselling or borrowing. ND reduced the problem to one variable by normalizing the y value to one, in which representation the stochastic process of y is passed onto x. This normalization makes it possible to deal with a redefined value function $\Psi(x) = V(x, 1)$[9].

[8] In the Constantinides (1986) formulation they appear only in the boundary conditions.

[9] By the homogeneity of V we have $V(x, y) = y^\gamma\Psi(x/y)$

Having proved that the NT region is a wedge in x and y coordinates, ND argue that it suffices to find (x_1, x_2) satisfying $y/x_1 = \bar{\lambda}$ and $y/x_2 = \underline{\lambda}$, given that y = 1, to fully determine the NT region[10]. ND did not find a closed form solution for (x_1, x_2); deriving the NT region involves a numerical solution of a set of two ordinary differential equations in unknown auxiliary functions f(x) and h(x). The boundary conditions of this set make it possible to derive x_1 and x_2. Having y fixed at one, it establishes the NT region.

The optimal consumption policy derived by ND for power utility takes the form:

$$c^*(x, y) = \frac{\gamma x h(x/y)}{(1-\gamma) f(x/y)}. \quad (2.11)$$

Numerical results derived by ND suggest that the optimal consumption rate may vary over a range of nearly two to one within the NT region; however, while comparing this result to Constantinides (1986) ND do not quantify how much of derived utility is lost by the simple consumption policy. The shape of the NT region as a function of the transaction cost rate found by ND shows that their results are not qualitatively different from those of Constantinides (1986).

Dumas-Luciano (1991) considered a dynamic portfolio choice under transaction costs of an investor who maximizes the derived utility of consumption taking place upon the liquidation of the portfolio holdings at some future time T. This assumption results in just two controls of the dynamic programming problem: $(\bar{\lambda}, \underline{\lambda})$, the upper and the lower

[10] The two rays forming the NT region pass through the origin of the (x, y) plane.

bound of the NT region. Dumas-Luciano (1991) considered a limiting case as the liquidation time T tends to infinity. They assumed the discount factor to be endogenous to the problem, i.e. they solved for the discount factor for which the partial derivative of the value function w. r. to time is zero. Dumas-Luciano (1991) results differed from those of Constantinides (1986) first, in that the NT region was found to be considerably wider; second, no shift towards the riskfree asset was found for increases in the transaction cost rate. The latter result Dumas-Luciano (1991) attributed to consumption not taking place[11] before the liquidation time T.

Zariphopoulou (2000) presented a non-technical summary of the expected utility method. The method is based in some sense on classical principles of stochastic dominance adapted to accommodate dynamic trading. Define the value function V of a holder of bond and stock accounts and the value function J of an owner of similar holdings save for writing one European derivative with maturity at T and payoff $g(S_T)$:

$$V[x, y, S, t] = E_0 \int_t^{\infty} e^{-\rho(T-t)} u(c(s)) ds ,$$

$$J[x, y, S, t] = E_0 \left[\int_t^T e^{-\rho t} u(c(s)) ds + e^{-\rho t} V(x_T - g(S_T), y_T, S_T, T) \right] . \quad (2.12)$$

The reservation state-dependent write price $\tilde{h}$ is defined as the quantity making the writer indifferent between receiving the price and writing the derivative, or not undertaking any action at all. For all states, $\tilde{h}$ satisfies:

$$V(x, y, S, t) = J(x + \tilde{h}(x, y, S, t), y, S, t) . \quad (2.13)$$

[11] Constantinides (1986) assumed consumption taking place from the bond account at a constant rate.

In frictionless markets, $\tilde{h}$ can be found independently from wealth, and it coincides with the Black-Scholes price. In the presence of frictions, as shown in Constantinides and Zariphopoulou (1999), (2.13) cannot hold for all states if the dependence of $\tilde{h}$ is removed. This leads to the definition of a universal write price $\tilde{C}$:

$$V(x, y, S, t) \leq J(x + \tilde{C}(S,t), y, S, t). \quad (2.14)$$

By a similar argument, Zariphoupoulou derives an expression defining the reservation purchase price $\underset{\sim}{C}$:

$$V(x, y, S, t) \leq J(x - \underset{\sim}{C}(S,t), y, S, t). \quad (2.15)$$

Before proceeding further with the derivative pricing under the expected utility approach, let's mention its important limitation regarding the number of modeled risky assets. As Magill and Constantinides (1976) pointed out, m risky assets imply 3^m transaction (Buy, Sell, NT) regions whereas there are no models establishing the portfolio optimality condition for more than one[12]. This restriction confines the method to assets interpreted as the market portfolio or an index as in Davis, Panas and Zariphopoulou (1993) and Constantinides and Zariphopoulou (1999). On the other hand, Constantinides and Zariphopoulou (2001) circumvented the restriction and presented results for multiple risky securities, but at the cost of restrictions on preferences and portfolio holdings.

Davis, Panas and Zariphopoulou (1993) considered the problem of pricing long derivatives in the presence of transaction costs by applying a utility maximization

[12] Davis and Norman (1990) mentioned that it might be feasible though difficult to obtain results for two or three risky assets.

approach within the framework developed by Norman and Davis (1990), but for exponential rather than power utility. In their approach the Buy, Sell and NT regions determine the optimal stock transactions of the writer. ` Q_j can be obtained by solving an appropriate partial differential equation numerically. The call write price derived by Davis at al (1993) clearly is a function of both risk aversion and the initial wealth of the writer, as it linearly varies with the risk-aversion coefficient, and the initial wealth enters (2.16) through the value function. Numerical results were presented for a given risk aversion coefficient γ.

A different approach than in Davis at al (1993) was applied by Constantinides and Zariphopoulou (1999). Their objective was to find write prices that depend only on S and t, the common variables to all investors. Given the objective, the call write price will be the minimum price satisfying:

$$J(x+\overline{C}(S,T,t),y,t) \geq V(x,y,t), \quad (2.17)$$

for a sufficiently general class of utility functions. Constantinides and Zariphopoulou (1999) did not assume any specific form of the utility function apart from some regularity assumptions. Since the shape of the NT region cannot be found under such general assumptions, Constantinides and Zariphopoulou (1999) assumed the NT region to be a convex subset of the non-negative quadrant of the (x, y) plane. The derived upper bound does not depend on the investors' initial portfolio but does depend on the preferences. Numerical results were not presented.

Constantinides and Zariphopoulou (2001) examined an investment environment with several risky assets. Contingent claims all expiring at the same time T are allowed to be American, exotic or path dependent. The portfolios considered by Constantinides and Zariphopoulou (2001) may contain riskless asset, stocks and derivatives. A portfolio initially worth \$1 has been constrained at time zero to hold only the riskless asset. The portfolio payoff h(τ, T) has been constrained to be non-negative at any exogenously given liquidation time τ. Further restrictions apply to the power utility function: the risk aversion coefficient, γ has to lie in an open interval (0, 1). Under the above set of assumptions, the J-value function will exceed the V-value function at time zero unless the following relation holds:

$$E\{e^{-\rho\tau}[h(\tau,T)]^{\gamma} - 1\} \leq 0\,, \quad (2.18)$$

where ρ is the impatience factor.

We illustrate the result (2.18) by the following lower bound on the price of a European put:

$$\underline{P} \geq e^{-\rho T} E\{[(K - S_T)^{+}]^{\gamma} \mid S_0\}^{1/\gamma}\ . \quad (2.19)$$

Before presenting the Constantinides-Perrakis (CP, 2002) results that form the basis of our own study, we review briefly the early results on option bound in incomplete markets that are closely related to the CP methodology. In the absence of transaction costs the CP results coincide with the results of these earlier studies, for which three different approaches have been presented.

Perrakis and Ryan (1984) derived one-period discrete-time option bounds later extended to multi-period setting in Perrakis[13] (1986). The derived bounds are functions of the stock and exercise prices, the riskless rate of interest, the time to maturity and the entire distribution of stock returns. The derivation uses a set of Rubinstein (1976) assumptions: the single-price law of markets, frictionless, Pareto-efficient financial markets, rational, time additive tastes, weak aggregation. An assumption specific to Perrakis and Ryan (1984) stipulates that the conditional mean marginal utility is non-increasing in the one-period stock price change Y. The derivation of the bounds compares the terminal payoffs of three portfolios consisting at t of: i. One share of stock at price S_t, ii. One call and S_t-C in riskless bond, iii. S_t/C call options. The relative dominance of expected payoffs of the aforementioned portfolios establishes the lower and upper bound on the call option:

$$\overline{C}(S_t, K) = S_t + \frac{S_t}{S_t + Y}\left[-K + \int_0^K F(w - S_t)dw\right], \quad (2.20)$$

$$\underline{C}(S_t, K) = \max\left\{0,\ S_t + R^{-1}\left[-K + \int_0^K F(w - S_t)dw\right]\right\}, \quad (2.21)$$

where F(.) denotes the cumulative distribution function of Y, $Y \in [-S_t, \infty)$.

To make it clear that the Perrakis and Ryan (1984) upper bound is equivalent to Levy (1985), Ritchken (1985) and, as we present further in the text, to Constantinides and Perrakis (2002) bounds, we re-express (2.20) by replacing the cumulative distribution function by the density function. It follows:

$$\overline{C}(S_t, t) = \frac{S_t}{S_t + E[Y]}\int_{K-S_t}^{\infty} (S_t + y - K) f(y)dy, \quad (2.20')$$

[13]Also, bounding the stock process from below and above in Perrakis (1986) resulted in tightening of the bounds derived in Perrakis and Ryan (1984).

where f denotes the density function of the change in the stock price. In this formulation, it is apparent that the call upper bound is the expected call payoff discounted by the expected stock return.

Ritchken (1985) derived single-period option bounds in incomplete risk-averse markets by solving a linear programming problem[14]. The upper bound on the call price coincided with the Perrakis and Ryan (1985) upper bound whereas the lower bound derived by Ritchken (1985) was tighter.

Levy (1985) derived an upper and a lower bound on European options coinciding with those of Ritchken (1985) by an explicit application of the second order stochastic dominance arguments in a single-period context. Levy's bounds were derived in discrete time and as such can incorporate market imperfections, as investors do not have to revise their portfolios continuously.

The derivation of the bounds relies on a theorem regarding the second order stochastic dominance (SSD) of risky portfolios containing riskless asset (SSDR). Let X and Y be random variables whose cumulative distributions are F and G, respectively. Then X dominates Y if and only if:

$$\int_{-\infty}^{z}[G(t)-F(t)]dt \geq 0 \quad (2.22)$$

for all z, with the strict inequality for at least one z_0, $z_0 \in (-\infty, z]$. The theorem, proven by Levy and Kroll (1978) states that if there is at least one α for which a portfolio whose

[14] Ritchken and Kuo (1988) generalized single-period linear-programming results into multi period setting.

return is $X_\alpha = \alpha X + (1-\alpha)R$ SSD dominates a portfolio whose return is $Y_\alpha = \alpha Y + (1 - \alpha)R$, then the entire set $\{X_\alpha\}$ dominates the set $\{Y_\alpha\}$.

The call price for which the investment in the call does not SSDR dominate the investment in stock establishes the lower bound on the call price $\underline{C} \equiv \max[0, C_L]$:

$$C_L = S - K/R - \frac{1}{Rp_c}\int_0^K (S_T - K) f(S_T) dS_T \ , \quad (2.23)$$

where $f(S_T)$ denotes the density of the terminal price of the underlying asset and p_c is the return on the underlying asset for which the investment in the call does not exceed the riskless return. The derived bound is tighter than the Merton (1973) lower bound since the integral is never positive; it coincides with the Ritchken (1985) lower bound.

Deriving the smallest call value for which the investment in the stock does not SSDR dominate the investment in the call establishes the upper call bound:

$$\overline{C} = \frac{S_t}{E(S_T)}\int_K^\infty (S_T - K) f(S_T) dS_T \ . \quad (2.24)$$

The upper bound has a straightforward interpretation as the expected call payoff discounted by the expected return on the underlying asset. It is the same as the Perrakis-Ryan result.

CP derived bounds for a multi-period economy for European cash-settled options by extending stochastic dominance arguments to incorporate transaction costs. Proposition 1 and Proposition 6 establish, respectively, at any time t prior to option expiration T, the

upper bound on the reservation write price of a call $\overline{C}(S_t, t)$ and the lower bound on the reservation purchase price of a put $\underline{P}(S_t, t)$ [15]:

$$\overline{C}(S_t, t) = \frac{1+k_1}{1-k_2}\left(\prod_{s=t}^{T-1} \hat{R}_s\right)^{-1} E[(S_T - K)^+ \mid S_t], t \le T-1 \quad (2.25)$$

and

$$\overline{C}(S_T, T) = [(1+\delta_T)S_T - K]^+,$$

where $\hat{R}_s \equiv E\left[\left(1+\frac{\delta_{s+1}}{1+k_1}\right)\frac{S_{s+1}}{S_s} \mid S_t\right]$ is one-period conditional mean return with the reinvestment of deterministic dividends that yield δ_s in period s.

$$\underline{P}(S_t, t) = \frac{1-k_2}{1+k_1}\left(\prod_{s=t}^{T-1} R_s\right)^{-1} E[(K - S_T)^+ \mid S_t],\ t \le T-1 \quad (2.26)$$

and

$$\underline{P}(S_T, T) = [K - S_T]^+,$$

where $R_s \equiv E\left[(1+\delta_{s+1})\frac{S_{s+1}}{S_s} \mid S_s\right]$ is the one-period conditional mean return.

The upper (lower) bound on the call (put) price is the expectation at time zero of the option payoff with the actual probability distribution of the stock price at T discounted by the return on the underlying asset inflated (deflated) by the round-trip transaction costs. Attractive features of (2.25) and (2.26) are that the bounds are independent from the frequency of trading, as the transaction costs enter only once, and that the bounds can be

[15] Results presented here include corrections for dividends added in Constantinides and Perrakis (2002a).

derived for any given arbitrary distribution of the stock price, provided that the first moment exists.

The derivation of the bounds stipulates that investors hold the riskless bond and one primary risky asset, which constraints the bounds applicability to index options. The proof of Proposition 1 compares the derived utility of an investor holding an optimal portfolio composed of the two primary assets and the derived utility[16] of an investor with similar holdings with an added short position in one call option. The position is composed of one short call sold at the upper bound, with the proceeds invested in the underlying asset; it entails zero initial cost. For a finite investment horizon T' (T' $\geq$ T), CP demonstrated that the derived utility J of the call writer has to be larger or equal than the derived utility V of an investor with identical holdings save for the position in the derivative[17]. The argument of an increase in derived utility applied in the proof together with the assumption that the utility function is concave and increasing establishes the Proposition 1 upper bound as a stochastic dominance result. The link with previous stochastic dominance results is clear once we observe that the term $\left(\prod_{s=t}^{T-1}\hat{R}_s\right)^{-1} E[(S_T - K)^+ \mid S_t]$ in the Proposition 1 upper bound is, by the definition of the expected call payoff, the upper bound on the call price derived by Perrakis and Ryan (1984) and re-derived by different arguments by Levy (1985) and Ritchken (1985).

[16] The utility function itself is plausibly assumed to be increasing and concave.
[17] The proof of Proposition 6 applies a similar argument with the zero-net-cost position composed of one long put purchased by shorting stock.

By the definition of the Proposition 1 upper bound, a risk-averse investor subject to the monotonicity of wealth condition can improve his expected utility by adding to his portfolio the zero-net-cost position in the derivative priced at or above the call bound. In our study we exploit this bound (2.25) in order to measure the improvement in expected utility by the definition of second order stochastic dominance (2.22) with the cumulative distribution functions of the returns of investors without and with the derivative, from (2.22). To achieve this, we need first to identify violations of the bound within traded European-style call options on some broad index, say S&P 500. As the portfolios have to be optimized before any trading in the derivative occurs, we apply the portfolio optimality conditions in the presence of transaction costs presented in Constantinides (1986) and Davis and Norman (1990). For both the above steps, the true return distribution of the underlying asset (the S&P index) is required. We proxy this true distribution by the distribution derived from historical data. The details of our methodology are provided in Part 3.

3. Methodology and Data

We consider a multi-period economy in which each investor makes sequential investment decisions at discrete trading dates t ($t = 0, 1 \ldots, T'$), where the terminal date is finite. Before the option is introduced, an investor may hold long or short positions in a stock (with a natural interpretation of an index) and/or in a zero-coupon risk-free bond. The bond trades do not incur transaction costs. At date t, the stock pays cash dividends $\delta_t S_t$, where the dividend yield parameters δ_t are assumed to be deterministic and known to the investor at time zero. We assume that the support of S_t is $(0, \infty)$ and that the successive rates of return on the stock are independently distributed with conditional mean return known to the investor at time zero:

$$R_t^S \equiv E\left[(1+\delta_{t+1}) \frac{S_{t+1}}{S_t} \mid S_t \right] . \quad (3.1)$$

We define the conditional mean return with the dividend reinvested in stock, net of transaction costs:

$$\hat{R}_t^S \equiv E\left[\left(1 + \frac{\delta_{t+1}}{1+k_1} \right) \frac{S_{t+1}}{S_t} \mid S_t \right] . \quad (3.2)$$

The distinction between $\hat{R}_t^S$ and R_t^S is negligible provided the dividend yield and the transaction cost rate are of the order of a few percent.

At t, the investor enters with x_t dollars in the bond account and with y_t/S_t *ex dividend* shares of stock. The investor increases (decreases) the stock dollar holdings from y_t to $y_t + v_t$ by decreasing (increasing) the bond account from x_t to $x_t - v_t - \max[k_1 v_t, -k_2 v_t]$,

where k_1 (k_2) is the proportional cost of buying (selling) of the risky asset. In our empirical work we assume $k_1 = k_2 \equiv k$. The investment decision variable v_t is constrained to be measurable with respect to the information set available at t.

Given v_t, the bond account dynamics are:

$$x_{t+1} = \{x_t - v_t - \max[k_1 v_t, -k_2 v_t]\}R + (y_t + v_t)\frac{d_{t+1}S_{t+1}}{S_t},\ t \le T' - 1, \quad (3.3)$$

and the stock account dynamics are:

$$y_{t+1} = (y_t + v_t)S_{t+1}/S_t\ . \quad (3.4)$$

At the terminal date, the stock account is liquidated. The net worth is defined as:

$$W_{T'} = x_{T'} + y_{T'} - \max[-k_1 y_{T'}, k_2 y_{T'}]. \quad (3.5)$$

The investors' objective is to maximize $E[u(W_T)]$. Even though this objective realistically represents the objective of a financial institution, the results are extendible to the case where consumption takes place at each trade date; this extension will be presented further in this chapter. The utility function, $u(.)$, is plausibly assumed to be concave and increasing in (x, y) and defined for both positive and negative terminal worth. The value function of an investor is defined recursively for $t \le T' - 1$:

$$V(x, y, t) \equiv$$
$$\max_v E\left[V\left(\{x_t - v - \max[k_1 v, -k_2 v]\}R + (y_t + v)\frac{\delta_{t+1}S_{t+1}}{S_t}, (y_t + v)\frac{S_{t+1}}{S_t}, t+1\right) \mid S_t\right] \quad (3.6)$$

and

$$V(x, y, T') = u(x_{T'} + y_{T'} - \max[-k_1 y_{T'}, -k_2 y_{T'}]).$$

Next, a European-style, cash settled option is added to the investment opportunity set. The option expires at T, $T \leq T'$. The value function of an investor holding a position in the derivative is defined as:

$$J(x, y, t) \equiv$$
$$\max_v E\left[J\left(\{x_t - j - \max[k_1 j, -k_2 j]\}R + (y_t + j)\frac{\delta_{t+1} S_{t+1}}{S_t}, (y_t + j)\frac{S_{t+1}}{S_t}, t+1 \right) \middle| S_t \right] \quad (3.7)$$

and

$$J(x, y, S, T) = V(x + g(S), y, T),$$

where g(S) is the terminal option payoff at the expiration T. Note that the optimal investment decision j_t of the J-investor may well differ from the optimal decision v_t of the V-investor.

For the multi-period economy, Constantinides-Perrakis (CP, 2002) state in the Proposition 1 the upper bound on the reservation write price of the call option:

$$\overline{C}(S_t, t) = \frac{1+k_1}{1-k_2}\left(\prod_{s=t}^{T-1} \hat{R}_s^S \right)^{-1} E[(S_T - K)^+ \mid S_t], t \leq T-1 \quad (3.8)$$

and

$$\overline{C}(S_T, T) = [(1+\delta_T)S_T - K]^+.$$

Under the assumption of the lognormal process of the stock price, the term $\left(\prod_{s=t}^{T-1} \hat{R}_s^S \right)^{-1} E[(S_T - K)^+ \mid S_t]$ in the Proposition 1 becomes the Black-Scholes (1973) model *with the true expected return* $\hat{R}_s^S$ replacing the riskless rate. In our empirical work we assume identically distributed expected *ex dividend* stock returns and identical

deterministic dividend yields δ for each time period. This simplifies the notation in Proposition 1 by defining $R_S \equiv \hat{R}_1^S = \hat{R}_2^S = \ldots = \hat{R}_T^S$. Under this new notation we have:

$$\overline{C}(S_t,t) = \{(1+k_1)/(1-k_2)\}E[(S_T - K)^+ \mid S_t]/R_S^{T-t}, t \le T-1 \quad (3.8')$$

and

$$\overline{C}(S_T,T) = [(1+\delta)S_T - K]^+ .$$

The proof of the Proposition 1 relies on the key property that the marginal utility is non-increasing in the stock price, which is preserved under the assumption of the monotonicity of wealth: the wealth at time $t \le T$ (including the payoff of the derivative) is a non-decreasing function of the stock price. Whereas intermediate trading may result in the violation of the monotonicity, increasing the initial wealth relative to the position in the derivative can make the probability of such violations arbitrarily small. Adding a zero-net-cost position in one (without a loss of generality) call option consisting of a short call and $\overline{C}(S_t,t)/(1+k_1)$ shares transforms the V-investor into the J-investor. The investor's expected utility increases at t by adopting the zero-net-cost position if, and only if:

$$J(x, y + \overline{C}(S_t,t)/(1+k_1),t) > V(x,y,t). \quad (3.9)$$

In equilibrium, we must have:

$$J(x, y + \overline{C}(S_t,t)/(1+k_1),t) \le V(x,y,t). \quad (3.10)$$

The CP proof exploits the condition that if the call write price exceeds the bound stated in Proposition 1, the equilibrium condition (3.10) is violated. The induction proof finds that

the equilibrium condition holds at T by the definition of the bound. The demonstration that if Proposition 1 holds at t it also holds at t-1 is much more difficult. The proof shows that the difference Δ_t of derived utilities of the J and V investors is nonnegative unless the reservation write price of the call is less than or equal to the bound given in (3.8). This result is derived by imposing the optimal policies of the V-investor onto the J-investor. This condition will remain valid for the remainder of our study. Hence, in what follows the J-investor will have the same holdings as the V-investor save for the added zero-net-cost position in the derivative. Note that if this imposed suboptimality of the j decision variable is relaxed then the Proposition 1 result is strengthened.

The key argument of the proof explores the concavity of the value function in the dollar value of the stock account. This, combined with the monotonicity assumption implies that the function $M \equiv \{y + v_{t-1} + E[(S_T - K)^+|S_{t-1}]/(1 - k_2) R_S^{T-t+1}\} S_t/S_{t-1} - \overline{C}(t)/(1 + k_1)$ is increasing in the stock price, while the partial derivative of V with respect to the second argument is decreasing. Demonstrating that $\Delta_t \geq 0$ implies $\Delta_{t-1} \geq 0$ completes the induction proof of Proposition 1.

The definition of the bound in (3.8) implies that the frequency of trading does not enter the CP Proposition 1. The bound does depend, however, on the monotonicity assumption. On the other hand, the likelihood of a violation of the monotonicity assumption clearly depends on the optimal policy of the V-investor, which in turn depends on the wealth of the call writer. Hence, we need to find how restrictive the condition is, or, how much wealth is needed to assure the satisfaction of the monotonicity

assumption. More to the point, we need to find the number of shares ($\equiv N_0$) that the J-investor must hold at the time of writing one call option at or above the Proposition 1 upper bound so as to bring down almost to zero the likelihood of the violation of the monotonicity assumption. Within the account dynamics described by (3.3) and (3.4), only the decision variable v remains unobservable. As already shown in Part 2, the presence of transaction costs establishes a region of the y to x proportions within which it is optimal to refrain from trading. Hence, we need to find how the no transaction (hereafter NT) region is defined.

We apply the Constantinides (1986) methodology to find $\underline{\lambda}$ and $\bar{\lambda}$, that are, respectively, the lower and the upper bound of the NT region. The Constantinides (1986) method assumes a simple consumption policy under which a constant proportion of the bound account is consumed every period. By the author's assertion, the loss in derived utility due to the assumption is small[18]. Also, the time horizon of the Constantinides (1986) is infinite. We consider that it is plausible to assume the final investment horizon T' is "much" longer than the derivative expiration time T; hence we consider the infinite horizon model as appropriate for the problem at hand.

The optimal simple investment policy of trading to the closest boundary $\underline{\lambda}$ or $\bar{\lambda}$ when we fall outside the NT region and the simple consumption policy at the constant rate β from the bond account, will fully determine the decision variable v of the V-investor. Because of the imposed suboptimality of the J function, the above holds true also for the J-

[18] George Constantinides in personal communication has confirmed the small impact of the simple consumption policy assumption on derived utility.

investor. In section 3.1 we combine the account dynamics in the multi-period economy with the simple investment and consumption policies.

The remainder of this part is organized as follows: Section 1 presents the derivation of the parameters of the NT region by the Constantinides (1986) methodology and the impact of the additional assumptions of this methodology on the assets dynamics and the derived utility in the multi-period economy, Section 2 focuses on the J-investor holdings satisfying the monotonicity of wealth in the share price assumption of Proposition 1, Section 3 deals with the estimation of the distribution of the S&P 500 (hereafter S&P or the Index) return and its application in a search of violations of the Proposition 1 upper bound among the traded Index options, Section 4 applies the second order stochastic dominance (SSD) arguments to measure the improvement in expected utility resulting from the addition of the zero-net-cost position in the call option violating the Proposition 1 upper bound to the V-investor's portfolio, Section 5 specializes the results in examining the relation of the portfolio composition to the improvement in expected utility. We use the SAS system and the Maple software to obtain numerical results.

3.1 Portfolio Optimality Conditions in the Presence of Transaction Costs

Merton (1969, 1971) solved the two-asset optimal investment problem in complete markets for the power utility function[19]. Under the following condition:

$$\rho > \gamma[r + (\mu - r)^2 / \sigma^2 (1-\gamma)], \quad (3.1.1)$$

[19] Merton (1969, 1971) presented a more general solution applicable to the HARA class of utility functions.

where ρ is the impatience (discount) factor, $\gamma < 1$, $\gamma \neq 0$ is the risk aversion coefficient, the optimal y to x proportion is:

$$\lambda^* = \left[\frac{\mu - r}{(1-\gamma)\sigma^2}\right]\left[1 - \frac{\mu - r}{(1-\gamma)\sigma^2}\right]^{-1}. \quad (3.1.2)$$

Note that the condition (3.1.1), which also applies to the portfolio selection problem with transaction costs, is automatically satisfied for $\gamma < 0$.

Under the assumption of a simple consumption policy stipulating consumption at a constant rate from the bond account, Constantinides (1986) solved a similar to Merton (1969) dynamic control problem of maximizing the derived (power) utility:

$$V[x, y; \beta, \underline{\lambda}, \overline{\lambda}] \equiv \max E_0 \int_0^\infty e^{-\rho t} u(c(t))dt = \max E_0 \int_0^\infty e^{-\rho t} \gamma^{-1} c^\gamma(t)dt, \quad (3.1.3)$$

where c(t) denotes consumption at t, $\beta \equiv c/x$ is the parameter of the simple consumption policy, E_0 is the current time expectation over the Wiener process ω. In the NT region, both risky and riskless assets follow:

$$\begin{aligned} dx &= rxdt - cdt = (r - \beta)xdt \\ dy &= \mu ydt + \sigma ydw \ . \end{aligned} \quad (3.1.4)$$

By virtue of the simple consumption and the simple investment policies, the dynamic programming problem controls are $(\beta, \underline{\lambda}, \overline{\lambda})$. The derived utility function $V(x, y; \beta, \underline{\lambda}, \overline{\lambda})$ satisfies the Bellman equation. Substituting $c = \beta x$ we get[20]:

$$(\beta x)^\gamma / \gamma + (r - \beta)xV_x + \mu yV_y + (\sigma^2/2)y^2 V_{yy} - \rho V = 0, \quad \underline{\lambda} \leq y/x \leq \overline{\lambda}. \quad (3.1.5)$$

[20] Subscripts denote partial derivatives.

It can be shown that V is homogenous of degree γ in x and y. The boundary conditions are set by the simple investment policy outside the NT region of trading to the closest boundary:

$$(1+k)V_x = V_y\,, \quad y/x \le \underline{\lambda}\,, \quad (3.1.6)$$

and

$$(1-k)V_x = V_y\,, \quad y/x \ge \overline{\lambda}\,. \quad (3.1.6a)$$

Based on the homogeneity argument and the continuity of derivatives of the value function, the following general solution can be obtained:

$$V(x,y;\beta,\underline{\lambda},\overline{\lambda}) = \frac{\beta^{\gamma}}{\rho-\gamma(r-\beta)}(x^{\gamma}/\gamma + A_1 x^{\gamma-s_1} y^{s_1} + A_2 x^{\gamma-s_2} y^{s_2})\,, \quad (3.1.7)$$

where (A_1, A_2) are free parameters and (s_1, s_2) are the roots of:

$$(\sigma^2/2)s^2 + (\mu-\sigma^2/2-r+\beta)s - [\rho-\gamma(r-\beta)] = 0\,. \quad (3.1.8)$$

Substituting the value function (3.1.7) into the boundary conditions (3.1.6) and dividing by x^{γ} yields a pair of equations in A_1, A_2:

$$(1+k)[1 + A_1(\gamma-s_1)\underline{\lambda}^{s_1} + A_2(\gamma-s_2)\underline{\lambda}^{s_2}] = A_1 s_1 \underline{\lambda}^{s_1-1} + A_2 s_2 \underline{\lambda}^{s_2-1} \qquad (3.1.9)$$

and

$$(1-k)[1 + A_1(\gamma-s_1)\overline{\lambda}^{s_1} + A_2(\gamma-s_2)\overline{\lambda}^{s_2}] = A_1 s_1 \overline{\lambda}^{s_1-1} + A_2 s_2 \overline{\lambda}^{s_2-1}\,. \quad (3.1.9a)$$

Constantinides (1986) proved that from the triplet of maximizing controls ($\beta, \underline{\lambda}, \overline{\lambda}$), ($\underline{\lambda}, \overline{\lambda}$) are independent from (x, y), given that β is set to be independent from (x, y) *by the assumption* of the simple consumption policy, and that the same pair ($\underline{\lambda}, \overline{\lambda}$) that satisfies the necessary optimality condition of A_1 also satisfies the same conditions for A_2. For a numerical solution, (3.20) and (3.20a) yield an expression for $A_1(\beta, \underline{\lambda}, \overline{\lambda})$. Then,

$A_1(\beta, \underline{\lambda}, \overline{\lambda})/[\rho - \gamma(r-\beta)]$ has to be maximized with respect to $(\underline{\lambda}, \overline{\lambda})$. The same pair also maximizes $V(x, y; \beta, \underline{\lambda}, \overline{\lambda})$. Constantinides (1986) defines:

$$V(x, y, \beta) \equiv \max_{(\underline{\lambda}, \overline{\lambda})} V(x, y; \beta; \underline{\lambda}, \overline{\lambda}), \quad (3.1.10)$$

and maximizes $V(x, y, \beta)$ w. r. to β at the value of $y/x = \lambda^*$ corresponding to the optimal portfolio proportions for the Merton (1971, 1969) two-asset optimal investment problem, which implies $\underline{\lambda} < \lambda^* < \overline{\lambda}$. By substituting λ^* into the value function we get:

$$V = x^\gamma \frac{\beta^\gamma}{\rho - \gamma(r-\beta)} (1/\gamma + A_1 \lambda^{*s_1} + A_2 \lambda^{*s_2}) \ . \quad (3.1.11)$$

Because of the homogeneity of V of degree γ in x and y, taking the partial derivative of (3.1.11) w. r. to the consumption rate β will yield the first order condition for the optimality of the value function independent from (x, y). This, together with the optimality conditions for A_1 in $(\underline{\lambda}, \overline{\lambda})$ results in three equations with three unknowns $\beta, \underline{\lambda}, \overline{\lambda}$:

$$\begin{aligned} &[\rho - \gamma(r-\beta)]^{-1} \frac{\partial A_1}{\partial \underline{\lambda}} = 0, \\ &[\rho - \gamma(r-\beta)]^{-1} \frac{\partial A_1}{\partial \overline{\lambda}} = 0, \quad (3.1.12) \\ &\frac{\partial V}{\partial \beta} = 0. \end{aligned}$$

The above set of equations can be solved for $(\beta, \underline{\lambda}, \overline{\lambda})$ numerically. Like Constantinides (1986), we assume the impatience factor ρ to be equal to the riskless rate r.

However, the method presented above does not apply to assets with dividend payouts as in the case of an index. Assuming continuous deterministic dividend yield at the rate δ, the assets' dynamics now become:

$$\begin{aligned} dx &= (r-\beta)xdt + \delta ydt \\ dy &= \mu ydt + \sigma yd\omega . \end{aligned} \quad (3.1.13)$$

For (3.1.13), the resulting Bellman equation has no known closed-form solution[21]. To apply the adjusted value function (3.1.7) in the case of dividends, we approximate the bond account dynamics by substituting in (3.1.13) instead of y the value of x multiplied by λ^*, the optimum y to x proportion inclusive of dividends without transaction costs:

$$\lambda^* = \left[\frac{\mu+\delta-r}{(1-\gamma)\sigma^2}\right]\left[1-\frac{\mu+\delta-r}{(1-\gamma)\sigma^2}\right]^{-1} . \quad (3.1.11')$$

The approximate bond dynamics are:

$$dx = (r-\beta+h)xdt , \quad (3.1.14)$$

where $h \equiv \delta\lambda^*$.

We now prove an auxiliary result.

Proposition. Define V^δ as the value function solving the dynamic optimization problem in the case of dividends under the approximate bond account dynamics (3.1.14). We have, instead of (3.1.7):

$$V^\delta(x,y;\beta,\underline{\lambda},\overline{\lambda}) = \frac{\beta^\gamma}{\rho-\gamma(r-\beta+h)}(x^\gamma/\gamma + A_1 x^{\gamma-s_1} y^{s_1} + A_2 x^{\gamma-s_2} y^{s_2}) , \quad (3.1.7')$$

[21] A review of the dynamic programming literature has revealed that a closed form solution to the Bellman equation rarely exists. See Fleming (1975).

where (s_1, s_2) are the roots of:

$$f(s) \equiv (\sigma^2/2)s^2 + (\mu - \sigma^2/2 - r + \beta - h)s - [\rho - \gamma(r - \beta + h)] = 0, \quad (3.1.8')$$

and (A_1, A_2) are free parameters derived from substituting the adjusted value function (3.1.7') into the boundary conditions (3.1.6) as in the case of assets' dynamics *without* dividends.

Proof. Under (3.1.14), the Bellman equation by Ito's lemma becomes:

$$(\beta x)^\gamma/\gamma + (r - \beta + h)xV_x^\delta + \mu y V_y^\delta + (\sigma^2/2)y^2 V_{yy}^\delta - \rho V^\delta = 0, \quad \underline{\lambda} \le y/x \le \overline{\lambda}. \quad (3.1.5')$$

If we substitute the value function (3.1.7') and its appropriate partial derivatives into (3.1.5') and simplify, we obtain $A_1 (y/x)^{s_1} f(s_1) + A_2 (y/x)^{s_2} f(s_2) = 0$, where f(.) is as defined in (3.1.8'). Since for (3.1.5') to hold we must have $f(s_1) = f(s_2) = 0$, the equation (3.1.8') follows immediately. Substituting the value function V^δ into the boundary conditions (3.1.6) and dividing by x^γ yields the same pair of equations (3.1.9) for (A_1, A_2) as in the no-dividend case, *QED*.

The inclusion of dividends shifts the NT region towards the risky asset since augmenting the expected risky return by the value of δ in the expression (3.1.11') for λ^* increases this number, which in turn enters the value function V^δ maximized w. r. to β, as in the case without dividends.

To obtain numerical results, we solve numerically for:

$$[\rho - \gamma(r - \beta + h)]^{-1} \frac{\partial A_1}{\partial \underline{\lambda}} = 0,$$

$$[\rho - \gamma(r - \beta + h)]^{-1} \frac{\partial A_1}{\partial \bar{\lambda}} = 0, \quad (3.1.12')$$

$$\frac{\partial V^{\delta}}{\partial \beta} = 0.$$

Since for the remainder of our study we examine exclusively a dividend-paying asset, we further refer to (3.1.12') in what follows as the adjusted Constantinides method.

The last complication with finding the NT region arises from the assumed positive weights on both assets included in the value function (3.1.5) or its adjusted form (3.1.5'). For a given risk aversion coefficient γ, a solution to (3.1.12') may or may not exist with positive[22] weights on both assets depending on the remaining problem parameters. In particular, the estimated risk premia of the S&P return would produce a solution to (3.1.12') for high γ only for borrowing (negative x) in most cases of the Index returns examined in sections 3.3-3.5. Finding systematically a negative weight on the riskless asset would contradict generally accepted equilibrium conditions since it implies that additional utility can be derived at the margin by borrowing and investing in the risky asset (Kocherlakota, 1996, Mehra-Prescott 1985). Hence, we search for a "reasonably" high[23] γ for which we can find $(\underline{\lambda}, \bar{\lambda})$ for positive (x, y) for the "majority" of the estimated risk premia. The Maple software has been used to obtain the numerical results upon pre-testing the code[24] by replicating the Constantinides (1986) results.

[22] Constantinides (1986) provided adjustments for negative weights on either asset: replace y with –y in (3.1.7) if y is negative, replace x with –x and β with $-\beta$ in (3.1.7) if x is negative.
[23] In fact, there is little agreement among economists regarding the permissible range of γ. See Kocherlakota (1996) for a survey.
[24] The code is given in Appendix 1.

Here, we combine the adjusted Constantinides method and the multi-period economy. This requires adjustments of the bond account dynamics (3.3) to the simple consumption policy at the constant rate β per period[25]. In the considered sequence of events, the consumption at t + 1 takes place immediately after cash dividends and interest have accrued to the portfolio and immediately before the trading decision v_{t+1}. Similarly to the continuous bond account dynamics, we proxy for the dividend accrual rate per period[26] by $h = \delta\lambda^*$.

$$x_{t+1} = (1+h)(1-\beta)\{x_t - v_t - \max[k_1 v_t, -k_2 v_t]\}R, \ t \le T' - 1. \quad (3.3')$$

Under the simple consumption policy, the V and J value functions (3.6) and (3.7) incorporate an additional control, the consumption rate parameter β.

$$V(x,y,t) \equiv$$
$$\max_{v,\beta}\left\{\frac{(\beta x_t)^\gamma}{\gamma} + E\left[V\left(\{x_t(1+h)(1-\beta) - v - \max[k_1 v, -k_2 v]\}R, (y_t + v)\frac{S_{t+1}}{S_t}, t+1\right) \mid S_t\right]\right\} \quad (3.6')$$

and

$$V(x,y,T') = u(x_{T'} + y_{T'} - \max[-k_1 y_{T'}, -k_2 y_{T'}]).$$

$$J(x,y,t) \equiv$$
$$\max_{j,\beta}\left\{\frac{(\beta x_t)^\gamma}{\gamma} + E\left[J\left(\{x_t(1+h)(1-\beta) - j - \max[k_1 v, -k_2 v]\}R, (y_t + j)\frac{S_{t+1}}{S_t}, t+1\right) \mid S_t\right]\right\} \quad (3.7')$$

and

[25] We hope that the same notation in what follows regarding both the period and annual rates will not confuse the reader.

[26] However, in our discrete time numerical work presented later in the text, we consider a more realistic approach regarding dividends.

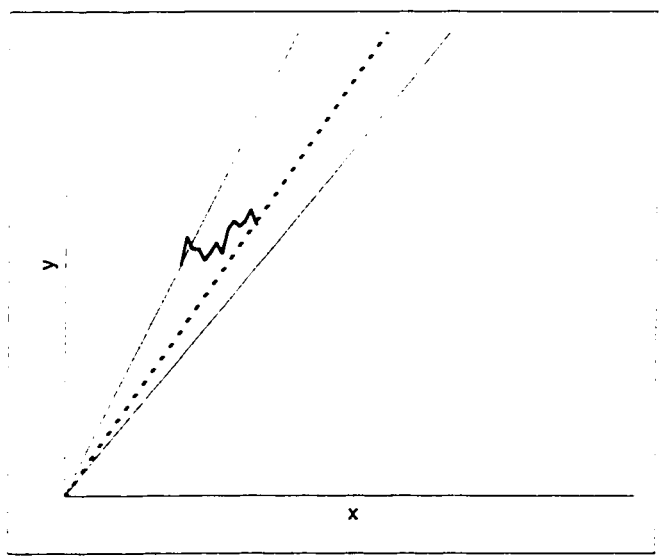
$$J(x, y, S, T) = V(x + g(S), y, T).$$

By combining the assets dynamics in the multi-period economy (3.3') and (3.4) with the simple investment policy, we have:

$$\begin{aligned} &(y_t + v_t)/[x_t(1+h)(1-\beta) + v_t(1+k_1)] = \underline{\lambda}, && y/x < \underline{\lambda}, \\ &v_t = 0, && y/x \in [\underline{\lambda}, \overline{\lambda}], \quad (3.1.15) \\ &(y_t + v_t)/[x_t(1+h)(1-\beta) + v_t(1-k_2)] = \overline{\lambda}, && y/x > \overline{\lambda}. \end{aligned}$$

Figure 1

The NT region and portfolio dynamics in (x, y) coordinates. The solid lines represent the boundaries of the region, the dashed line represents the Merton line, and the irregular curve represents the stochastic portfolio process.



Last, we present several remarks regarding the portfolio dynamics in the NT region applicable also to later sections. Figure 1 may be helpful[27]. We approximate the expected slope of the line joining any two points on the irregular curve of the stochastic portfolio process by $(1 + \mu)/(1 + r + h)(1 - \beta)$, where $h \equiv \delta\lambda^*$. For plausible parameter values, this quantity is close to one. This implies that a portfolio originally situated at the upper boundary $\bar{\lambda}$ of the NT region will tend in time towards the area inside the region if the value of $\bar{\lambda}$ "significantly" exceeds one, thus implying in expectations no sales from the stock account after a "sufficiently" long period of time. The Constantinides (1986) results and our numerical work imply that a value of $\bar{\lambda}$ lower than one can be found as a result of low values of the risk aversion coefficient γ or for other parameter values for which it is plausible to assume a significant portfolio shift towards the riskfree asset, but which rarely can be observed in the market data we deal with in later sections. A value of $\bar{\lambda}$ lower than 1 for a portfolio at the upper boundary of the NT region would imply frequent sales from the stock account.

3.2 The Monotonicity of Wealth in Stock Price

Recall from the induction proof of the Proposition 1 (CP) that the following expression representing the net value of assets of the J-investor as a function of the stock price has to be monotone increasing in the stock price S_t by the monotonicity assumption:

$$M \equiv y_t + [\bar{C}(S_{t-1}, t-1)/(1+k_1)]S_t / S_{t-1} - \bar{C}(S_t, t)/(1+k_1)\,. \quad (3.2.1)$$

[27] We represent a three-dimensional problem with this figure.

where $\overline{C}(S_t, t)$ is given by (3.8'), $y_t = y_{t-1}S_t/S_{t-1}$ and $y_{t-1} = y_{t-2}S_{t-1}/S_{t-2} + v_{t-1}$. Substituting for y_t into (3.2.1) yields the partial derivative w. r. to S_t of the first two parts of M. By substituting the maximum of the partial derivative of $\overline{C}(S_t, t)$ in S_t, which is $(1+k_1)/(1-k_2)$, we derive the following sufficient condition for the preservation of the monotonicity:

$$\frac{\partial M}{\partial S_t} \geq M' \equiv [y_{t-1} + \overline{C}(S_{t-1}, t-1)/(1+k_1)]/S_{t-1} - 1/(1-k_2) > 0\,, \quad (3.2.2)$$

where $y_{t-1} = N_{t-1}S_{t-1}$ and N_t denotes the number of shares optimally held at t.

Note that if the investor holds at t-1 $N_{t-1} > 1/(1-k_2)$ the monotonicity assumption is automatically satisfied, since the derivative of $\overline{C}(S_t, t)$ can never exceed this number. To derive numerical results, we simulate paths of the Index price via Monte Carlo simulations according to a discretized geometric Brownian motion model (Hull, 2002):

$$S_t = S_{t-1}\exp[(\mu - \sigma^2/2)\Delta t + \sigma\varepsilon\sqrt{\Delta t}\,]\,, \quad (3.2.3)$$

where ε is a random sample from N(0,1), Δt is the length of an assumed revision interval in years, μ is the expected *ex dividend* return on the Index. Random numbers from a standard normal distribution were obtained through a SAS pseudo-random number generator leaving all the results readily replicable.

At time t = 0, we form two portfolios, both containing N_0 shares of the Index, $N_0 > 1$. The dollar value of the riskless asset at t = 0 is determined by $y_0/x_0 = \overline{\lambda}$ ($y_0/x_0 = \underline{\lambda}$) for the portfolio formed at the upper (lower) boundary of the NT region. We assume that the

boundaries of the NT region remain unchanged throughout the life of an option. Dividends accrue to the bond account at the rate δ per period and the investor consumes at the rate β per period from this account. Immediately before a revision we have[28]:

$y_t = N_{t-1}S_t$ and $x_t = (1-\beta)\{[x_{t-1} - v_{t-1} - \max(k_1 v_{t-1}, -k_2 v_{t-1})]R + N_{t-1}\delta S_t\}$, t = 1...T. Let x'_t and y'_t denote the dollar value of, respectively, the bond and stock accounts immediately after the adjustments due to the simple investment policy have taken place. For each revision opportunity, at the prevailing Index price, the V-investor makes the following adjustments:

i. If $y/x > \overline{\lambda}$:

$$\begin{aligned} \Delta N_t &= -v_t / S_t = (y_t - \overline{\lambda} x_t) / S_t [1 + \overline{\lambda}(1-k_2)], \\ N_t &= N_{t-1} - \Delta N_t, \\ y'_t &= N_t S_t, \\ x'_t &= x_t + \Delta N_t S_t (1-k_2). \end{aligned} \tag{3.2.4}$$

ii. If $y/x < \underline{\lambda}$:

$$\begin{aligned} \Delta N_t &= v_t / S_t = (\underline{\lambda} x_t - y_t) / S_t [1 + \underline{\lambda}(1+k_1)], \\ N_t &= N_{t-1} + \Delta N_t, \\ y'_t &= N_t S_t, \\ x'_t &= x_t - \Delta N_t S_t (1+k_1). \end{aligned} \tag{3.2.4a}$$

iii. If $\overline{\lambda} \geq y/x \geq \underline{\lambda}$ the investor does not trade.

Recall that the call upper bound $\overline{C}(S_t, t)$ under the lognormal process of the Index price is given by the same expression as the Black-Scholes (1973) model multiplied by the

[28] This formulation implies that consumption takes place right after receiving the interest and cash dividends.

round-trip transaction costs with the expected return on the Index (including the reinvestment of dividends) R_S replacing the riskless rate. It follows:

$$\overline{C}(S_t,t) \equiv \overline{C}_0 = \frac{(1+k_1)}{(1-k_2)}[S_0 e^{-\delta(T-t)}N(d_1) - Ke^{-R_S(T-t)}N(d_2)] \ ,$$
$$d_1 = [\ln(S_t/K) + (R_S - \delta + \sigma^2/2)(T-t)]/\sigma\sqrt{T-t} \ , \qquad (3.2.5)$$
$$d_2 = d_1 - \sigma\sqrt{T-t} \ ,$$

where t = 1 ... T denotes the time of a current revision opportunity and N(.) denotes the standard normal cumulative distribution function (CDF).

At t = 1, for each path of the stock price, we check for: a) $N_{t-1} > 1/(1-k_2)$, b) $\frac{\partial M}{\partial S_t} > 0$,

where $\frac{\partial M}{\partial S_t}$ is given by (3.2.2). If at least one of the conditions a) and b) holds we assign the value of 0 to a counting function, otherwise the counting function assumes the value of 1. At t = 2 ... T if at least one of the conditions a) and b) holds, the counting function assumes the value of the function at t-1, otherwise the counting function assumes the value of 1. Aggregating the counting function across all the paths for each portfolio revision yields at t the estimate of the probability of the event that at any t' ($0 < t' \leq t$), the monotonicity assumption has been violated. We set the controlling variable to be the ending probability of the monotonicity violation. A gradual increase in the number of shares at t = 0 will result in the estimate of N_0 for which the mean ending probability of the monotoncity assumption violation at the option expiration T approaches zero. The decision variable, v_t is determined as shown in (3.2.4), to keep the portfolio inside the NT region. The selected problem parameters presented further in this section result in a width of the NT region that precludes any of the two portfolios from reaching a bound of

the NT region other than the point in which the portfolio process has originated[29]. Since this finding assures non-positive (non-negative) v_t for the portfolio formed at the upper (lower) bound of the NT region, it is clear that the number of shares satisfying the monotonicity assumption has to be higher (lower) for the portfolio formed at the upper (lower) bound of the NT region. This implies that the case of a starting point at the upper boundary of the region is the only interesting one for the purpose of our study.

50,000 paths of the Index price were simulated. The expected *cum dividend* return, the dividend yield and the riskless rate remain fixed at, respectively, 0.11, 0.02 and 0.05. We apply 31 and 45 days to maturity. For an increase in the time to maturity, by the argument presented at the end of section 3.1 regarding the relation of the sales from the stock account to $\bar{\lambda}$, we expect the ending probability of a monotonicity violation increase significantly less for $\bar{\lambda} > 1$ than for $\bar{\lambda} < 1$. We use values of the standard deviation of return of 0.15 and 0.20, transaction costs parameters of the underlying of 0.5% and 1%, moneyness of an option ranging from 0.9 to 1.05 and finally, we assume that the J-investor can revise his portfolio 1 or 3 times a day. An increase in the volatility, *ceteris paribus*, is expected to cause an increase in the ending probability of a monotonicity violation in two ways: first, a higher volatility increases the likelihood of sales from the stock account to revert to the upper bound of the NT region; second, via a decrease in $\bar{\lambda}$ it causes sales of bigger fractions of shares to reach the bound. A decrease in the moneyness, *ceteris paribus*, is expected to cause an increase in the ending probability of a violation, as smaller fractions of shares can be deposited on the stock account from the

[29] This occurs because the width of the NT region is large with respect to the observed variability in the risky asset per transaction period. This was also a characteristic of the data in our empirical work.

proceeds of the short sale of a call. Last, an increase in the frequency of intraday trading, *ceteris paribus*, is expected to cause an increase in the ending probability of violations. At the same time, we also estimate the mean proportion of shares that has to be sold from the original holding to adjust to the portfolio optimality condition[30].

3.3 Estimation of Violations of the Upper Bound on the Call Price

Having derived the conditions for assuring the monotonicity of wealth assumption, we measure the improvement in utility resulting from adopting zero-net-cost positions in call options violating the Proposition 1 bound and adding them to the holdings of the V-investor. As the natural interpretation of the underlying asset in the Proposition 1 is an index, we use as our data base Chicago Mercantile Exchange (CME) data containing intraday quotes on European-style S&P options[31] for the period January 1990 – August 1996 (with the exception of April 1991 – December 1991, for which the data is missing) for calls with 31 and 45 days to maturity violating the Proposition 1 and intraday S&P quotes for the same period. The search is based on fitted lognormal distributions of 31- and 45-day returns on the Index assumed to proxy for the true return distribution.

We assume that the subjective distribution of the Index returns of a representative investor[32] is estimated from the past data. Hereafter we refer to data and/or results for 31 (45) days to expiration as to the Set 1(2) for short. We obtained 71 (72) sampling dates for the Set 1 (2). As estimating the distribution of the gross return on the Index would

[30] A pertinent SAS code is given in Appendix 2.
[31] The investigated options conform to the cash settlement method assumed by CP model.
[32] A similar though nonparametric methodology was applied by Jackwerth (2000).

lead to spurious results due to the variability in the riskless rate, we estimate risk premia rather than the log relative of the S&P value. Under the geometric Brownian model for the Index price we have:

$$\log\frac{S_{T_i}}{S_0} - r \sim \Phi\left[(\theta - \frac{\sigma^2}{2})T_i, \sigma\sqrt{T_i}\right],\ i = 1,2, \quad (3.3.1)$$

where θ denotes the estimated expected market risk premium and r denotes the observed riskless rate. Given the well-documented instability in the distribution of returns on the stock market (Black, 1986, Rubinstein, 2000), we vary the number of 31 (45) day periods used to obtain estimates by going backward from the day prior to observing call data for the Set 1 (2) to find estimates both satisfying $\theta > 0$ and conforming to lognormality. To test for the lognormality of the estimated market risk premium, we apply the Kolmogorov D-test. The test uses the highest absolute difference between the *normal* cumulative distribution function *with the estimated parameters* and the *empirical* cumulative distribution function. The difference is assigned a value D taken from the λ-Kolmogorov distribution. A high value of D leads to rejection of the null of normality. We proceed further with only those sampling dates for which the p-value for rejection of the null is 0.1 or larger, after checking for consistency of the test results with the Shapiro-Wilk and Anderson-Darling normality tests.

In the course of searching for the optimal number of periods for the density estimation, we faced the following trade-off: the number of observations yielding the maximum number of estimated distributions conforming to lognormality (50 and 40 for Sets 1 and 2, respectively) brought about several negative market risk premia, while an increase in the number of periods brought an increase in the number of departures from

lognormality. Settling for a middle ground, we apply 72 31-day periods and 50 45-day periods for the Set 1 and Set 2, respectively[33]. While in our choice there are no negative market risk premia, on the other hand the number of distributions conforming to lognormality is lower than its observed maximum, thus reducing the sample of options available for our study. Daily S&P returns were obtained from the CRSP tape, daily 3-month T-bill rates were obtained from Bloomberg. SAS Insight software has been used to estimate the parameters of the distributions and to perform the tests.

We gauge the performance of the estimated distributions by aggregating a standardized variable $z = [\mu_{SP} - \mu]/\sigma$, where μ_{SP} is the realized return on the Index throughout the life of an option violating the Proposition 1 upper bound, and μ and σ are the estimates of the parameters of the Index return distribution, the former resulting from adding the estimated risk premium to the current period riskless rate. The test cannot confirm the appropriateness of the method applied to search for violations of the Proposition 1 upper bound; however, finding a z value significantly greater than zero would indicate that the observed violations may have resulted from biased estimates of the parameter values.

Restating Proposition 1 under the assumed stochastic process for the Index price with assumed deterministic dividends throughout the life of an option, we have:

$$\begin{aligned} &\overline{C}(S_0,0) \equiv \overline{C}_0 = \frac{(1+k_1)}{(1-k_2)}[S_0 e^{-\delta T_i} N(d_1) - K e^{-R_S T_i} N(d_2)], \\ &d_1 = [\ln(S_0/K) + (R_S - \delta + \sigma^2/2)T_i]/\sigma\sqrt{T_i}, \qquad (3.3.2) \\ &d_2 = d_1 - \sigma\sqrt{T_i}, i = 1, 2, \end{aligned}$$

[33] Note that the total length of time to estimate the parameters of each density function is approximately equal for both sets.

where S_0 denotes the Index price contemporaneous to the observed call price, δ denotes the realized dividend yield till the option expiration, $R_S = \theta + r + \delta/(1 + k)$, where θ is the estimated risk premium and r is the observed 3-month T-bill rate at time zero, σ is the estimated volatility of the S&P return, N(.) denotes standard normal cumulative distribution function. Finding a call bid quote (C_0) greater than $\overline{C}_0$ indicates a violation of the Proposition 1 upper bound. For our analysis, we set k = 0.005, which is assumed to realistically approximate the cost of trading in the Index. Daily dividend yields on the Index were obtained from Bloomberg.

Since the CME option tape records ask (bid) flags only if an ask (bid) is lower (higher) than its immediate predecessors for a given strike price, we enrich the data set by recovering flags by concluding that if, for a given strike, a non-flagged quote is higher (lower) than its immediate predecessor, the quote is an ask (bid). Table 1 below presents basic quantitative characteristics of the CME option data. Upon finding call quoted prices at bid, we proceed further only with those that can be matched within 10 seconds with an S&P quote. As a safeguard against misreading flags, we compare proportions of violating calls for the original and the recovered data[34].

[34] See Appendix 3 for pertinent SAS codes.

Table 1

CME Call Data (Calls at Bid)

	31 Days to Expiration (71 sampling dates)			45 Days to Expiration (72 sampling dates)		
Year	Original	Recovered	Total	Original	Recovered	Total
1990	392	415	807	238	188	426
1991	111	66	177	161	97	258
1992	326	372	698	157	138	295
1993	279	434	713	214	177	391
1994	311	421	732	230	177	407
1995	286	506	792	257	226	483
1996	235	482	717	195	227	422
Total	1940	2696	4636	1452	1230	2682

3.4 Improvement in Expected Utility

To measure the improvement in expected utility resulting from the added zero-net-cost position in the call violating the CP upper bound, we apply the second order stochastic dominance (SSD) arguments. SSD does not require stronger assumptions than increasing and concave utility. We assume that the trading decisions, of the V- and J-investors are determined by the portfolio optimality conditions derived by the adjusted Constantinides (1986) method with the inputs of the parameters of the estimated S&P return distribution, the observed current period riskless rate, and the dividend yield at the option expiration. Hence, we form the portfolios of the V- and J-investors within the NT region and simulate their returns by means of Monte Carlo simulation of the model (3.2.3) conditional on the validity of the estimated distribution derived by the method described in section 3.3. The J-portfolio will consist of the V-portfolio holdings plus one zero-net-cost position based on a call option whose price has violated the Proposition 1. To satisfy

the monotonicity assumption, we consider portfolios containing at least the minimum value for N_0 that corresponds to an insignificantly small probability of violation of the assumption as derived by the methodology presented in 3.2.

We make several assumptions simplifying our numerical work. First, we fix the NT region for the life of an option at the boundaries derived for the set of parameters observed or estimated at time zero. Second, we set dividends to arrive each day in equal dollar amounts per one share $d_T \equiv \delta S_{SP} \Delta t$, where δ and S_{SP} are, respectively, the dividend yield realized at the option expiration T, and the closing S&P price at T. Third, we invest dividends at the riskless rate observed at time zero.

Under the suboptimality conditions of trading imposed on the J-investor, the quantitative difference between the j and v decision variables may result only from a change in the proportion of y to x caused by the inclusion of extra shares at t = 0. To underscore this difference[35], in what follows we keep j as denoting the trading decisions of the J-investor, even though they are derived by the same optimality criteria as the v trading decisions.

In model 3.4.1 we derive at T values of the V and J portfolios inclusive of the value of the consumption stream accrued from t = 1 to t = T to compare expected utilities of the V- and J-investors. We compound the intermediate consumption until time T with the discount rate ρ, which is assumed to be equal to the riskless rate observed at time zero[36]. Since for the numerical derivation of expected utility, consumption is not distinct from

[35] The difference is small, given that the number of extra shares is of the order of a few percent.
[36] Alternatively, we can discount expected utilities at T to derive the improvement at time zero.

the bond account holdings, we do not separate it in model (3.4.2) below. This representation will simplify the derivation of several analytical approximations later in this section and in 3.5.

Under the numerical simplifications above, we specify the assets dynamics with the simple consumption and investment policies:

$$\begin{aligned} x_t &= (1-\beta)\{[x_{t-1} - v_{t-1} - \max(k_1 v_{t-1}, -k_2 v_{t-1})]R + N_{t-1} d_T\}, \\ y_t &= N_{t-1} S_{t-1}, \\ t &= 1 \ldots T. \end{aligned} \tag{3.4.1}$$

The decision variable v and the portfolio adjustments due to the simple investment policy[37] are determined as shown in (3.2.4).

Let R_{kT} and N_{kT}, k = v, j denote, respectively, the holding period returns and the number of shares held at the call expiration T. To include the transaction costs into the portfolio return, thus making R_{vT} comparable to R_{jT}, we compute the returns by applying the liquidating values of the V- and J-portfolios at time zero and at the option expiration T, inclusive of compounded consumption. Let λ denote the y to x proportion right before the zero-net-cost position has been adopted by the J-investor at time zero, and let $n_0 \equiv C_0/S_0(1+k_1)$ denote the number of extra shares acquired by the J-investor with the proceeds from writing one call option.

[37] In the considered sequence of events, the portfolio adjustments occur immediately after consumption has taken place

$$R_{vT} = \frac{x_0 R_i - \sum_t \{v_t - \max[k_1 v_t, -k_2 v_t]\} R_i^{T_i - t} + D_{vi} + N_{vT} S_T (1-k)}{N_0 S_0 (1 - k + 1/\lambda)},$$

$$R_{jT} = \frac{x_0 R_i - \sum_t \{j_t - \max[k_1 j_t, -k_2 j_t]\} R_i^{T_i - t} + D_{ji} + N_{jT} S_T (1-k) - [S_T + d_T - K]^+}{N_0 S_0 (1 - k + 1/\lambda)}. \tag{3.4.2}$$

$$x_0 = N_0 S_0 / \lambda \ ,$$
$$D_{ki} = \sum_t N_{ki} d_T R_i^{T_i - t} \ ,$$
$$N_{vT} = N_0 + \sum_t v_t / S_t \ ,$$
$$N_{jT} = N_0 + n_0 + \sum_t j_t / S_t, \ k = v, j, \ i = 1, 2,$$

where R_i is the riskless return observed at t = 0 for the expiration time T_i, d_T denotes an equal in each period dividend payment per one share, D_k is the dollar value at T_i of dividends arriving in a portfolio throughout the life of an option. The first term in the numerator of the expressions for R_k in (3.4.2) above implies we measure the value at T_i of the bond holdings at time zero. The next two terms, however, subtract (add) the values *at* T_i of any changes due to the bond account dynamics incurred in the course of the life of an option.

To derive returns at the option expiration, we apply (3.4.2) to 200,000 price paths of the Index simulated through the discretized geometric Brownian motion model (3.2.3) to which we apply the distribution parameters estimated by (3.3.1). The Index value was simulated starting from the price contemporaneous to each of the observed violations included in a representative sample described later in this section.

In the numerical applications, we sample the rates of return on the J- and V-portfolios at the option expiration:

$$r_k = \frac{1}{T_i} \log R_{kT} \,, \qquad (3.4.3)$$
$$\mathrm{k} = \mathrm{v}, \mathrm{j}, \mathrm{i} = 1, 2.$$

Although the Constantinides-Perrakis (2002) are "stochastic dominance" results insofar as they apply to all risk-averse utility functions, their link with the traditional stochastic dominance approach to portfolio selection is less clear. The traditional approach[38] compares the probability distribution of the terminal wealth of two alternative portfolios. An SSD relation of one portfolio by the other requires that an integral condition such as (3.4.5) below holds *for all* values z_0 in the distribution domain. Here we demonstrate that the CP results are in fact an application of the traditional SSD relations, modified to incorporate the intermediate trading and transaction costs.

Let F_k , k = v, j denote the cumulative distribution functions of the portfolio rate of return. Having obtained the terminal stock prices from the Monte Carlo simulation and the rates of return by (3.4.2), we estimate F_k at the discrete intervals η as follows:

$$\hat{F}_{kw} = m_{kl} / M_{S_T} \,, \qquad (3.4.4)$$
$$m_{kw} : r_k \le w\eta \,,$$

where w = -80, -79, … , 89, 90, η = 0.01, M_{S_T} is the number of simulated paths of the Index price, k = v, j. The covered interval of the observed return satisfies for its lower

[38] E.g. Hadar-Russel (1969), Hanoch-Levy (1969), Levy-Kroll (1979); see Levy (1992) for a survey.

(upper) bound $F_k \cong 0(1)$. Setting the same variable of integration for both investors, we measure the improvement in expected utility[39]:

$$H(z_0) \equiv \int_{-\infty}^{z_0} (F_v - F_j) dz. \quad (3.4.5)$$

It also follows that $H(\infty)$ (alternatively, the value of H at which $F_v = F_j = 1$) is the expected excess return across all the states. For the J-investor portfolio to show SSD over the V-investor portfolio $H(z_0)$ must be nonnegative for all z_0. A sufficient condition for this is that F_v and F_j cross only once at some stock price denoted by S_T^0, since F_v is clearly above F_j for "small" values of z_0. We demonstrate that such a single crossing holds under some simplifying assumption and verify it empirically in our results *without* these simplifying assumptions.

We estimate $H(z_0)$ by numerical integration via the trapezoid approximation:

$$\hat{H}(z) \equiv \hat{H}(s\eta) = \sum_{w=-79}^{s} (\Delta_{F,w-1} + \Delta_{F,w})\eta/2, \quad (3.4.6)$$

where $\Delta_{F,w} = \hat{F}_{vw} - \hat{F}_{jw}$ and $-79 \le s \le 90$. By the SSD argument, finding a non-negative $\hat{H}$ for each s will provide evidence for the improvement in expected utility from adopting the zero-net-cost policy.

To evaluate the stock price S_T^0 for which the excess return of the J-investor turns negative, in the theoretical discussion below we simplify model (3.4.2) by setting the decision variables v and j to zero, which implies disregarding the transaction costs of the intermediate restructuring of the portfolios. Alternatively, we may assume that the

[39] It is apparent that (3.4.5) must be non-negative for all z_0 to satisfy the SSD condition (2.22).

cumulative sums of the v- and j-terms in (3.4.2) are equal and cancel out. To justify this assumption, we measure the excess return under the simplified model over the full model (3.4.2) applied to the J-investor whose portfolio is set at the upper boundary of the NT region, where the sales from the stock account and the resulting transaction costs are the highest in the region. We found that this number is of order 0.1% annually or lower[40] implying that the difference in excess return due to transaction costs of the V-investor over the J-investor by the full model (3.4.2) is even smaller, thus justifying its neglect in computing R_E, the excess return of the J-investor. Under the simplifying assumptions for model (3.4.2) we have:

$$R_v = \frac{x_0 R_i + N_0[S_T(1-k)+D]}{N_0 S_0(1-k+1/\lambda)},$$

$$R_j = \frac{x_0 R_i + (N_0+n_0)[S_T(1-k)+D] - [S_T + d_T - K]^+}{N_0 S_0(1-k+1/\lambda)}, \qquad (3.4.7)$$

$$R_E \equiv R_j - R_v = \frac{n_0[S_T(1-k)+D] - [S_T + d_T - K]^+}{N_0 S_0(1-k+1/\lambda)}, \qquad i = 1, 2.$$

where D is the dollar value at T_i of dividends *per one share* accruing in a portfolio throughout the life of an option. By inspection of (3.4.7) we note that the excess return of the J-investor reaches its maximum at $S_T = K - d_T$ and decreases afterwards. It is positive (negative) to the left (right) of the unique terminal stock price:

$$S_T^0 \equiv (K - d_T + n_0 D)/[1-(1-k)n_0]. \quad (3.4.8)$$

This value is independent from the initial wealth. Hence, this result demonstrates that the improvement in expected utility will occur approximately independently from the initial

[40] Results are presented in Part 4.

wealth[41], provided that the monotonicity of wealth in the stock price is satisfied. Note also that large values of n_0 inherent to calls with high moneyness can more than compensate the J-investor for the high risk of the exercise provided it shifts S_T^0 sufficiently to the right in terms of the true distribution of the Index. Even though approximate, the critical value of S_T^0 does not depend on the assumed stock process and can be derived by an investor *ex ante*. For instance, one can apply a different stochastic process of the Index price than the Brownian motion to estimate the likelihood[42] of the terminal price lying to the left of S_T^0.

S_T^0 will also uniquely determine the portfolio return for which the excess return is zero. By substituting the expression (3.4.8) for S_T in (3.4.7) we derive the unique portfolio return R^0 for which $R_{vT} = R_{jT}$, which, like the critical stock price S_T^0, is also approximately independent from the initial wealth:

$$R_i^0 = \frac{R_i}{\lambda(1-k+1/\lambda)} + \frac{S_T^0(1-k)+D}{S_0(1-k+1/\lambda)}, \quad i = 1, 2. \quad (3.4.9)$$

To derive the numerical results[43] we consider two cases: i. The V and J portfolios are situated at the midpoint of the NT region in terms of the y/x proportion by setting $\lambda = (\bar{\lambda} + \underline{\lambda})/2$ in (3.4.2), ii. The V and J portfolios are situated at the upper boundary of the NT region. For case ii., we apply 1 and 9 portfolio revisions per day and derive the

[41] The return difference R_E does, however, depend on the initial wealth.
[42] Note, however, that the NT region and the optimal policy will no longer be the same.
[43] See Appendix 4 for pertinent SAS codes.

difference in the improvements in expected utility due an increase in the revision frequency.

Case ii. causes a complication concerning the asset proportions. Consider that adding the zero-net-cost position to a portfolio already containing the y/x proportion $\bar{\lambda}$ would cause the investor to sell the extra shares by virtue of the optimal investment policy. As a solution, we set the y/x proportion so that the acquisition of the zero-net-cost position brings λ to $\bar{\lambda}$ exactly. It implies $\lambda = N_0\bar{\lambda}/(N_0 + n_0)$, to be substituted in (3.4.2) for case ii. We use returns derived by (3.4.2) to estimate by (3.4.6) the improvement in expected utility as defined by (3.4.5). Section 3.5 deals with the impact of changing the portfolio composition on improvements in expected utility.

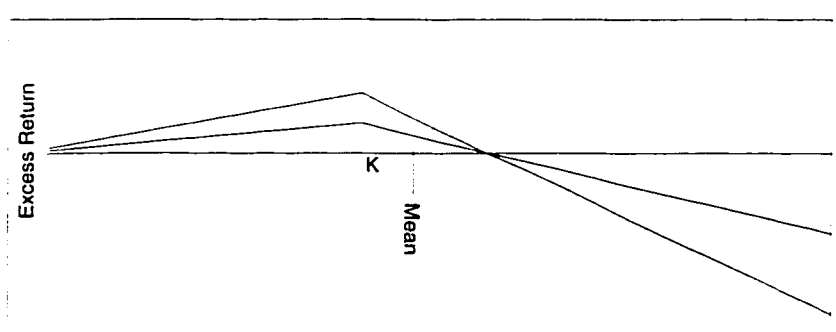
We demonstrate below numerical results in four representative cases for each maturity: i. Low moneyness, low violation (LL), ii. Low moneyness, high violation (LH), iii. High moneyness, low violation (HL), iv. High moneyness, high violation (HH). We define the violation size as the ratio of the observed call price C_0 to the Proposition 1 upper bound $\bar{C}_0$. A subscript i = 1, 2 to LL, LH, HL and HH refers to the corresponding option set. We anticipate any expected utility maximizer to prefer case HH to any other regardless of risk preferences.

3.5 Portfolio Composition and Improvements in Expected Utility

In this section we examine the contribution to the improvement in expected utility of the portfolio composition of the V- and J-investors. This improvement depends on the number of shares N_0 at time zero *per one zero-net-cost position* strategy represented by the writing of one call option, as well as on the proportion of risky to riskless asset y/x. Here we examine separately these two factors. We vary N_0 to measure the resulting improvements in expected utility due to the first factor, and measure the effects of the second factor by comparing *improvements* in expected utility derived for portfolios situated within the NT region for $\lambda = (\bar{\lambda} + \underline{\lambda})/2$ and for $\lambda = N_0\bar{\lambda}/(N_0 + n_0)$, the two cases of our empirical work discussed in the previous section.

Figure 2

The excess return of the J-investor R_E as a function of the terminal stock price. The interval of the terminal stock price represents "almost" the entire distribution. The lower curve to the left of the crossing point with the X-axis S_T^0 represents the R_E of an investor whose wealth at time zero has increased by a factor of two relative to an investor represented by the upper curve.



The approximate independence of the result (3.4.8), the value S_T^0 for which the excess return R_E is zero, from the initial wealth means that the point where the CDF of the J-investor's return crosses from below its equivalent for the V-investor would not change if the initial number of shares N_0 per one zero-net-cost position increases. We consider the quantitative impact of an increase in N_0 on the improvement in expected utility. Figure 2 may be helpful. It is apparent that the expected excess return is positively (negatively) related to the area between each curve and the X-axis to the left (right) of the crossing point with the X-axis S_T^0. From the definition of R_E in (3.4.7) it can be easily seen that increasing the number of shares N_0 by a factor decreases proportionally both areas by its reciprocal. Since stochastic dominance implies that the area to the left of S_T^0 is larger than the area to the right, the gains in expected utility resulting from an increase in the number of shares to the right would never exceed the losses to the left of S_T^0. In terms of the cumulative distribution functions of returns of both investors from which we derive improvements in expected utility, an increase in N_0 will turn clockwise the CDF of the J-investor return around the point we proxy for by R_i^0, decreasing the areas between the CDF curves of the J- and V-investors in a similar way and proportion as shown in Figure 2 for the excess returns. Hence, we expect that increasing the initial number of shares N_0 by a factor will decrease the improvement in expected utility approximately by its reciprocal. In our numerical results, we derive the improvement in expected utility for increases in N_0 for the representative set of options described in the previous section.

For the effect of an increase in λ, the y to x proportion, we generally expect to find a positive *change* in the *improvement* of expected utility as λ increases. An approximately

sufficient condition for a positive change in the improvement of expected utility across all the states as λ increases is a shift to the right of the quantity R^0 as defined by (3.4.9). We derive a condition on the number of shares n_0 purchased with the proceeds from writing one call option for a positive shift in R^0 to occur. By inspection of the partial derivative of R^0 w. r. to λ, which is equal to

$$\frac{\frac{S_0R_i}{\lambda}+\frac{(K-d_T+n_0D)(1-k)}{1-(1-k)n_0}+D}{S_0(1-k+1/\lambda)\lambda^2}-\frac{R_i}{(1-k+1/\lambda)\lambda^2}$$

, we observe that this quantity is monotone increasing in n_0. Hence, finding n_0 for which the partial derivative of R^0 w. r. to λ is zero ($\equiv n^{\lambda}$) sets a critical value above which R^0 increases for an increase in λ. By solving for n^{λ}, we get:

$$n_i^{\lambda}=\frac{S_0R_i+d_T-K}{S_0R_i(1-k)}-\frac{D}{1-k}\ ,i=1,2. \quad (3.5.1)$$

This result is approximately independent from the initial wealth, provided the monotonicity condition is satisfied.

We derive the numerical results for the set of representative options described earlier for portfolios with $\lambda=N_0\bar{\lambda}/(N_0+n_0)$ and $\lambda=(\bar{\lambda}+\underline{\lambda})/2$, and compare the improvements in expected utility.

4. Results

Table 2 below presents the estimates for the mean ending probability of the monotonicity violation. This probability is positively related to the volatility and to the transaction cost rate. The sensitivity to these variables; however, increases in the moneyness and/or in the initial number of shares N_0. The impact of the trading frequency on the increase in the mean ending probability is of a lower order. The mean ending probability decreases rapidly, as expected, with increases in the moneyness; however, for low N_0 at the low moneyness values this sensitivity vanishes. The numerical results in Table 2 imply that the mean ending probability as a function of the moneyness, is concave for the low and convex for the high moneyness. In Table 2, we observe a decrease in the mean ending probability to almost zero for the initial number of shares N_0 of approximately equal to 1.075.

A comparison of the results in Table 2 for 31 days to maturity with those for call maturing in 45 days, provides evidence for the relatively low sensitivity of the mean ending probability to the time to maturity. The "threshold" initial number of shares for which the mean ending probability approaches zero has not changed with an increase in the time to maturity, save for the case of the lowest moneyness, the highest volatility and the highest transaction cost rate, where we observe this probability of the order of 10%.

A percentage decrease in the number of shares held at the option expiration derived in this step to proxy for the effects of the dynamic trading, ranges from 1.55% to 2.69%.

Table 2
Mean Ending Probabilities of the Monotonicity Violation

N_0	n	S/K	T = 45 days				T = 31 days			
			Pr*		Pr**		Pr*		Pr**	
			σ=0.20	σ=0.15	σ=0.20	σ=0.15	σ=0.20	σ=0.15	σ=0.20	σ=0.15
1.025	3	0.90	0.7289	0.5906	0.6473	0.4712	0.6659	0.4910	0.5689	0.3593
		0.95	0.7246	0.5638	0.6343	0.4126	0.6146	0.3314	0.4483	0.1298
		1.00	0.3828	0.1322	0.2611	0.0555	0.1433	0.0166	0.0730	0.0038
		1.05	0.1000	0.0104	0.0577	0.0030	0.0158	0.0003	0.0063	0.0000
	1	0.90	0.7026	0.5668	0.6216	0.4463	0.6325	0.4619	0.5369	0.3324
		0.95	0.6929	0.5271	0.5961	0.3582	0.5328	0.2493	0.3560	0.0881
		1.00	0.3120	0.0994	0.2056	0.0388	0.0995	0.0097	0.0470	0.0021
		1.05	0.0739	0.0065	0.0415	0.0017	0.0094	0.0001	0.0033	0.0000
1.050	3	0.90	0.3699	0.1264	0.3089	0.0789	0.2503	0.0219	0.1797	0.0045
		0.95	0.2075	0.0120	0.1287	0.0035	0.0452	0.0001	0.0206	0.0000
		1.00	0.0359	0.0004	0.0192	0.0000	0.0029	0.0000	0.0010	0.0000
		1.05	0.0042	0.0000	0.0019	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.90	0.3465	0.1121	0.2881	0.0598	0.2194	0.0124	0.1426	0.0025
		0.95	0.1581	0.0071	0.0944	0.0021	0.0283	0.0001	0.0113	0.0000
		1.00	0.0252	0.0001	0.0126	0.0000	0.0017	0.0000	0.0006	0.0000
		1.05	0.0026	0.0000	0.0012	0.0000	0.0000	0.0000	0.0000	0.0000
1.075	3	0.90	0.1077	0.0003	0.0631	0.0001	0.0139	0.0000	0.0053	0.0000
		0.95	0.0132	0.0000	0.0063	0.0000	0.0004	0.0000	0.0002	0.0000
		1.00	0.0010	0.0000	0.0004	0.0000	0.0000	0.0000	0.0000	0.0000
		1.05	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.90	0.0789	0.0002	0.0439	0.0000	0.0077	0.0000	0.0026	0.0000
		0.95	0.0083	0.0000	0.0039	0.0000	0.0002	0.0000	0.0001	0.0000
		1.00	0.0007	0.0000	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000
		1.05	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
1.100	3	0.90	0.0044	0.0000	0.0019	0.0000	0.0001	0.0000	0.0000	0.0000
		0.95	0.0002	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
		1.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		1.05	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.90	0.0028	0.0000	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000
		0.95	0.0001	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000
		1.00	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
		1.05	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

* k = 1% ** k = 0.5%

T is the time to expiration, N_0 is the number of shares held at t = 0, n is the number of revisions per day, S/K is the option moneyness, k is the cost of trading in the underlying index, σ is the volatility of an index. Other parameters are as follows: expected *ex dividend* rate of return 0.11, dividend yield 0.02, riskless rate 0.05. For σ = 0.20 and k = 1% (0.5%) the upper bound of the NT region $\overline{\lambda}$ is 0.742 (0.736) and the consumption rate β is 0.088 (0.089). For σ = 0.15 and k = 1% (0.5%), the upper bound of the NT region $\overline{\lambda}$ is 1.935 (1.894) and the consumption rate β is 0.122 (0.124). The risk aversion coefficient γ is –6.

Tables 3 and 4 below present the data used to search for violations of the Proposition 1 upper bound. The lognormality criterion applied to select observations reduced the sampling dates available to study for the Set 1 (2) from 71 (72) to 27 (28). The presented distribution estimates confirm the instability of the distribution since a difference between sampling dates of, say, two months may result in the difference in market risk premia of the order of 2%. In this regard, a more stable parameter for small time differences is the Index volatility.

We observed in total 147 (209) violations of the Proposition 1 upper bound at 7 (11) sampling dates for the Set 1 (2). The z-statistics applied to control for possible estimation biases resulting in spurious violations is 0.09 (0.06) for the Set 1 (2), and 0.08 for the data aggregated for dates when violations were observed. In neither case are the z-statistics significantly different than zero.

Table 5 below present descriptive statistics regarding the proportions of the call quotes, which are in violation of the Proposition 1 upper bound. The proportions of options included in the sample of violations with the original and recovered flags for bids are similar; thus indicating the appropriateness of the method used to recover the bid flags. In Table 5 we observe that the proportion of violations for the Set 2 is more than twice as big as the proportion for the Set 1. This may be caused by a higher degree of mispricing of options with longer maturity, larger errors of the distribution estimates, or both.

Table 3

Characteristics of Data Used to Search for Violations of the Proposition 1 Upper Bound for 31-day to Expiration CME S&P 500 European Call Options

Date	D-stat.	p-value	σ	$\theta - \sigma^2/2$	z	Yield	T-bill Rate	n
910116	0.085	> 0.15	0.1794	0.0395	2.687	0.0328	0.0612	20
910213	0.077	> 0.15	0.1842	0.0613	0.131	0.0324	0.0602	1
920219	0.089	> 0.15	0.1723	0.0452	0.071	0.0300	0.0393	n.a.
920320	0.093	0.121	0.1841	0.0414	-0.117	0.0300	0.0413	n.a.
920415	0.086	> 0.15	0.1585	0.0352	-0.391	0.0300	0.0369	n.a.
920819	0.068	> 0.15	0.1590	0.0275	0.192	0.0293	0.0313	n.a.
920916	0.081	> 0.15	0.1650	0.0312	-0.469	0.0301	0.0293	n.a.
921021	0.052	> 0.15	0.1578	0.0233	0.517	0.0290	0.0296	n.a.
930120	0.090	> 0.15	0.1490	0.0290	-0.012	0.0285	0.0306	n.a.
930421	0.087	> 0.15	0.1580	0.0098	0.090	0.0280	0.0286	n.a.
931222	0.057	> 0.15	0.1207	0.0503	0.322	0.0265	0.0311	n.a.
940119	0.075	> 0.15	0.1178	0.0556	-0.549	0.0269	0.0303	n.a.
940216	0.078	> 0.15	0.1136	0.0422	-0.219	0.0270	0.0334	n.a.
940316	0.076	> 0.15	0.1076	0.0450	-1.783	0.0285	0.0358	n.a.
940420	0.087	> 0.15	0.1147	0.0258	0.787	0.0279	0.0382	12
940615	0.086	> 0.15	0.1036	0.0390	-0.580	0.0283	0.0421	n.a.
940817	0.092	0.133	0.1020	0.0333	0.335	0.0274	0.0468	n.a.
940921	0.088	> 0.15	0.1091	0.0392	0.127	0.0278	0.0490	n.a.
941019	0.080	> 0.15	0.1067	0.0386	-0.720	0.0279	0.0503	n.a.
950118	0.084	> 0.15	0.1031	0.0329	0.754	0.0272	0.0587	n.a.
950215	0.083	> 0.15	0.1123	0.0367	0.582	0.0266	0.0589	n.a.
950322	0.089	> 0.15	0.0987	0.0304	0.786	0.0259	0.0588	3
950419	0.094	0.116	0.1068	0.0366	0.783	0.0254	0.0576	n.a.
950621	0.086	> 0.15	0.0993	0.0434	0.474	0.0242	0.0557	n.a.
960214	0.083	> 0.15	0.1225	0.0513	-0.740	0.0220	0.0492	1
960619	0.082	> 0.15	0.1233	0.0607	-1.156	0.0225	0.0522	66
960717	0.079	> 0.15	0.1190	0.0423	1.247	0.0223	0.0525	44

D-statistics is Kolmogorov-Smirnoff test statistics for normality, p-value is for rejection of H_0 for normality, σ is the estimated standard deviation of the estimated market risk premium θ, T-bill rate is 3-month T-bill rate, $z = [\mu_{SP} - \mu]/\sigma$, where μ_{SP} and μ are, respectively, the realized and the estimated expected S&P returns throughout the option life, n represents the number of found violations of the Proposition 1 upper bound for a given date.

Table 4

Characteristics of Data Used to Search for Violations of the Proposition 1 Upper Bound for 45-day to Expiration CME S&P 500 European Call Options

Date	D-stat.	p-value	σ	$\theta - \sigma^2/2$	z	Yield	T-bill Rate	n
900103	0.0885	> 0.15	0.1676	0.0573	-1.400	0.0332	0.0786	n.a.
900705	0.0881	> 0.15	0.1558	0.0618	-1.630	0.0352	0.0792	n.a.
900808	0.0975	> 0.15	0.1747	0.0543	-1.467	0.0380	0.0763	1
901003	0.0902	> 0.15	0.1621	0.0377	0.238	0.0373	0.0739	11
901107	0.1067	> 0.15	0.1708	0.0341	1.277	0.0365	0.0731	94
901205	0.0980	> 0.15	0.1565	0.0438	0.029	0.0364	0.0720	10
910130	0.1125	0.113	0.1805	0.0480	1.351	0.0324	0.0637	n.a.
940302	0.0722	> 0.15	0.0979	0.0464	-1.356	0.0285	0.0353	15
940406	0.1045	> 0.15	0.1150	0.0383	0.260	0.0279	0.0365	n.a.
940504	0.1007	> 0.15	0.1175	0.0303	0.268	0.0277	0.0411	n.a.
940601	0.0560	> 0.15	0.1019	0.0368	-0.340	0.0283	0.0427	1
940907	0.0883	> 0.15	0.1168	0.0363	-0.427	0.0278	0.0465	n.a.
941005	0.1111	0.124	0.1152	0.0295	0.340	0.0279	0.0508	n.a.
941207	0.0634	> 0.15	0.1059	0.0258	0.710	0.0282	0.0583	28
950104	0.1065	> 0.15	0.1181	0.0288	1.002	0.0272	0.0586	n.a.
950308	0.0969	> 0.15	0.1074	0.0352	1.241	0.0259	0.0592	1
950405	0.1030	> 0.15	0.1132	0.0335	0.565	0.0254	0.0583	n.a.
950705	0.1047	> 0.15	0.1086	0.0419	0.431	0.0239	0.0564	n.a.
950906	0.1091	0.140	0.1032	0.0405	0.687	0.0231	0.0548	n.a.
951004	0.1125	0.113	0.1108	0.0350	0.698	0.0226	0.0544	n.a.
951206	0.1135	0.104	0.0978	0.0387	-0.534	0.0225	0.0546	2
960103	0.0829	> 0.15	0.1064	0.0494	0.962	0.0213	0.0518	n.a.
960131	0.1114	0.122	0.1194	0.0463	0.066	0.0220	0.0505	n.a.
960306	0.0871	> 0.15	0.1012	0.0523	-0.482	0.0219	0.0501	44
960403	0.0736	> 0.15	0.1023	0.0626	0.333	0.0211	0.0513	11
960508	0.0953	> 0.15	0.1100	0.0553	0.695	0.0213	0.0511	15
960605	0.0998	> 0.15	0.0994	0.0633	-1.953	0.0225	0.0521	16
960703	0.0622	> 0.15	0.0962	0.0646	-0.554	0.0223	0.0522	7

D-statistics is Kolmogorov-Smirnoff test statistics for normality, p-value is for rejection of H_0 for normality, σ is the estimated standard deviation of the estimated market risk premium θ, T-bill rate is 3-month T-bill rate for a given date, $z = [\mu_{SP} - \mu]/\sigma$, where μ_{SP} and μ are, respectively, the realized and the estimated expected S&P returns throughout the option life, n represents the number of found violations of the Proposition 1 upper bound for a given date.

Table 5

Number and Proportion to All Observed Quotes of Violations of the Proposition 1 Upper Bound

	Set of Options	
	1	2
Quotes, Total	1592	1032
Quotes, Original	654	545
Quotes, Recovered	938	487
Violations, Total	147	209
Pr, Total	0.092	0.203
Pr, Original	0.086	0.217
Pr, Recovered	0.097	0.187

The descriptive statistics regarding the moneyness and the violation size in Table 6 below do not indicate significant differences between the sets. The large mean of the violation size for the Set 1 is an effect of a few far-out-the-money outliers. Generally, the violation size is decreasing in the moneyness, which is presented in Figure 3 below, in which the violation size is measured as the logarithm of $C_0/\overline{C}_0$. This finding is consistent with the volatility smile literature since far from the money calls generally exhibit larger violation size[44] than at- or in-the-money options.

Table 6

Mean Size of Violation of the Proposition 1 Upper Bound and Mean Moneyness

	Panel A: Set 1				
	N	Mean	Std Dev.	Minimum	Maximum
$C_0/\overline{C}_0$	147	0.907	6.911	0.005	83.876
S_0/K	147	0.971	0.023	0.876	1.015
	Panel B: Set 2				
	N	Mean	Std Dev.	Minimum	Maximum
$C_0/\overline{C}_0$	209	0.240	0.294	0.002	1.759
S_0/K	209	0.970	0.031	0.877	1.061

[44] Recall that the Proposition 1 upper bound is a constant volatility model.

Table 7

Representative Sample of Options

Case	Date	σ	$\theta - \sigma^2/2$	Yield	T-bill Rate	Index	K
LL_1	960619	0.1233	0.0607	0.0225	0.0522	66404	71000
LH_1	960717	0.1190	0.0423	0.0223	0.0525	63471	67500
HL_1	940420	0.1147	0.0258	0.0279	0.0382	44071	44500
HH_1	960717	0.1190	0.0423	0.0223	0.0525	63612	64000
LL_2	941207	0.1059	0.0258	0.0282	0.0583	45058	48000
LH_2	901003	0.1621	0.0377	0.0373	0.0739	31140	35000
HL_2	941207	0.1059	0.0258	0.0282	0.0583	45259	45500
HH_2	901107	0.1708	0.0341	0.0365	0.0731	30755	31000
Case	C_0	$\overline{C}_0$	$C_0/\overline{C}_0$	S_0/K	C_0/S_0	$\underline{\lambda}$	$\overline{\lambda}$
LL_1	70	63.17	1.208	0.935	0.001	2.287	5.823
LH_1	140	65.61	2.134	0.940	0.002	1.641	3.266
HL_1	530	526.32	1.007	0.990	0.012	0.905	1.694
HH_1	1170	991.34	1.280	0.994	0.018	1.641	3.266
LL_2	70	65.58	1.067	0.939	0.002	1.735	3.664
LH_2	65	29.92	2.173	0.890	0.002	0.541	1.268
HL_2	860	830.20	1.036	0.995	0.019	1.735	3.664
HH_2	1030	881.07	1.269	0.992	0.033	0.326	1.085
Case	$(\underline{\lambda}+\overline{\lambda})/2$	β	y	x_m	x_u	R_m^0	R_u^0
LL_1	4.055	0.580	79685	19651	13684	1.0580	1.0601
LH_1	2.454	0.218	76165	31043	23321	1.0473	1.0506
HL_1	1.300	0.124	52885	40697	31219	1.0146	1.0157
HH_1	2.454	0.218	76334	31112	23372	1.0218	1.0231
LL_2	2.700	0.247	54070	20029	14757	1.0498	1.0525
LH_2	0.905	0.157	37368	41313	29470	1.0625	1.0719
HL_2	2.700	0.247	54311	20119	14823	1.0238	1.0249
HH_2	0.706	0.132	36906	52312	34015	1.0236	1.0270

σ is the estimated standard deviation of the estimated market risk premium θ, T-bill rate is 3-month T-bill rate. Index represents S&P 500 prices contemporaneous to observed violations of the Proposition 1 upper bound. $\underline{\lambda}$ and $\overline{\lambda}$ are, respectively, the upper and the lower bound of the NT region, β is the optimal consumption rate. The risk aversion coefficient γ is -6. y is the dollar value of the Index in a portfolio at time zero containing 1.2 shares of the Index, x_m (x_u) is the dollar value of the riskless asset at time zero in a portfolio satisfying at time zero $y/x = (\underline{\lambda}+\overline{\lambda})/2$ ($y/x = \overline{\lambda}$). R_m^0 (R_u^0) is the critical portfolio return as derived by (3.4.9) for a portfolio satisfying at time zero $y/x = (\underline{\lambda}+\overline{\lambda})/2$ ($y/x = \overline{\lambda}$).

We now turn to the numerical results for the improvement in expected utility. Table 7 above presents the data and partial results for a representative sample of options. Graphs of the improvements in expected utility plotted in Figure 4 and Figure 5 below indicate that for high-moneyness options violating the Proposition 1 upper bound (cases HL and HH), the maximum possible improvement in expected utility far exceeds this derived from adopting the zero-net-cost in low-moneyness calls. This expected result appears also in potential SSD relations; for instance, Case HH_1 in Figure 4 below, Panel (a) clearly exhibits a SSD over all the other cases. On the other hand, the graph of the improvement in expected utility in case HL_1 crosses both graphs of the low moneyness cases.

Note that Case HH_2 in Figure 4, Panel (a) does not second order dominate Case HL_2 since the graphs cross. For an increase of the risky to riskless asset proportion, however, Case HH_2 exhibits SSD over Case HL_2, as observed in Fig 5, Panel (b). We explain this result by a change in the critical return R^0 as derived by (3.4.9), due to a change in the asset proportion. For the midpoint portfolio, R^0 is larger for Case HL_2 than for Case HH_2 (1.0238 to 1.0236, respectively), as seen in Table 7 while this relation becomes reversed at the upper boundary of the NT region with the corresponding values for R^0 respectively of 1.0249 and 1.0270. Note that for Case HH_1 clearly dominating Case HL_1, the critical portfolio return R^0 is larger than for Case HL_1 irrespective of the assets proportion. By annualizing the critical portfolio return R^0 (3.4.9) we obtain a value that corresponds to the portfolio rate of return where the improvement in expected utility becomes a

decreasing function. For instance, these annualized values of R^0 for Cases HL_2 and HH_2 are, respectively, 0.2020 and 0.2190 for a portfolio satisfying at time zero $y/x = \bar{\lambda}$. These analytical values are similar to the simulated portfolio rates of return that can be read from the Fig 5, Panel (b).

As seen in Table 7, the critical portfolio return R^0 increases for all the cases in the representative sample of options for an increase in the risky to riskless asset proportion λ. This corresponds to a positive change in the improvement in expected utility as λ increases, derived by model (3.4.2) and represented in Fig 6. Note that the magnitude of this change is significantly larger for the options representing the high-moneyness cases. The results presented in Table 7 and Figure 6 provide evidence that an investor whose portfolio exhibits a higher risky to riskless asset proportion has more of an incentive to adopt the zero-net-cost policy. We do not provide results for a critical value for the number of extra shares n^{λ} as derived by (3.5.1) bought with the proceeds from writing one call option; this is because this value proved to be negative for all the cases, which means an automatic increase in the improvement of expected utility for an increase in λ. We hypothesise that we could obtain a meaningful critical value for n^{λ} for far-in-the-money calls, which are missing in our sample[45]

Figure 7 presents an example of a change in the improvement in expected utility due to a change in wealth. The improvement decreases in proportion to increases in wealth. Figure 8 represents the loss in expected utility due to transaction costs. The low

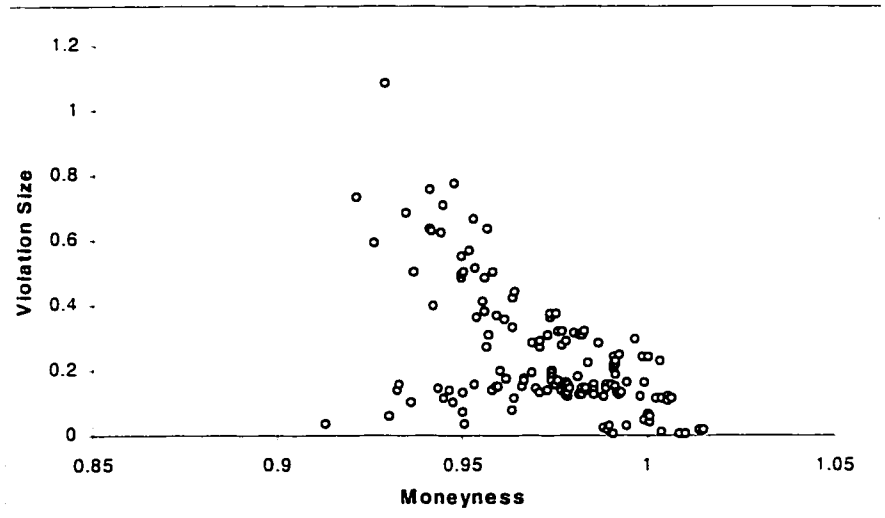
[45] The highest observed moneyness for either sample is app. 1.01.

magnitude of this loss provides justification to the analytical approximate results (3.4.8), (3.4.9) and (3.5.1). Finally, the low magnitude of the difference in the improvements in expected utility due to increased trading frequency presented in Fig 9 provides evidence that this improvement is relatively insensitive to the trading frequency. This result is expected, since any decrease in the improvement in expected utility of the J-investor might occur only because of an increase in the proportion of the risky to riskless asset proportion λ due to the inclusion of the extra shares at time zero. Since this increase is small, the results for the improvement in expected utility show little change at the higher trading frequency.

Figure 3

The Size of Violations of the Proposition 1 Upper Bound versus Moneyness. The X-axis represents the S/K ratio, the Y-axis represents $\log(C_0 / \overline{C}_0)$, where C_0 is the observed call price, and $\overline{C}_0$ is the Proposition 1 upper bound. Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days.

(a)



(b)

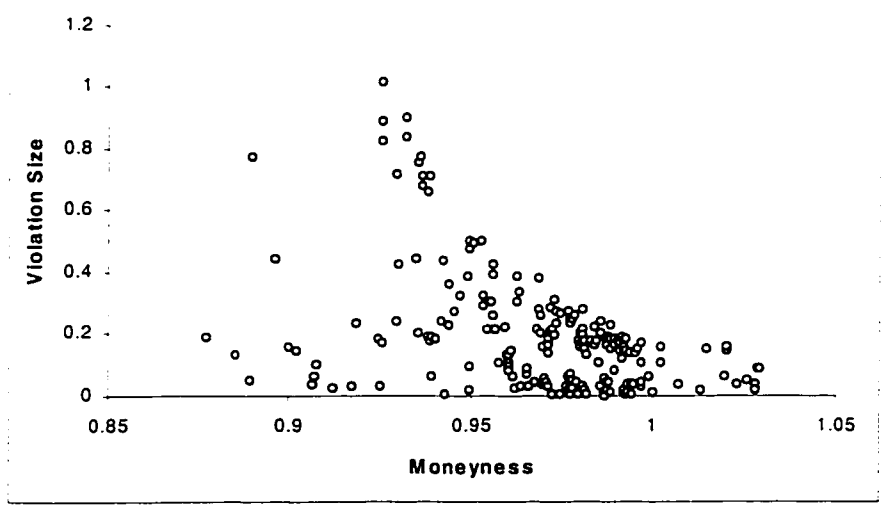


Figure 4

The improvement in expected utility $\hat{H}(z)$ resulting from adopting the zero-net-cost position in a call option violating the Proposition 1 upper bound for the representative set of options for portfolios satisfying $y/x = (\underline{\lambda} + \overline{\lambda})/2$ estimated by models (3.4.2) and (3.4.6). Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days. The original number of shares N_0 is 1.2. X-axis represents the observed portfolio return simulated by the Monte Carlo model (3.2.3).

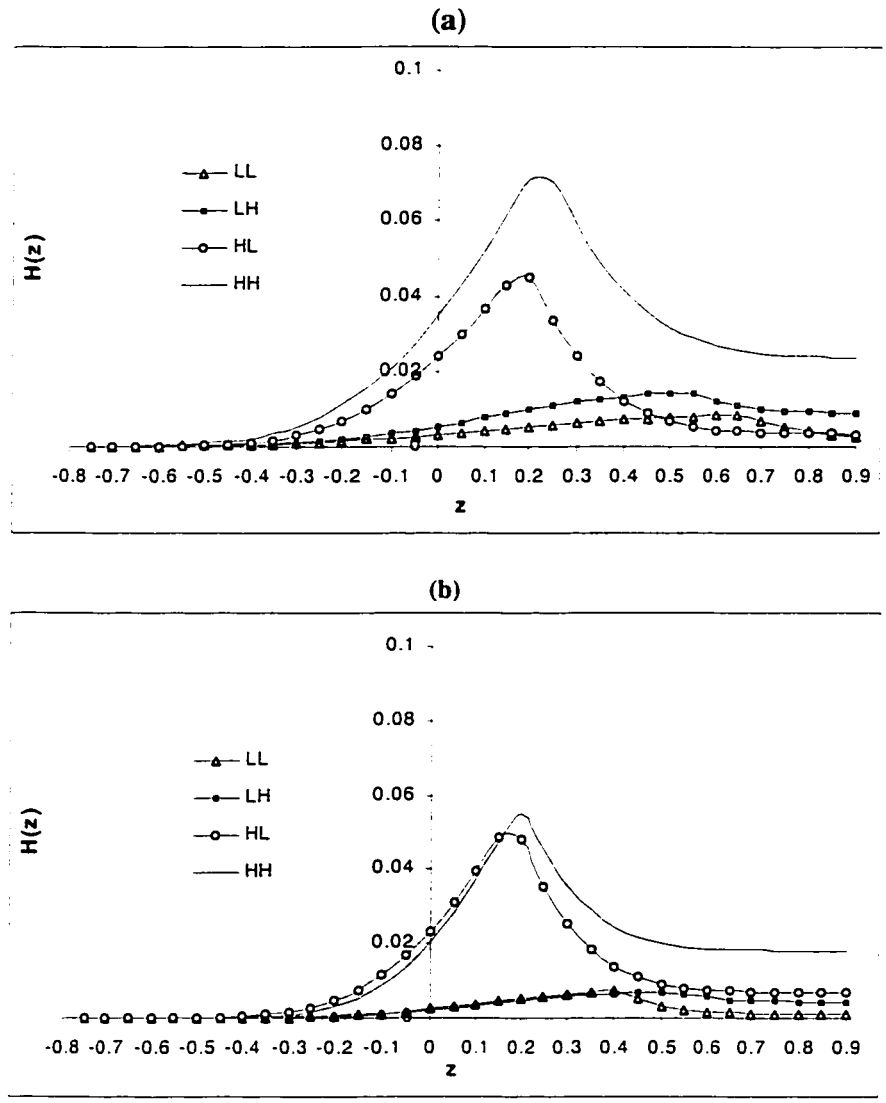


Figure 5

The improvement in expected utility $\hat{H}(z)$ resulting from adopting the zero-net-cost position in a call option violating the Proposition 1 upper bound for the representative set of options for portfolios satisfying y/x = $\bar{\lambda}$ estimated by models (3.4.2) and (3.4.6). Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days. The original number of shares N_0 is 1.2. X-axis represents the observed portfolio return simulated by the Monte Carlo model (3.2.3).

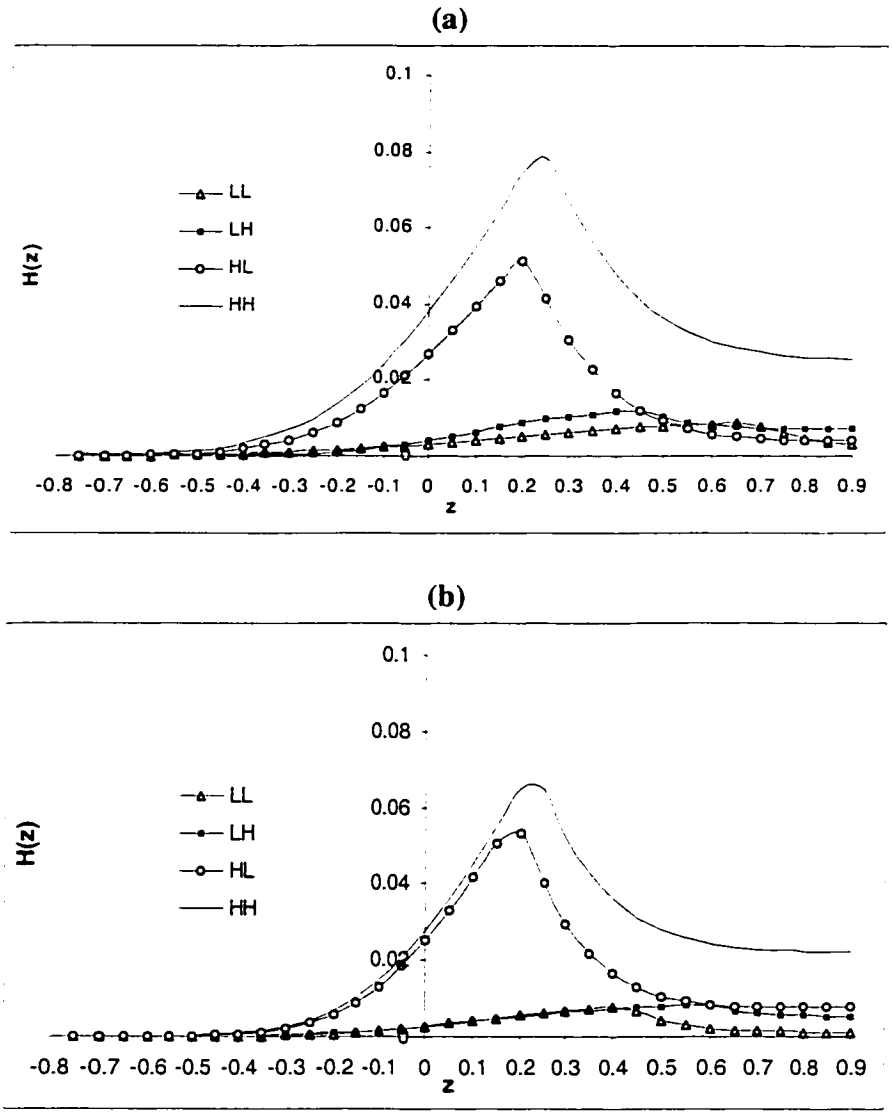


Figure 6

The difference in the improvements in expected utility $\hat{H}(z)$ resulting from adopting the zero-net-cost position in a call option violating the Proposition 1 upper bound for the representative set of options between portfolios satisfying $y/x = (\underline{\lambda}+\overline{\lambda})/2$ and $y/x = \overline{\lambda}$ estimated by models (3.4.2) and (3.4.6). Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days. The original number of shares N_0 is 1.2. X-axis represents the observed portfolio return simulated by the Monte Carlo model (3.2.3).

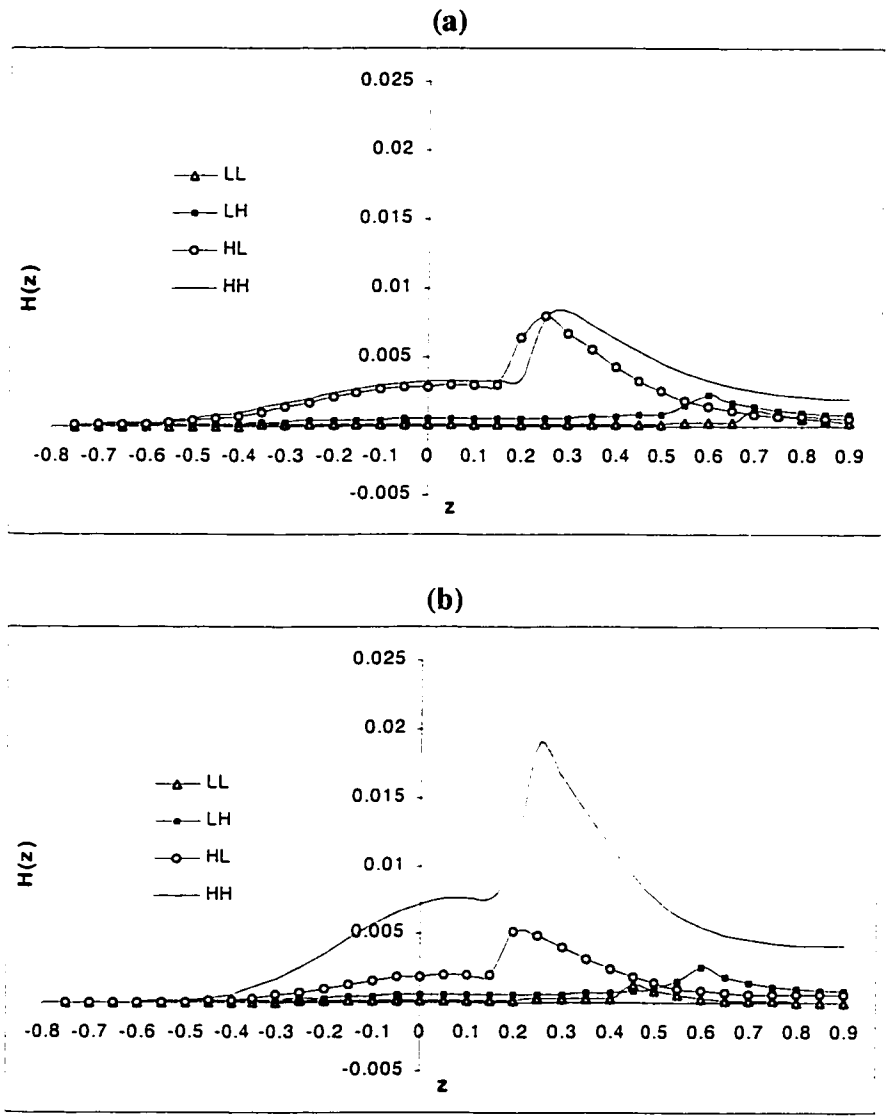
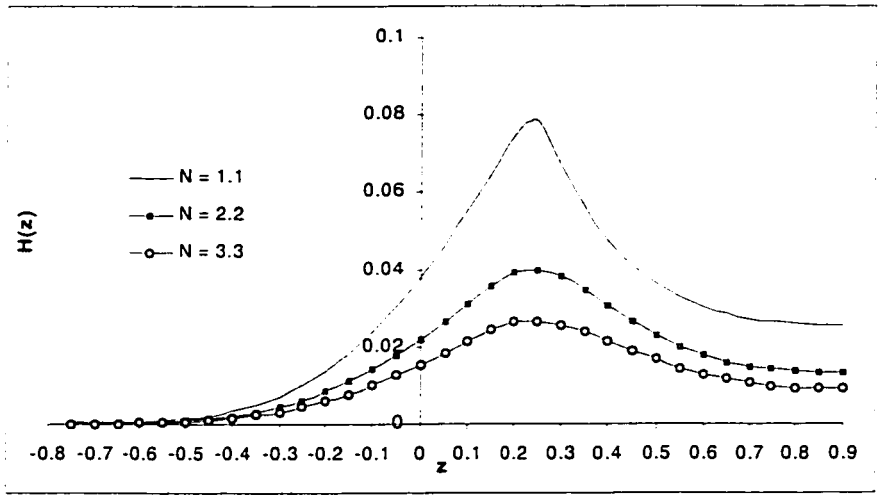


Figure 7

The impact of a change in the original number of shares N_0 on the improvement in expected utility $\hat{H}(z)$ resulting from adopting the zero-net-cost position in a call option violating the Proposition 1 upper bound for the representative set of options for portfolios satisfying $y/x = \bar{\lambda}$ estimated by models (3.4.2) and (3.4.6). Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days. The selected case represents high moneyness and high violation. X-axis represents the observed portfolio return simulated by the Monte Carlo model (3.2.3).

(a)



(b)

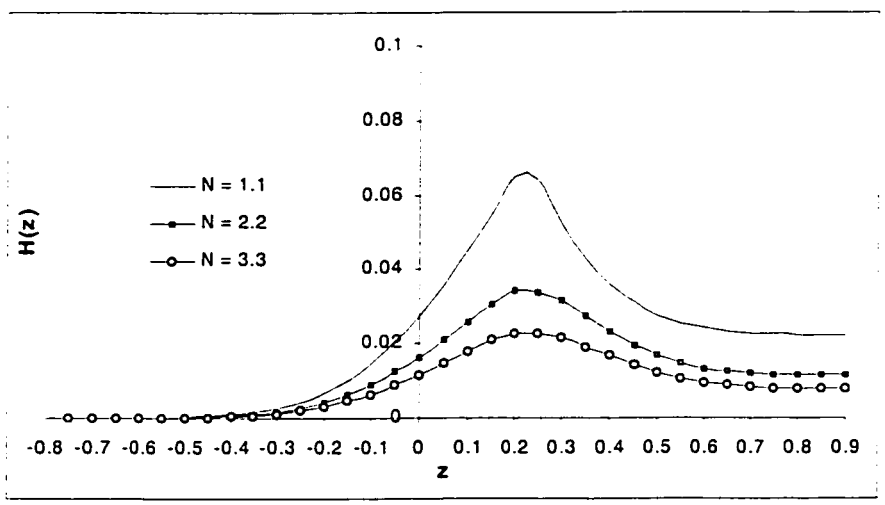


Figure 8

The loss in expected utility $\hat{H}(z)$ due to transaction costs caused by portfolio revisions for the representative set of options for portfolios satisfying $y/x = \bar{\lambda}$ estimated from the from the difference in the cumulative distribution functions of models (3.4.2) and (3.4.7) by model (3.4.6). Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days. The original number of shares N_0 is 1.2. X-axis represents the observed portfolio return simulated by the Monte Carlo model (3.2.3).

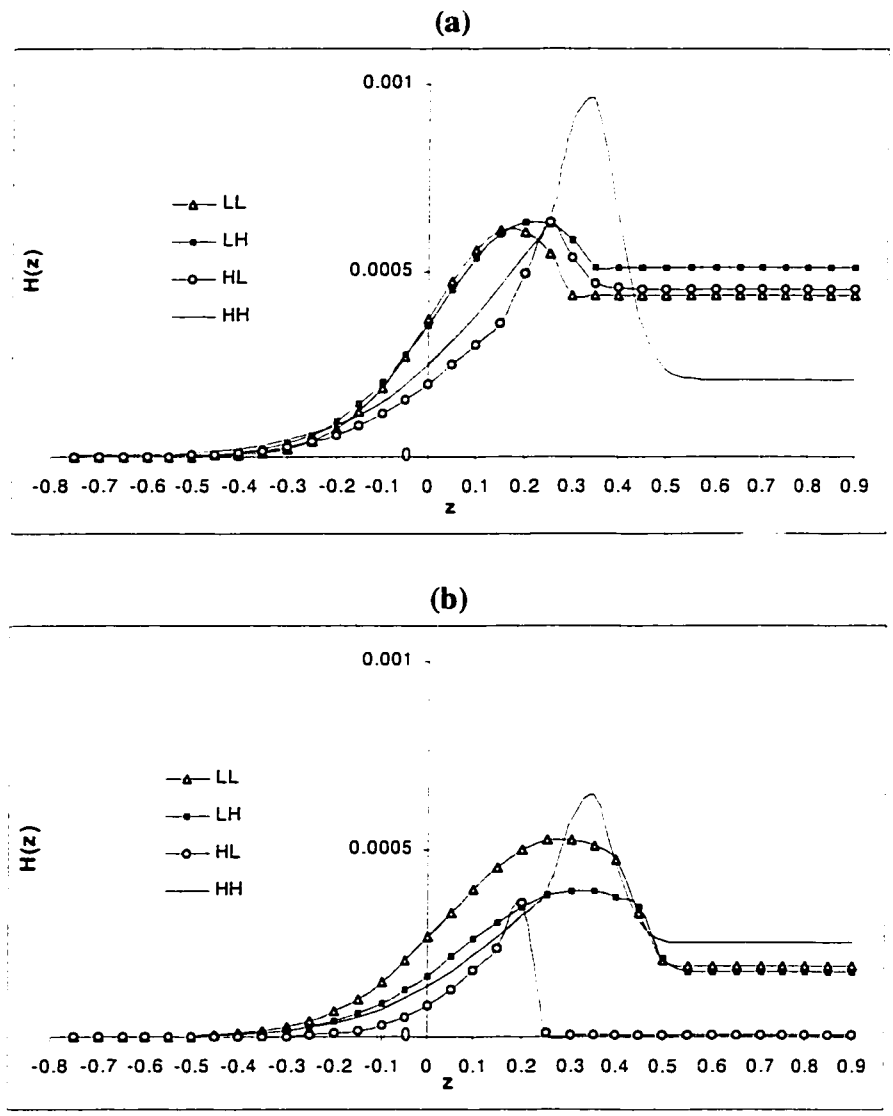
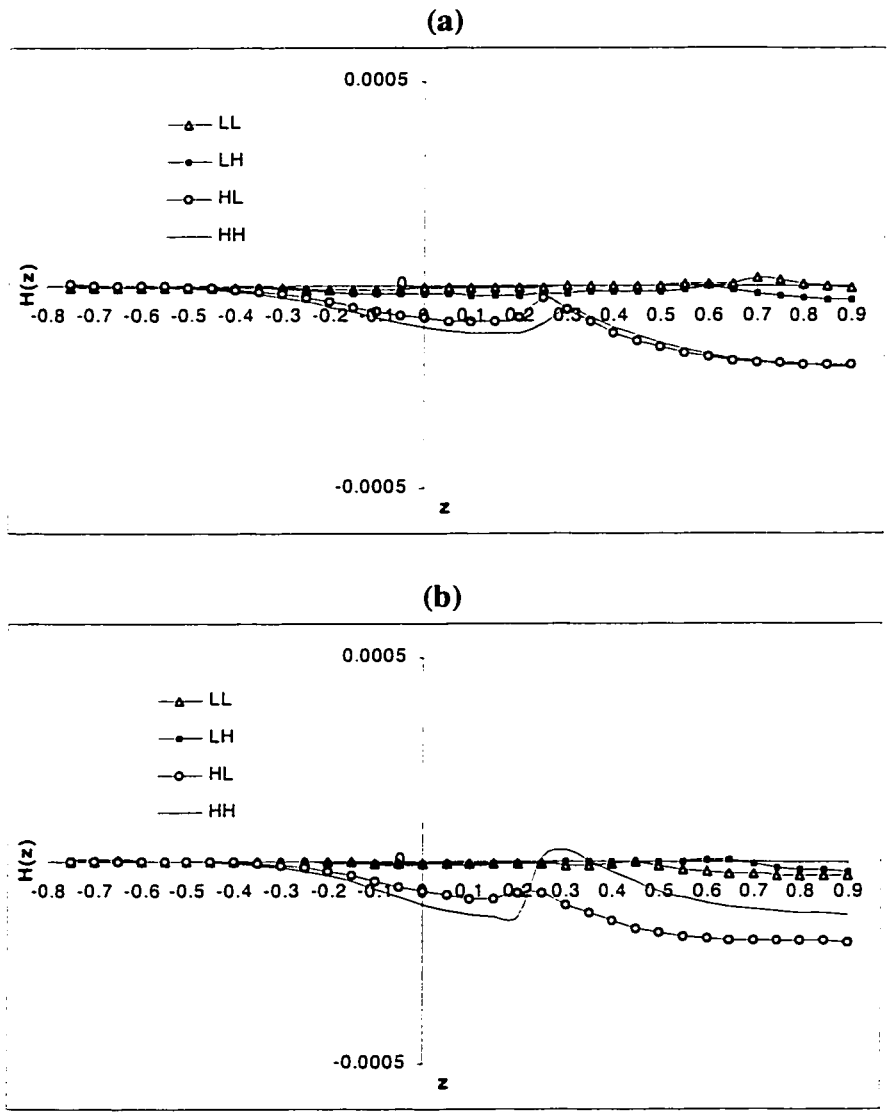


Figure 9

The difference in the improvements in expected utility $\hat{H}(z)$ resulting from adopting the zero-net-cost position in a call option violating the Proposition 1 upper bound for the representative set of options between portfolios satisfying $y/x = \bar{\lambda}$ for an increase in the portfolio revision frequency from 1 to 9, estimated by models (3.4.2) and (3.4.6). Panels (a) and (b) represent, respectively, options maturing in 31 and 45 days. The original number of shares N_0 is 1.2. X-axis represents the observed portfolio return simulated by the Monte Carlo model (3.2.3).



5. Concluding Remarks

The results presented in this thesis confirm that the Constantinides-Perrakis model does not impose "much" stronger restrictions on the wealth of the call writer than the well-established finding that hedging one derivative in the presence of transaction costs requires holding the same number of units of the underlying asset. Our empirical findings provide evidence that the stochastic dominance argument in the Proposition 1 upper bound results in the traditional SSD relation in the presence of dynamic trading in discrete time. The result of a clear second order stochastic dominance of the J-investor holdings over the V-investor for options priced almost exactly at the Proposition 1 upper bound raises a question if the bound can be further tightened; for instance, by specifying the J-value function.

The CP methodology represents a pricing mechanism since every risk-averse investor whose wealth satisfies the monotonicity condition[46] agrees with its results. As such, it can be applied to pricing index options in emerging markets, for which no organized derivatives markets exist.

The research presented in this thesis has apparent extensions. First and foremost, the methodology presented here can be applied to the lower bound on the European put (CP Proposition 6), which is also a stochastic dominance result. Second, some modifications

[46]The condition imposed on wealth is not restrictive since financial intermediation removes concerns regarding a "sizable" wealth necessary to engage in trading described in this thesis.

are also worth exploring; for instance, the sensitivity of both the empirical and analytical results presented here to an increase in the transaction costs rate can be examined.

The Constantinides-Perrakis (2002) model can accommodate other processes than the lognormal that was assumed in this; mixed jump-diffusion, GARCH or stochastic volatility. For those processes, however, the improvement in expected utility cannot be quantified with theV-function used in this thesis. New theoretical work w. r. to the value function is required to derive numerical results under the above processes of the underlying asset.

The method used to derive the improvement in expected utility is applicable without modifications whenever the anticipated distribution is lognormal, regardless of whether it is the past data that is used to derive the distribution estimates. For instance, this method can be applied if superior private information, such as analyst recommendations exists, provided that lognormality can be reasonably assumed.

References

Bensaid, B., J-P. Lesne, H. Pagés and J. Scheinkman, 1992. "Derivative Asset Pricing with Transaction Costs." *Mathematical Finance* 2, 63-86.

Black, Fischer, 1986. "Presidental Address: Noise." *Journal of Finance* 41, 529-544.

Black, F. and M. S. Scholes, 1973. "The Pricing of Options and Corporate Liabilities." *Journal of Political Economy* 81, 637-654.

Boyle, P. P. and T. Vorst, 1992. "Option Replication in Discrete Time with Transaction Costs." *Journal of Finance* 47, 271-293.

Constantinides, G. M., 1979. "Multiperiod Consumption and Investment Behavior with Convex Transactions Costs." *Management Science* 25, 1127-37.

Constantinides, G. M., 1986. "Capital Market Equilibrium with Transactions Costs." *Journal of Political Economy* 94, 842-862.

Constantinides, G. M., and S. Perrakis, 2002. "Stochastic Dominance Bounds on Derivative prices in a Multiperiod Economy with Proportional Transaction Costs", *Journal of Economic Dynamics and Control* 26, 1323-1352.

Constantinides, G. M., and S. Perrakis, 2002a. "Stochastic Dominance Bounds on American Option Prices in Markets with Frictions", working paper.

Constantinides, G. M. and T. Zariphopoulou, 1999. "Bounds on Prices of Contingent Claims in an Intertemporal Economy with Proportional Transaction Costs and General Preferences." *Finance and Stochastics* 3, 345-369.

Constantinides, G. M. and T. Zariphopoulou, 2001. "Bounds on Derivative Prices in an Intertemporal Setting with Proportional Transaction Costs and Multiple Securities." *Mathematical Finance* 11, 331-346.

Cox, J. C., S. Ross and M. Rubinstein, 1979. "Option Pricing: A Simplified Approach." *Journal of Financial Economics* 7, 229-263.

Cvitanic, J. and I. Karatzas, 1996. "Hedging And Portfolio Optimization Under Transaction Costs: A Martingale Approach." *Mathematical Finance* 6, 133-165.

Cvitanić, J. H. Pham and N. Touzi, 1999. "A Closed-form Solution to the Problem Of Super-Replication under Transaction Costs." *Finance and Stochastics* 3, 35-54.

Davis, M. H. A. and A. R. Norman, 1990. "Portfolio Selection with Transaction Costs." *Mathematics of Operations Research* 15, 676-713.

Davis, M. H. A., V. G. Panas and T. Zariphopoulou, 1993. "European Option Pricing with Transaction Costs." *SIAM Journal of Control and Optimization* 31, 470-493.

Dumas, B. and E. Luciano, 1991. "An Exact Solution To A Dynamic Portfolio Choice Problem Under Transactions Costs." *Journal of Finance* 46, 577-596.

Fleming W. L. and R. W. Rishel. *Deterministic and Stochastic Optimal Control* (Berlin, Heidelberg, New York: Springer-Verlag, 1975).

Hadar, J. and W. R. Russell, 1969. "Rule For Ordering Uncertain Prospects." *American Economic Review* 59, 25-34.

Hanoch, G. and H. Levy, 1969. "The Efficiency Analysis Of Choices Involving Risk." *Review of Economic Studies* 36, 335-346.

Hull, J. C. *Options, Futures, & Other Derivatives* (Upper Saddle River: Prentice-Hall, 2003).

Jackwerth, J. C., 2000. "Recovering Risk Aversion from Option Prices and Realized Returns." *Review of Financial Studies* 13, 433-451.

Kocherlakota, N. R., 1996. "The Equity Premium: It's Still a Puzzle." *Journal of Economic Literature* 34, 42-71.

Kroll, Y. and H. Levy, 1979. "Stochastic Dominance With A Riskless Asset: An Imperfect Market." *Journal of Financial and Quantitative Analysis* 14, 179-204.

Leland, H. E., 1985. "Option Pricing and Replication with Transactions Costs." *Journal of Finance* 40, 1283-1301.

Levy, H., 1985. "Upper and Lower Bounds of Put and Call Option Value: Stochastic Dominance Approach." *Journal of Finance* 40, 1197-1217.

Levy, H., 1992. "Stochastic Dominance end Expected Utility: Survey and Analysis." *Management Science* 38, 555-593.

Magill, M. J. P., and G. M. Constantinides, 1976. "Portfolio Selection with Transactions Costs." *Journal of Economic Theory* 13, 245-263.

Mehra, R. and E. C. Prescott, 1985. "The Equity Premium: A Puzzle," *Journal of Monetary Economics* 15, 145-162.

Merton, R. C., 1969. "Lifetime Portfolio Selection under Uncertainty: The Continuous-Time Case", *Review of Economics and Statistics* 51, 247-257.

Merton, R. C., 1971. "Optimum Consumption and Portfolio Rules in a Continuous-Time Model", *Journal of Economic Theory* 3, 373-413.

Merton, R. C., 1973. "Theory of Rational Option Pricing." *Journal of Economics and Management Science* 4, 141-183.

Merton, R. C., 1989. "On the Application of the Continuous-Time Theory of Finance to Financial Intermediation and Insurance." *The Geneva Papers on Risk and Insurance* 14, 225-261.

Perrakis, S., 1986. "Option Bounds in Discrete Time: Extensions and the Pricing of the American Put." *Journal of Business* 59, 119-141.

Perrakis, S., 1988. "Preference-free Option Prices when the Stock Return Can Go Up, Go Down, or Stay the Same." *Advances in Futures and Options Research* 3, 209-235.

Perrakis, S., and J. Lefoll, 1997. "Derivative Asset Pricing with Transaction Costs: an Extension." *Computational Economics* 10, 359-376.

Perrakis, S., and J. Lefoll, 2000. "Option Pricing and Replication with Transaction Costs and Dividends." *Journal of Economic Dynamics and Control* 24, 1527-1561.

Perrakis, S., and P. J. Ryan, 1984. "Option Pricing Bounds in Discrete Time." *Journal of Finance* 39, 519-525.

Ritchken, P. H., 1985. "On Option Pricing Bounds." *Journal of Finance* 40, 1219-1233.

Ritchken, P.H. and S. Kuo, 1988. "Option Bounds with Finite Revision Opportunities." *Journal of Finance* 43, 301-308.

Rubinstein, M., 1976. "The valuation of Uncertain Income Streams and the Pricing of Options." *Bell Journal of Economics and Management Science* 7, 407-425.

Zariphopoulou, T., 1999. "Comment on 'The Valuation of Contingent Claims under Portfolio Constraints: Reservation Buying and Selling Prices'." *European Finance Review* 3, 389-392.

Appendices

Appendix 1: Maple Code for the Derivation of the NT Region.

```
restart;r:=.0393:q:=r:m:=.0845:v:=.0148:d:=.0302:g:=-6:
k:=0.005:
eq1:=(1+k)*(1+a1*(g-s1)*l1^s1+a2*(g-
s2)*l1^s2)=a1*s1*l1^(s1-1)+a2*s2*l1^(s2-1):
eq2:=(1-k)*(1+a1*(g-s1)*l2^s1+a2*(g-
s2)*l2^s2)=a1*s1*l2^(s1-1)+a2*s2*l2^(s2-
1):a:=solve({eq1,eq2},{a1,a2}):
a1:=subs(a,a1): a2:=subs(a,a2):
a1_l1:=diff(a1/(q-g*(r-b+h)),l1):
a1_l2:=diff(a1/(q-g*(r-b+h)),l2):
val:=(b^g)*(1/g+a1*(l^s1)+a2*(l^s2))/(q-g*(r-
b+h)):val_b:=diff(val,b):
l:=((m+d-r)/((1-g)*2*v))/(1-((m+d-r)/((1-g)*2*v)));h:=l*d:
s2,s1:= solve(v*s^2+(m-v-r+b-h)*s-(q-g*(r-b+h))=0,s):
sol:=fsolve({a1_l1,a1_l2,val_b},{l1=.5*l,l2=1.5*l,b=.15},{l
1=0.3*l..0.90*l,l2=1.2*l..2.2*l,b=.07..0.45});val_b;
```

Appendix 2: SAS Code Simulating Stock Paths, SAS Code Deriving Estimates for the Likelihood of the Monotonicity violation.

```
libname out 'u:\mczerwonko\thesis\results';
data check;
do const=1 to 500;
output;end; run;
data check; set check;
call symput('m',trim(left(put(const,8.))));run;
%put m= ***&m***;
%macro seed;
%do i=1 %to &m;
seed&i=abs(int(120076573*rannor((&i+1)**2+&i**3))-
int(383745455*rannor(&i*abs(&i-&i**2))));
%end;%mend;
%macro rand;
%do j=1 %to &m;
```

```
call rannor(seed&j,x&j);
x&j=abs(int(x&j*10000000));
%end;%mend;
data out.random1;keep k x1-x&m;
%seed;
do k=1 to 300;
%rand;
output;
end;
run;
%do k=1 %to &n;
call rannor(s&k, x&k);
%end; %mend;
%macro paths;
%do i=1 %to &nn;
data temp; keep x%eval(&i+0); set tmp1.random1(obs=&n);
call symput('seed'||left(_n_),trim(left(x&i+0)));run;
%put seed3= ******&seed3*****;
data path_%eval(&i);keep x1-x&n;
s1   =    &seed1    ;
/*More seeds*/
s405 =    &seed405  ;
%rand1;
output;
end;
run;
%end; %mend; %paths; run;


/* Code for estimating Pr of violations of the monotonicity
assumption*/
/*# of paths used.*/
%let NP=50000;
/*N calls*/
%let NC=128;
/*N of max revisions to maturity*/
%let NDAYS=132;
libname out 'u:\mczerwonko\thesis\newjob\final';
%macro plain;
/*Computations of bounds/(1+k1) for all days for all
calls.*/
%let str=%sysevalf(1000/&moneyness);
%let xx=%sysevalf(&N*1000/&l1);
/*Computation of bounds*/
%do k=0 %to &NDAYS;
c=(&NDAYS-&k+1)/(3*365);x= &std*sqrt(c);
```

```
d1=( log(s&k/&str) + (.11*(1+.02/(1+&kk))-.02+&std**2)*c
)/x; d2= d1 - x;
nd1= cdf('GAUSS',d1); nd2= cdf('GAUSS',d2);
e= s&k*exp(-1*0.02*c); f= &str*exp(-
1*.11*(1+.02/(1+&kk))*c);
b&k= (e*nd1 - f*nd2)/(1-&kk);
%end;
/*beginning values for bond account, N shares, counting
function*
f0=0;
N=&N;
xx=&xx;
%do j=%eval(3/&nrev) %to &NDAYS %by %eval(3/&nrev);
/*Derivative condition computation*/
cond= (N*s%eval(&j-3/&nrev) + b%eval(&j-
3/&nrev))/s%eval(&j-3/&nrev)-1/(1-&kk);
temp1= N*s&j;
temp2= (xx*exp(0.05/(&nrev*365))+N*s&j*.02/(&nrev*365))*(1-
&beta/(&nrev*365));
check= temp1/temp2;
if check gt &ll then do;
temp3=(temp1 -&ll*temp2)/(s&j*(1+&ll*(1-&kk)));
N= N - temp3;
xx= temp2 + s&j*(1-&kk)*temp3;
end;
else do;
xx= temp2;
end;
/*Counting function (cf) assumes the value of 1 if the
previous period cf was 1*/
if f0=1 then f1=f0;
/* If not, the other conditions are checked*/
else do;
/* If both deriv. cond and N shares are violated, cf
assumes 1*/
if N le 1/(1-&kk) & cond le 0 then f1=1;
/* if the previous period cf was 0, and either derivative
or N shares was satisfied,
cf assumes 0*/
else f1=0;
end;
f0=f1;
%if &j=90 %then %do;
f31=f1; N31=N;%end;
%end;
f=f1;
%mend;
```

```
%macro total;
%do ii=1 %to &NC;
/* ingesting inputs*/
data temp; set out.mon_inputs;
if ll=. then delete;
if _n_=&ii then do;
call symput('std',trim(left(std)));
call symput('kk',trim(left(kk)));
call symput('ll',trim(left(ll)));
call symput('N',trim(left(n)));
call symput('nrev',trim(left(nrev)));
call symput('moneyness',trim(left(moneyness)));
call symput('beta',trim(left(beta)));
end;
run;
/* launching macros computing Pr*/
%if &std=0.2 %then %let vvv=highvol; %else %let vvv=lowvol;
data t&ii; keep p31 nn31 p nn; set out.stock_&vvv;
%plain;
nn31 + N31;
p31 + f31;
nn + N;
p + f;
if _n_=&NP then do;
p31=p31/&NP;
nn31=nn31/&NP;
p = p/&NP;
nn=nn/&NP;
output;
end;
run;
%end;
%mend;
%total; run;
%macro handle;
%do i=1 %to &NC;
t&i
%end;
%mend;
data out.mon; set %handle;
data out.mon; merge out.mon_inputs out.mon;run;
run;
```

Appendix 3: SAS Code Recovering and Assigning Bid/Ask Flags and Matching Option and Index Quotes, SAS Code Searching for Violations of the Proposition 1 Upper Bound.

```
LIBNAME out1 'u:\mczerwonko\thesis\raw\res31';
%MACRO rawdata;
%DO i=0 %TO 6;
DATA ind9&i;
INFILE "u:/mczerwonko/thesis/raw/ind9&i..csv" DLM=','
FIRSTOBS=2 MISSOVER;
INPUT date itime ihour imin isec ival;
i_time= (ihour*60 + imin)*60 + isec;
mins= ihour*60 + imin;
PROC SORT data=ind9&i;
BY date mins;RUN;
DATA sel31_9&i;
INFILE "u:/mczerwonko/thesis/raw/op9&i..csv" DLM=','
FIRSTOBS=2 MISSOVER;
INPUT date otime hour min sec mat ab_cme $1. strike price
cp $1. ;
DROP temp1-temp3;
temp1= lag(price);temp2=lag(date);temp3=lag(strike);
IF (ab_cme=',' & date= temp2 & strike= temp3) THEN DO;
IF (price GT temp1) THEN ab='A';
IF (price LT temp1) THEN ab='B'; END;
IF ab_cme NE "," THEN n_cme + 1;
IF ab_cme NE ',' THEN ab= ab_cme; ELSE ab=ab;
IF ab NE " " THEN count + 1;
IF ab=" " THEN DELETE;
IF cp="C" & ab="A" THEN DELETE;
IF cp="P" & ab="B" THEN DELETE;
o_time= (hour*60 + min)*60 + sec;
mins= hour*60 + min;
lagmins= mins -1;
leadmins= mins +1; RUN;
%END; %MEND; %rawdata; RUN;
%MACRO name66(name= , num= );
%DO k=1 %TO &num;
&name&k %END; %MEND;
%MACRO catch22;
%DO i=0 %TO 6;
%IF &i=0  %THEN     %LET n=   2008;
%ELSE %IF &i=1      %THEN     %LET n=   410;
%ELSE %IF &i=2 %THEN     %LET n=   1714;
```

```
%ELSE %IF &i=3 %THEN     %LET n=   1747;
%ELSE %IF &i=4 %THEN     %LET n=   2054;
%ELSE %IF &i=5 %THEN     %LET n=   2009;
%ELSE %IF &i=6 %THEN     %LET n=   1541;
%DO j=1 %TO &n;
DATA temp&j ; SET sel31_9&i(FIRSTOBS=&j OBS=&j);
DATA ptemp&j; SET temp&j; mins=lagmins;
DATA ftemp&j; SET temp&j; mins=leadmins;
DATA temp&j; SET ptemp&j temp&j ftemp&j;
PROC SORT; BY date mins;
DATA temp&j; MERGE temp&j ind9&i; BY date mins;
IF price NE .; diff= ABS(i_time - o_time);
PROC SORT; BY o_time;
PROC RANK DATA= temp&j out= rtemp&j ties=low; BY o_time;
VAR diff; RANKS punks; RUN;
DATA rtemp&j; SET rtemp&j; BY o_time; WHERE punks=1;
IF otime=lag(otime) THEN delete;
%END;
DATA out1.put_call9&i; KEEP date otime mat cp strike price
ival diff;
SET %name66(name= rtemp, num= &n); RUN;
%END; %MEND; %catch22; RUN;


libname job 'u:\mczerwonko\thesis\newjob';
libname final 'u:\mczerwonko\thesis\newjob\final';
%macro viol;
%do i=1 %to 2;
%if &i=1 %then %let id=31; %else %let id=45;
DATA a&i;keep date price ival strike bound viol moneyness
dyield rint3m m1 std1 ab z;
set job.call&id.sel;
iadj= ival*exp(-1*dyield*&id/365);
d1= (log(ival/strike) + (1/.995-1)*dyield*&id/365 + mt1 +
rint3m*&id/365 + stdt1**2)/stdt1;
d2= d1 - stdt1;
N1= CDF('GAUSS', d1);
N2= CDF('GAUSS', d2);
bound= (iadj*N1 - N2*strike*EXP(-1*((1/.995)*dyield*&id/365
+ mt1 + rint3m*&id/365 + stdt1**2)))*(1.005/.995);
check= bound - price;
IF check LT 0 THEN viol = abs(check)/bound; ELSE viol=0;
moneyness= ival/strike;
DATA final.viol&id; set a&i;if viol gt 0;
if date ne lag(date) then id +1;if ab='B' then ncme+1;
proc sort; by date;
```

```
data final.charv&id; keep date nviol dyield rint3m m1 std1
z;set final.viol&id; by date;
if first.date then nviol=0;
nviol+1;
if last.date then output;
%end;
%mend;
%viol; run;
```

Appendix 4: SAS Code Deriving the Improvement in Expected Utility for the Portfolio at the Upper bound of the NT Region.

```
/*#of paths*/
%let NP=200000;
/*number of shares held*/
%let ntrials=3;
/*Libraries used for input/output.*/
libname job 'u:\mczerwonko\thesis\newjob';
libname final 'u:\mczerwonko\thesis\newjob\final';
/* Macro computing portfolio revisions and final returns.*/
%macro job;
N= &N + &fr;
NV= &N;
cj=0;cv=0;
%let coeff= %sysevalf((&N+&fr)/&l2+&N*.995);
xx= (&N+&fr)*s0/&l2;
xxv= xx;
%do j=1 %to &nn;
/*J-investor portfolio dynamics*/
temp1= N*s&j;
temp2= (xx*&r + N*&divday)*(1-&beta/365);
cj=cj*&r + (xx*&r + N*&divday)*&beta/365;
check= temp1/temp2;
if check gt &l2 then do;
temp3=(temp1 - &l2*temp2)/(s&j*(1+&l2*.995));
N= N - temp3;
xx= temp2 + s&j*.995*temp3;
end;
else do;
xx= temp2;
end;
/*V-investor portfolio dynamics*/
```

```
temp4= NV*s&j;
temp5= (xxv*&r + NV*&divday)*(1-&beta/365);
cv=cv*&r + (xxv*&r + NV*&divday)*&beta/365;
check1= temp4/temp5;
if check1 gt &l2 then do;
temp6=(temp4 - &l2*temp5)/(s&j*(1+&l2*.995));
NV= NV - temp6;
xxv= temp5 + s&j*.995*temp6;
end;
else do;
xxv= temp5;
end;
%end;
jret=365/&nn*log((xx + cj + .995*N*s&nn+&divday-
max(s&nn+&divday-&str,0))/(s0*&coeff));
vret=365/&nn*log((xxv + cv +
.995*NV*s&nn+&divday)/(s0*&coeff));
temp=jret - vret;
tempv= (&N-NV)/&N;
tempj= (&N+&fr-N)/(&N+&fr);
%mend;
/*Macro computing CDF's of the returns*/
%macro cdf;
data cdf&id._&sss; keep dnj dnv excess jcdf1-jcdf171 vcdf1-
vcdf171 iretj1-iretj171 iretv1-iretv171; set cdf;
excess+temp;
dnv+tempv;
dnj+tempj;
%do i=1 %to 171;
if jret le (&i-81)*.01 then jcdf&i + 1;
%end;
%do i=1 %to 171;
if jret gt (&i-81)*.01-.005 & jret lt (&i-81)*.01+.005 then
do;
iretj&i+iret; nj&i+1; end;
%end;
%do i=1 %to 171;
if vret le (&i-81)*.01 then vcdf&i + 1;
%end;
%do i=1 %to 171;
if vret gt (&i-81)*.01-.005 & vret lt (&i-81)*.01+.005 then
do;
iretv&i+iret; nv&i+1; end;
%end;
if _n_=&NP then do;
excess=excess/&NP;
dnv=dnv/&NP;
```

```
dnj=dnj/&NP;
%do i=1 %to 171;
jcdf&i = jcdf&i/&NP;
iretj&i = iretj&i/nj&i;
vcdf&i = vcdf&i/&NP;
iretv&i = iretv&i/nv&i;
%end;
output; end;
%mend;
/*Macro deploying a loop over all the above macros, and
reading in input data.*/
%macro total;
%do id=1 %to 8;
data a; set final.sample;
divday=dyield*endind/365;
r=1+rint3m/365;
if _n_=&id then do;
call symput('nn',trim(left(id2)));
call symput('r',trim(left(r)));
call symput('divday',trim(left(divday)));
call symput('l1',trim(left(l1)));
call symput('l2',trim(left(l2)));
call symput('str',trim(left(strike)));
call symput('fr',trim(left(fr)));
call symput('beta',trim(left(beta)));
end; run;
%do sss=1 %to &ntrials;
%let N=%sysevalf(1.2*&sss);
data cdf;set final.both_&id(obs=&NP);
%job;
run;
%cdf;
%end;
%end;
%mend;
%total;run;
%macro m;
%do sss=1 %to &ntrials;
data final.up&sss;
set %do id=1 %to 8;
cdf&id._&sss %end;;
%end; %mend; %m; run;
/*Macro selecting each fifth observation to produce
graphs*/
%macro kkk;
%do i=1 %to 34;
v%eval(1+5*&i)
```

```
%end;
%do i=1 %to 34;
j%eval(1+5*&i)
%end;
%do i=1 %to 34;
d%eval(1+5*&i)
%end;
%do i=1 %to 34;
d%eval(1+5*&i)
%end;
%do i=1 %to 34;
i%eval(1+5*&i)
%end;
%do i=1 %to 34;
iretj%eval(1+5*&i)
%end;
%do i=1 %to 34;
iretv%eval(1+5*&i)
%end;
%mend;
/*Macro deriving the integral condition for the utility
improvement*/
%macro int;
%do i=1 %to 171;
v&i=vcdf&i;
%end;
%do i=1 %to 171;
j&i=jcdf&i;
%end;
%do i=1 %to 171;
d&i=v&i-j&i;
%end;
a1=d1/2;
i1=a1;
%do i=2 %to 171;
a&i=(d&i + d%eval(&i-1))*0.005;
i&i=sum(of a1-a&i);
%end;
%mend;
%macro short;
%do j=1 %to 8;
%do k=1 %to &ntrials;
data qqq&k; keep %kkk;set cdf&j._&k;
%int;run;
proc transpose data=qqq&k out=a&k prefix=v&j._&k;run;
%end;
data aa&j; merge _null_
```

```
%do k=1 %to &ntrials;
a&k %end;;
%end;
data final.up_revised;
merge _null_
%do i=1 %to 8;
aa&i %end;;
%mend;
%short;run;
```