# MODELLING LIQUIDITY AND ITS EFFECTS

J. D. Evans, L. C. G. Rogers and Surbjeet Singh

July 25, 2005

## 1 Introduction

After credit risk, liquidity risk is probably the next most important risk faced by the finance industry, and yet the study of liquidity is comparatively under-developed. This may be in part due to the fact that there is no agreed definition of what liquidity is, even in qualitative terms; everyone would agree that the effect of illiquidity is to make it costly or difficult to trade large volumes of the underlying asset in small times, but there are different approaches to modelling this. There is for example the approach of Longstaff (2001) who takes a conventional Black-Scholes model for the asset, and supposes that the portfolio process must be differentiable with bounded derivative. Another approach (see for example Frey & Stremme (), Frey (2000), Schönbucher & Wilmott ()) takes the market price of the asset to be a function of certain fundamentals, and the amount of the asset held by a 'large' investor. Such models should be considered as models of feedback effects, rather than models of liquidity effects, since they suffer a number of disadvantages in modelling of liquidity[1]. Another contribution of a quite different nature is the work of Çetin, Jarrow & Protter (2003).

Our viewpoint here is to start from a simple discrete-time microeconomic story, and try to derive the dynamics of an agent's wealth process from that. Dividing the time axis into equal intervals of length $\Delta t$, we suppose that the number of shares offered for sale or purchase in each interval is of the order of $\Delta t$. An agent who wishes to buy a large number of shares in any one time interval will have to pay an inflated price, because he is trading in a shallow market. On the other hand, the amount he actually trades, itself of order $\Delta t$, is small relative to the total number of shares in the firm, so will have negligible impact on the price of the shares in the next period. What transpires is that the impact of liquidity modelled in this fashion is like a transaction cost, but not one which is proportional to the amounts traded. In the conventional proportional transaction cost model (see Magill & Constantinides (1976), Davis & Norman (1990)) the bid-offer spread (the difference between the lowest offer price and highest bid price) is proportional to the current asset price, but the depth of the limit order book is

[1] It is hard to see how such models might deal with the presence of many 'large' agents, each pursuing his own hedging objectives with different time horizons (with or without knowledge of the other participants' intentions?). These feedback models typically present the solution to a hedging problem in feedback form, exhibiting the hedge as a function of time, and current stock price - but if the initial portfolio is not at the exactly correct value, it is not clear how it is to be moved to that value.

not taken into account. Our approach is to model illiquidity as a transaction cost taking into account the depth of the book and the trade size; this is the subject of Section 2.

As might be expected, once such illiquidity costs are introduced there is little hope of exact solutions to any of the natural questions one might ask. However, provided the illiquidity costs are small, we may hope to derive a reasonably close approximation to the solution of any problem. We illustrate this philosophy in Section 3, where we see how the results of the Merton investment/consumption model are modified in the presence of small illiquidity costs. We start with a brief summary of the classical Merton model where the utility is constant relative risk aversion. Then we use some scaling arguments to simplify the HJB but a closed form solution seems impossible. Instead, since the Merton model is the limiting case of our model as the illiquidity parameter $\varepsilon$ goes to zero we look for a series expansion for the optimal solution. We derive the first correction term and compare the results to numerical calculations.

As another application of our approach to modelling liquidity, in Section 3 we shall investigate the cost of a European put option in such a market, together with the associated hedging strategy. Since the market is incomplete once liquidity effects are taken into account, there is no unique replication price, so the first issue is to decide how the price is to be defined. The approach taken here is to use the utility-indifference price proposed by Hodges & Neuberger (), Davis (), for an investor with constant absolute risk-aversion utility. This is a rational approach to pricing, though not without its limitations[2]. Exact solution is again impossible, but once again it is possible to derive an expansion for the solution, and we find the leading-order form of the correction, and the hedging strategy.

## 2 The Discrete Time Model

We shall consider a single asset, whose price at time $t$ is denoted by $S_t$. Fixing for the moment some $\Delta t > 0$, we let $p_n$ denote $\log S(n\Delta t)$, which we suppose evolves as a random walk,

$$p_n = p_{n-1} + \xi_n, \tag{2.1}$$

where the $\xi_n$ are independent and identically distributed. We imagine that in each period $\lambda_a \Delta t$ shareholders consider whether to sell their shares, and $\lambda_b \Delta t$ agents consider whether to buy shares. A hedger also comes to the market with the intention of buying $\Delta H_n$ shares during the $n^{\text{th}}$ period $((n-1)\Delta t, n\Delta t] \equiv (t_{n-1}, t_n]$. The log-price $\tilde{p}_n$ at which he will trade is determined by the equalisation of supply and demand:

$$\lambda_a \Delta t\, f_s(\tilde{p}_n - p_n) = \lambda_b \Delta t\, f_d(\tilde{p}_n - p_n) + \Delta H_n, \tag{2.2}$$

where the supply function $f_s$ is continuous and strictly increasing, the demand function $f_d$ is continuous and strictly decreasing. We therefore find that the log-price at which the hedger trades is determined by

$$\tilde{p}_n - p_n = \psi\left(\frac{\Delta H_n}{\Delta t}\right), \tag{2.3}$$

[2]For example, the price assigned to a European call option will be infinite.

where $\psi$ is the inverse function to $x \mapsto (\lambda_a f_s(x) - \lambda_b f_d(x))\Delta t$. It is natural to suppose that $\psi(0) = 0$.

If we now let $H_t$ denote the number of shares held by the hedger at time $t$, and $K_t$ denote the amount of cash held by the hedger at time $t$, then the wealth[3] of the hedger at time $t$ is $w_t = H_t S_t + K_t$. The change in wealth over the $n^{\text{th}}$ period is therefore

$$\begin{aligned} w_{t_n} - w_{t_{n-1}} &= H_{t_{n-1}}(S_{t_n} - S_{t_{n-1}}) + S_{t_n}\Delta H_n - e^{\tilde{p}_n}\Delta H_n \\ &= H_{t_{n-1}}(S_{t_n} - S_{t_{n-1}}) - S_{t_n} h_n(\exp(\psi(h_n)) - 1)\Delta t, \end{aligned}$$

where we use the notation $h_n = \Delta H_n/\Delta t$. If we let $\Delta t \downarrow 0$, and suppose that $H$ is differentiable, with derivative $h_t = dH_t/dt$, then we derive the (continuous-time) dynamics for wealth in the form

$$\begin{aligned} dH_t &= h_t dt, \\ dw_t &= H_t dS_t - h_t S_t f(h_t) dt, \end{aligned} \tag{2.4}$$

where $f(x) \equiv \exp(\psi(x)) - 1$ is continuous and increasing, equal to 0 at 0.

We shall therefore now concentrate on the continuous-time dynamics (2.4), and let the discrete-time construction by which we arrived at it slip into the background. Various forms for $f$ can be considered; since we intend to treat the illiquidity costs as small, we shall introduce the small parameter $\varepsilon$ into the argument of $f$, leading to wealth dynamics

$$dw_t = H_t dS_t - h_t f(\varepsilon h_t) S_t dt. \tag{2.5}$$

We shall always assume that $f$ is continuous, increasing, bounded below by -1, and $f(0) = 0$.

# 3 The Merton problem

## 3.1 Review of the Merton problem without liquidity effects

To establish notation, we briefly review the classical Merton problem (see Merton()). When liquidity effects are taken into account, we will see a perturbation of the solution recalled here.

An investor may invest in two assets, a money market account with constant interest rate $r$, and a share with price process $(S_t)_{t\geq 0}$ satisfying

$$dS_t = S_t(\sigma dW_t + \mu dt) \tag{3.1}$$

for constants $\sigma$ and $\mu$, where $(W_t)_{t\geq 0}$ is a standard Brownian motion.

The investor chooses to hold the value $\theta_t$ in shares and to consume at a rate $C_t$ so that his wealth evolves as

$$dw_t = (rw_t - C_t)dt + \theta_t(\sigma dW_t + (\mu - r)dt). \tag{3.2}$$

[3] We are talking here of the *paper* or *nominal* wealth at time $t$, which is not the same as the *liquidation value* of the hedger's portfolio.

Subject to the constraint $w_t \geq 0$ for all $t$, the investor's objective is to achieve

$$v_0(w) = \sup \mathbb{E}\left[ \int_0^\infty e^{-\rho t} U(C_t) dt \,\middle|\, w_0 = w \right] \tag{3.3}$$

where $\rho$ is some positive constant. If the utility has the form $U(x) = x^{1-R}/(1-R)$ for some $R > 0$ different from 1, Merton finds

$$v_0(w) = \gamma^{-R} U(w), \tag{3.4}$$

$$\theta_t = \pi w_t, \tag{3.5}$$

$$C_t = \gamma w_t, \tag{3.6}$$

where

$$\pi = \frac{\mu - r}{\sigma^2 R} \tag{3.7}$$

$$\gamma = \frac{\rho + (R-1)(r + (\mu - r)^2/2R\sigma^2)}{R} \tag{3.8}$$

$$= \frac{\rho + (R-1)(r + \frac{1}{2}\sigma^2 R \pi^2)}{R}. \tag{3.9}$$

In order that the problem is well posed, we shall need that $\gamma > 0$, which is only an issue if $0 < R < 1$.

## 3.2 The Merton problem in an illiquid market

Using (2.5) we have the following equations for the evolution of the asset and the wealth of the investor:

$$dS_t = S_t(\mu dt + \sigma dW_t)\,, \tag{3.10}$$

$$dw_t = r w_t dt + H_t(dS_t - S_t r dt) - C_t dt - h_t f(\varepsilon h_t) S_t dt\,. \tag{3.11}$$

In this situation the investor controls $C_t$ and $h_t$ to achieve[4]

$$V(w, H, S) \equiv V(w, H, S; \varepsilon) = \sup \mathbb{E}\left[ \int_0^\infty e^{-\rho t} U(C_t) dt \,\middle|\, w_0 = w, H_0 = H, S_0 = S \right], \tag{3.12}$$

with the restriction $w_t \geq 0$ for all $t$. Because the share price can with positive probability rise or fall arbitrarily high or low in arbitrarily short time, this positive wealth constraint implies that the investor can never short shares or cash;

$$H_t \geq 0, \quad w_t - H_t S_t \equiv K_t \geq 0.$$

Because of this requirement, we shall make the

$$\textsc{Assumption:} \quad 0 < \pi \equiv (\mu - r)/\sigma^2 R < 1$$

[4]We will usually omit the explicit dependence of $V$ on $\varepsilon$ from the notation.

so that the optimal solution with liquidity effects can be treated as a perturbation of the optimal solution of the original Merton problem.

Given initial values for $H$, $w$ and $S$, we say that $(C, h)$ is *admissible* if $w_t \geq 0$ for all $t$. We can exploit the scaling in (3.11), (3.10) and (3.12) to give

$$\begin{aligned} V(w, H, S) &= v(z, H)S^{1-R}, \\ C^*(w, H, S) &= c^*(z, H)S \end{aligned}$$

where $v(z, H) = V(z, H, 1)$, $c^*(z, H) = C^*(z, H, 1)$, $z = w/S$, and $C^*(w, H, S)$ is the optimal consumption rate as a function of current wealth, current holding of the share, and current share price. To see this note that a policy $(C, h)$ is admissible for starting values $H_0$, $w_0$ and $S_0$ if and only if $(\lambda C, h)$ is admissible for starting values $H_0$, $\lambda w_0$ and $\lambda S_0$ where $\lambda > 0$.

It appears therefore that the variable $z_t \equiv w_t/S_t$ is important for the analysis. It is of course the wealth of the agent expressed in units of the share, and by routine calculations using Itô's formula we are able to derive

$$dz_t = \sigma(H_t - z_t)dW_t - \Big\{ (\sigma^2 + r - \mu)(H_t - z_t) + h_t f(\varepsilon h_t) + c_t \Big\}dt, \tag{3.13}$$

where $c_t \equiv C_t/S_t$. Even more simply, if we introduce the non-negative process $Y_t \equiv z_t - H_t = K_t/S_t$, we have

$$dY_t = Y_t\Big( -\sigma dW_t + (\sigma^2 + r - \mu)dt \Big) - \Big\{ h_t + h_t f(\varepsilon h_t) + c_t \Big\}dt, \tag{3.14}$$

a simple linear SDE for $Y$. This suggests that we re-express the value function as a function of $Y$ and $H$:

$$V(w, H, S) = v(z, H)S^{1-R} = F(Y, H)S^{1-R}, \tag{3.15}$$

where we define $F(Y, H) = v(H+Y, H)$. The next stage of the analysis is of course to derive the Hamilton-Jacobi-Bellman (HJB) equation; once again, routine but lengthy calculations using Itô's formula lead us to the conclusions

$$\sup_{c,h} \left\{ U(c) - \tilde{\rho} F + \frac{1}{2}\sigma^2 Y^2 F_{YY} - (h + hf(\varepsilon h) + c + \alpha Y)F_Y + hF_H \right\} = 0, \tag{3.16}$$

or equivalently

$$\sup_{c,h} \left\{ U(c) - \tilde{\rho} v + \frac{1}{2}\sigma^2 (H - z)^2 v_{zz} - (hf(\varepsilon h) + c + \alpha(z - H))v_z + hv_H \right\} = 0, \tag{3.17}$$

where $\alpha = \mu - r - \sigma^2 R = \sigma^2 R(\pi - 1)$ and $\tilde{\rho} = \rho + (R-1)(\mu - \frac{1}{2}\sigma^2 R) = R\gamma - \frac{1}{2}\sigma^2 R(R-1)(\pi-1)^2$. The maximisation over $c$ and $h$ is easily done, leading to

$$c^* = F_Y^{-1/R} = v_z^{-1/R}, \tag{3.18}$$

$$v_H = F_H - F_Y = v_z \bar{f}'(\varepsilon h^*) = F_Y \bar{f}'(\varepsilon h^*) \tag{3.19}$$

where $\bar{f}(t) \equiv tf(t)$. The optimal $c^*$ and $h^*$ can be substituted back into (3.17) and (3.16) to give the non-linear PDEs

$$0 = \tilde{U}(F_Y) - \tilde{\rho}F + \frac{1}{2}\sigma^2 Y^2 F_{YY} - \alpha Y F_Y + \frac{F_Y}{\varepsilon}\Phi\left(\frac{F_H - F_Y}{F_Y}\right) \tag{3.20}$$

$$0 = \tilde{U}(v_z) - \tilde{\rho}v + \frac{1}{2}\sigma^2 (H-z)^2 v_{zz} - \alpha(z-H)v_z + \frac{v_z}{\varepsilon}\Phi\left(\frac{v_H}{v_z}\right) \tag{3.21}$$

where $\tilde{U}(y) \equiv \sup_x\{U(x) - xy\}$ is the convex dual of $U$, and where the convex function $\Phi$ is defined by

$$\Phi(a) \equiv \sup_t\{at - tf(t)\}. \tag{3.22}$$

Notice that $\Phi(a) \geq \Phi(0) = 0$, and $\Phi(a) = +\infty$ if $a < -1$.

Exact solution of (3.21) or (3.20) appears impossible, but in the next two sections we show that progress may nevertheless be made, firstly by looking for a power-series solution (where the form (3.17) works best), and secondly by numerical solution (where the form (3.16) works best).

In this Section, we seek an expansion of the solution in powers of $\varepsilon$; first we present a simple heuristic which tells us what powers of $\varepsilon$ to be looking at.

We expect that for $\varepsilon$ small, the optimal behaviour will be to try to keep $H/z = HS/w$ close to $\pi_*$, that is, $H_t \doteqdot \pi_* z_t$. If the proportion of wealth in the risky asset in the original Merton problem were $p \neq \pi_*$, we would expect to be making a loss $O((p-\pi_*)^2)$; we would also expect that the main issue in making $H_t$ track $\pi_* z_t$ is to track the martingale part of $\pi_* z_t$. So if we consider the allied problem of finding a differentiable adapted process $X$ to achieve

$$\inf E\left[\int_0^T \left\{(W_t - X_t)^2 + \varepsilon \dot{X}_t^2\right\} dt\right]$$

we can solve this problem explicitly by setting $\xi_t \equiv W_t - X_t$, and solving the HJB equations for this problem; the solution for the value function is

$$V(\tau, \xi) = \sqrt{\varepsilon}\tanh(\tau/\sqrt{\varepsilon})\xi^2 + \varepsilon \log\cosh(\tau/\sqrt{\varepsilon})$$

expressed in terms of time-to-go $\tau = T - t$. The optimal control is $\dot{X}_t = \tanh(\tau/\sqrt{\varepsilon})\xi_t/\sqrt{\varepsilon}$, which is clearly $O(1/\sqrt{\varepsilon})$, and the value function itself is also seen to be $O(\sqrt{\varepsilon})$.

Treating $\varepsilon$ as a variable, we therefore seek a solution of the form

$$v(z, H; \varepsilon) = \sum_{k=0}^{\infty} \delta^k G_k(z, H) \tag{3.23}$$

where $\delta = \varepsilon^{1/2}$, and we make the dependence of $F$ on $\varepsilon$ explicit.

Now there is a further scaling of the wealth equation we can exploit, namely, that

$$v(\lambda z, \lambda H; \varepsilon/\lambda) = \lambda^{1-R} v(z, H; \varepsilon) \tag{3.24}$$

To see this, refer back to (3.14) and note that an admissible strategy $(C, h)$ with intial values $z_0$, $H_0$ and $\varepsilon$ is admissible if and only if the strategy $(\lambda C, \lambda h)$ with initial values $\lambda z_0$, $\lambda H_0$ and $\varepsilon/\lambda$ is admissible where $\lambda > 0$.

This scaling should be present in our power series solution so

$$\sum_{k=0}^{\infty} \frac{\delta^k}{\lambda^{k/2}} G_k(\lambda z, \lambda H) = \lambda^{1-R} \sum_{k=0}^{\infty} \delta^k G_k(z, H) \tag{3.25}$$

Now equating term by term we get

$$G_k(\lambda z, \lambda H) = \lambda^{1-R+k/2} G_k(z, H) \tag{3.26}$$

Put $g_k(x) = G_k(1, x)$ so that

$$G_k(z, H) = z^{1-R+k/2} G_k(1, H/z) = z^{1-R+k/2} g_k(H/z) \tag{3.27}$$

We have succeeded in reducing the dimension of the problem to 1. To proceed further, we perform an asymptotic analysis of the HJB equation, and obtain equations for the various functions $g_k$ in succession. We find that

$$g_0(x) = \gamma^{-R} U(x), \tag{3.28}$$

$$g_1(x) = -\gamma^{-R} \sqrt{\frac{\sigma^2 R}{2}} (\kappa_1 + (x - \pi)^2), \tag{3.29}$$

where

$$\kappa_1 = \frac{8\sigma^2 \pi^2 (1-\pi)^2}{3(4\gamma - \sigma^2(1-\pi)^2)} \tag{3.30}$$

(3.31)

An alternative approach is to attempt to express

$$v(z,H) \simeq v_0(z) + \delta g(z,H), \tag{3.32}$$

and then to obtain some bounds. Specifically, we intend to propose some explicit approximation $\bar{v} = v_0 + \delta g$ to the true value function, and then to examine

$$\mathcal{L}\bar{v} \equiv \tilde{U}(\bar{v}_z) - \tilde{\rho}\bar{v} + \alpha(H-z)\bar{v}_z + \frac{1}{2}\sigma^2(H-z)^2\bar{v}_{zz} + \frac{\bar{v}_z}{\varepsilon}\,\Phi\left(\frac{\bar{v}_H}{\bar{v}_z}\right). \tag{3.33}$$

Provided we can prove that there is a constant $\kappa$ such that $|\mathcal{L}\bar{v}(z,H)| \le \kappa\delta^2$ for all $z, H$, then we will be able to deduce that $\bar{v}$ is near to the true value function.

It is easy to verify that the Merton value function $v_0$ satisfies

$$\begin{aligned} 0 &= \tilde{U}(v_{0z}) - \tilde{\rho}v_0 + \alpha z(\pi-1)v_{0z} + \frac{1}{2}\sigma^2 z^2(\pi-1)^2 v_{0zz} \\ &= \gamma R v_0 - \tilde{\rho}v_0 + \alpha z(\pi-1)v_{0z} + \frac{1}{2}\sigma^2 z^2(\pi-1)^2 v_{0zz}, \end{aligned}$$

so returning to (3.33), we have

$$\begin{aligned} \mathcal{L}\bar{v} &= \tilde{U}(\bar{v}_z) - \tilde{\rho}\bar{v} + \alpha(H-z)\bar{v}_z + \frac{1}{2}\sigma^2(H-z)^2\bar{v}_{zz} + \frac{\bar{v}_z}{\varepsilon}\,\Phi\left(\frac{\bar{v}_H}{\bar{v}_z}\right) \\ &= \tilde{U}(v_{0z}+\delta g_z) - \tilde{U}(v_{0z}) - \delta\tilde{\rho}g + \delta\alpha z(p-1)g_z + \frac{1}{2}\delta\sigma^2 z^2(p-1)^2 g_{zz} \\ &\quad + \frac{v_{0z}+\delta g_z}{\varepsilon}\,\Phi\left(\frac{\delta g_H}{v_{0z}+\delta g_z}\right) + \alpha(H-\pi z)v_{0z} + \frac{1}{2}\sigma^2\{(H-z)^2 - z^2(\pi-1)^2\}v_{0zz} \\ &= \tilde{U}(v_{0z}+\delta g_z) - \tilde{U}(v_{0z}) - \delta\tilde{\rho}g + \delta\alpha z(p-1)g_z + \frac{1}{2}\delta\sigma^2 z^2(p-1)^2 g_{zz} \\ &\quad + \frac{v_{0z}+\delta g_z}{\varepsilon}\,\Phi\left(\frac{\delta g_H}{v_{0z}+\delta g_z}\right) - \gamma^{-R}z^{1-R}(p-\pi)^2\frac{\sigma^2 R}{2}, \end{aligned}$$

where we write $p = H/z$ for short. Now

$$\begin{aligned} |\tilde{U}(v_{0z}+\delta g_z) - \tilde{U}(v_{0z}) - \delta g_z\tilde{U}'(v_{0z})| &= \tilde{U}(v_{0z})\left|\left(1+\frac{\delta g_z}{v_{0z}}\right)^{(R-1)/R} - 1 - \frac{R-1}{R}\frac{\delta g_z}{v_{0z}}\right| \\ &\le \tilde{U}(v_{0z})C\delta^2 \end{aligned}$$

for some constant $C$ provided we have $|g_z/v_{0z}|$ bounded. Recalling that we want a bound of the form $|\mathcal{L}\bar{v}| \le \kappa\delta^2$, it is therefore enough to establish such a bound for the expression

$$\begin{aligned} Lv &\equiv \delta g_z\tilde{U}'(v_{0z}) - \delta\tilde{\rho}g + \delta\alpha z(p-1)g_z + \frac{1}{2}\delta\sigma^2 z^2(p-1)^2 g_{zz} \\ &\quad + \frac{v_{0z}+\delta g_z}{\varepsilon}\,\Phi\left(\frac{\delta g_H}{v_{0z}+\delta g_z}\right) - \gamma^{-R}z^{1-R}(p-\pi)^2\frac{\sigma^2 R}{2}\,. \\ &= \delta g_z\tilde{U}'(v_{0z}) - \delta\tilde{\rho}g + \delta\alpha z(p-1)g_z + \frac{1}{2}\delta\sigma^2 z^2(p-1)^2 g_{zz} \\ &\quad + \frac{v_{0z}+\delta g_z}{\varepsilon}\,\Phi\left(\frac{\delta g_H}{v_{0z}+\delta g_z}\right) - \frac{1}{2}\sigma^2 R(1-R)(p-\pi)^2 v_0(z). \end{aligned} \tag{3.34}$$

We now plan to represent $g$ as

$$g(z, H) = \frac{F_0(z,p) + f(z)}{1 + \delta G_1(z,p)} \tag{3.35}$$

and find out what forms of the functions $F_0$, $G_1$ and $f$ work well.

If we take $\Phi(x) = x^2/4$, which corresponds to taking $f(x) = x$, then the choice

$$F_0(z,p) = -\sigma\gamma^{-R}\sqrt{R/2}z^{3/2-R}(p-\pi)^2 \tag{3.36}$$

removes all terms that are $O(1)$ in (3.34). Turning next to terms that are $O(\delta)$, we obtain an expression in $z$, $p$ and the unknown functions $f$ and $G_1$ and various derivatives. Setting $p = \pi$, we find that this expression vanishes if and only if $f$ solves a second-order linear differential equation, whose general solution is

$$f(z) = -\frac{4\pi^2\sigma\gamma^{-R}\sqrt{2R}z^{3/2-R}}{3q} + A\gamma^{-R}z^{-R} + B\gamma^{-R}z^{3/2-R+q/2}, \tag{3.37}$$

where

$$q \equiv \frac{4\gamma}{\sigma^2(\pi-1)^2} - 1,$$

and $A$ and $B$ are constants to be determined. Assuming that $f$ has this form, the terms of order $\delta$ reduce to

$$\varphi_2(z,p)G_{1,p}(z,p) + \varphi_1(z,p)G_1(z,p) + \varphi_0(z,p), \tag{3.38}$$

where

$$\begin{aligned}
\varphi_2(z,p) &= (A + Bz^{(3+q)/2})(p-\pi)\sqrt{\frac{\sigma^2 R}{2z}} - \frac{\sigma^2 Rz(p-\pi)(3q(p-\pi)^2 + 8\pi^2)}{6q} \\
&= at - \frac{1}{2}\sigma^2 Rzt^3, \\
\varphi_1(z,p) &= -\sigma^2 Rz(p-\pi)^2 \\
&= -\sigma^2 Rzt^2,
\end{aligned}$$

where we have written $t = p - \pi$. For $\varphi_0$ we obtain a long expression which is a polynomial of degree 4 in $t$, with $t = 0$ as a root. This first-order linear ODE is solved by finding the integrating factor, which by inspection is $(p-\pi)^{-1}$, leading to the differential equation

$$\frac{\partial}{\partial p}\left((a - \frac{1}{2}\sigma^2 Rz(p-\pi)^2)G_1(z,p)\right) + \frac{\varphi_0(z,p)}{p-\pi} = 0.$$

---

## 3.3 Numerical Results

We use a Markov chain approximation to find an approximate numerical solution to the problem, based on the technique described in Kushner and Dupuis (1992). Then policy improvement is used to compute the optimal policy for this Markov chain approximation.

In the classical Merton solution the optimal holding is from (3.7) $H^* = \theta^* Z$ or in our transformed co-ordinates $H^* = \frac{\theta^*}{1-\theta^*} Y$. In our Markov chain the grid is scaled so that $\Delta h = \frac{\theta^*}{1-\theta^*} \Delta y$. Also for each $y_i$, $i \in \{1, 2, ..., n\}$ the point $(y_i, \frac{\theta^*}{1-\theta^*} y_i)$ is a grid point. In the plots below we take the parameter values as $\varepsilon = 0.1$, $\mu = 0.05$, $r = 0.03$, $\sigma = 0.4$, $R = 1/3$, $p = 0.04$, $ymin = 1/3$, $ymax = 2$, $hmin = 0$ and $hmax = \frac{\theta^*}{1-\theta^*} ymax + \Delta h$. We plot the percentage decrease in the value function as compared to the classical Merton solution after 10 iterations and for 4 successive grid refinements, starting with initial guesses for the optimal policy to be identically zero for $h$ and $C$.

Intuitively, it is clear that the optimal $h$ will be positive if $h_i < \frac{\theta^*}{1-\theta^*} y_j$ and negative if $h_i > \frac{\theta^*}{1-\theta^*} y_j$ so that we don't have to worry about the boundary conditions for extreme values of $h_j$ since the probability of a transition outside these values is zero. For the upper and lower values of cash, $ymin$ and $ymax$, respectively we take the power series with the first order correction as an approximate value. As we go through the policy iteration we expect the value function to converge to some value on the grid points. If the boundary conditions were exactly correct we would expect the value function to converge to the true value for the continuous time problem as we refine our grid further and further. Since our boundary conditions are only approximate it is difficult to give a quantitative estimate of how accurate the numerical calculation is. However, it can be seen from figure 5 that the numerical calculation and grid values from the power series solution with the first correction term are very close. In fact, the maximum difference is less than 0.1 %.

# References


[1] Frey, R.: *Market Illiquidity as a Source of Model Risk in Dynamic Hedging,* 2000.

[2] Kushner, H.J. and Dupuis, P.G.:*Numerical Methods for Stochastic Control Problems in Continous Time,* Volume **24** of Applications of Mathematics, Springer-Verlag, 1992.

[3] Longstaff, F.A.: *Optimal Portfolio Choice and the Valuation of Illiquid Securities,* 2001.

[4] Magill, M.J.P. and Constantinides, G. M.: *Portfolio selection with transaction costs,* J. Economic Theory **13** 245-263, 1976.

[5] Munk, Claus.:*Numerical Methods for Continous-Time, Continous-State Stochastic Control Problems,* 1997.

[6] Munk, Claus.:*The Markov Chain Approximation Approach for Numerical Solution of Stochastic Control Problems: Experiences from Merton's Problem,* 1998

[7] Rogers, L.C.G. and Williams D.:*Diffusions, Markov Processes and Martingales, vol. 2,* Cambridge University Press 2000.

[8] Rogers, L.C.G.:*The relaxed investor and parameter uncertainty,* Finance and Stochastics **5** 131-154, 2001

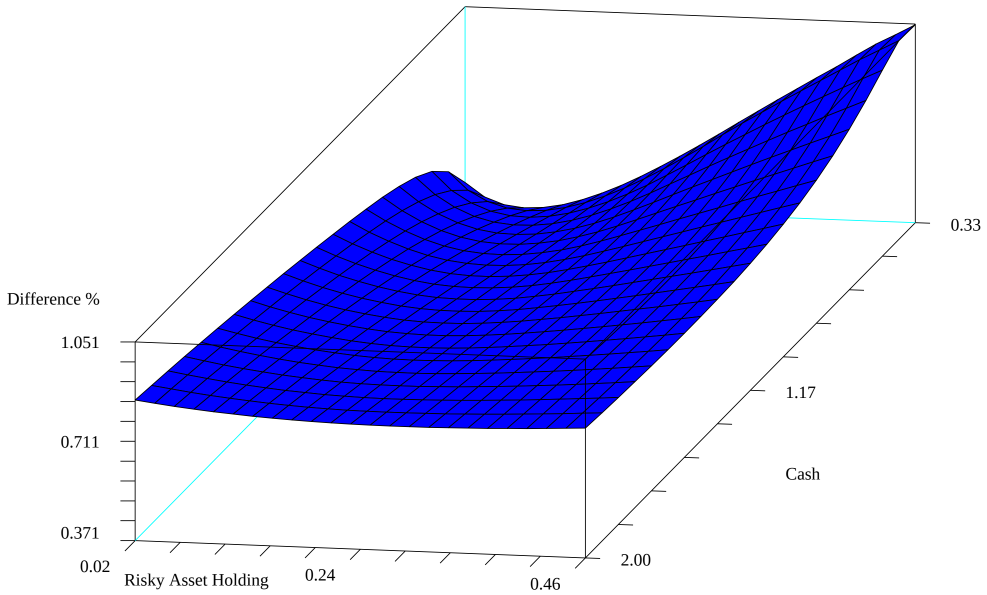


Figure 1: (Merton Soln-V)/Merton Soln n=21, m=24

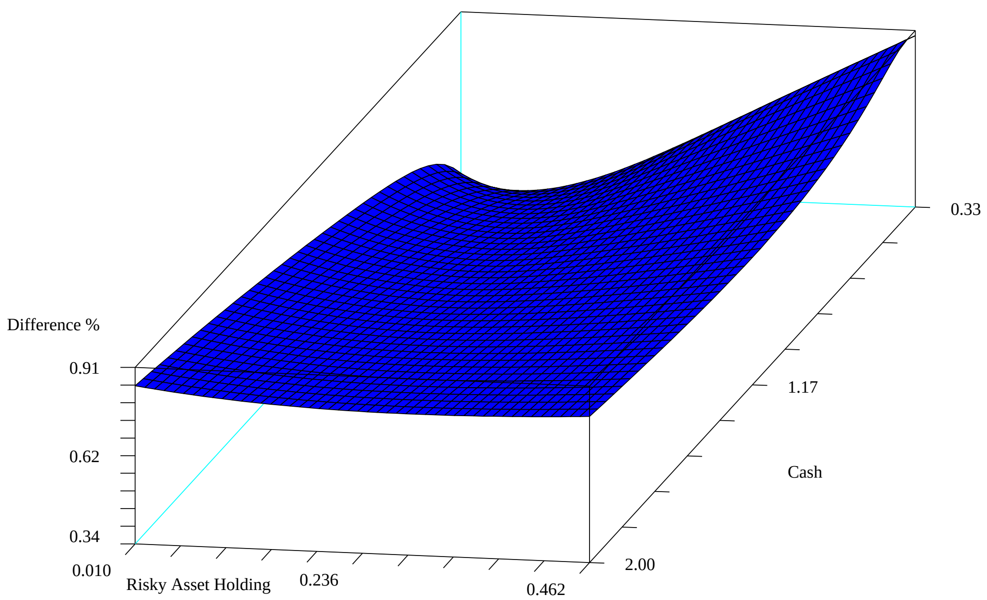


Figure 2: (Merton Soln-V)/Merton Soln n=41, m=48

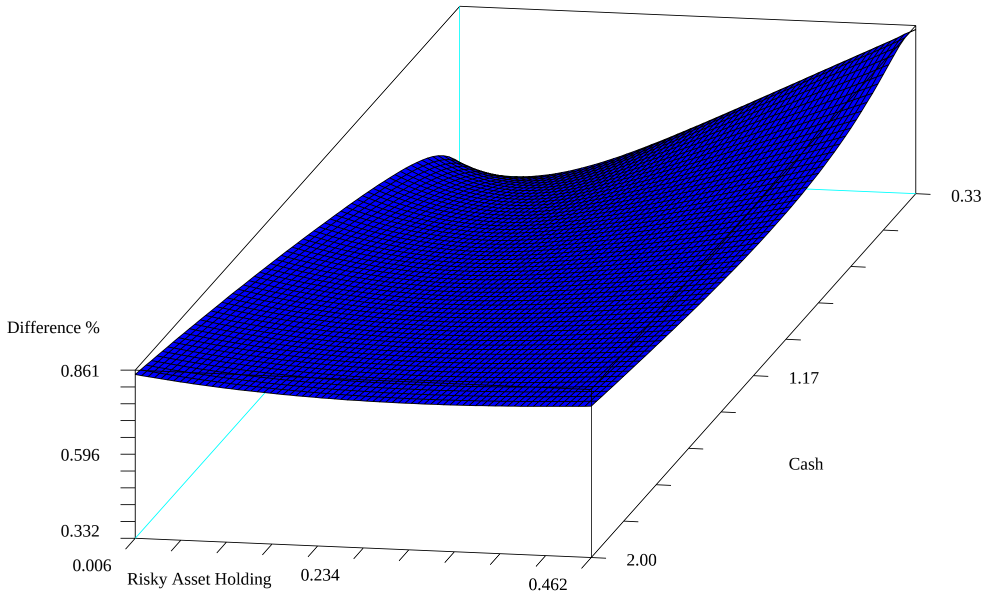


Figure 3: (Merton Soln-V)/Merton Soln n=61, m=72

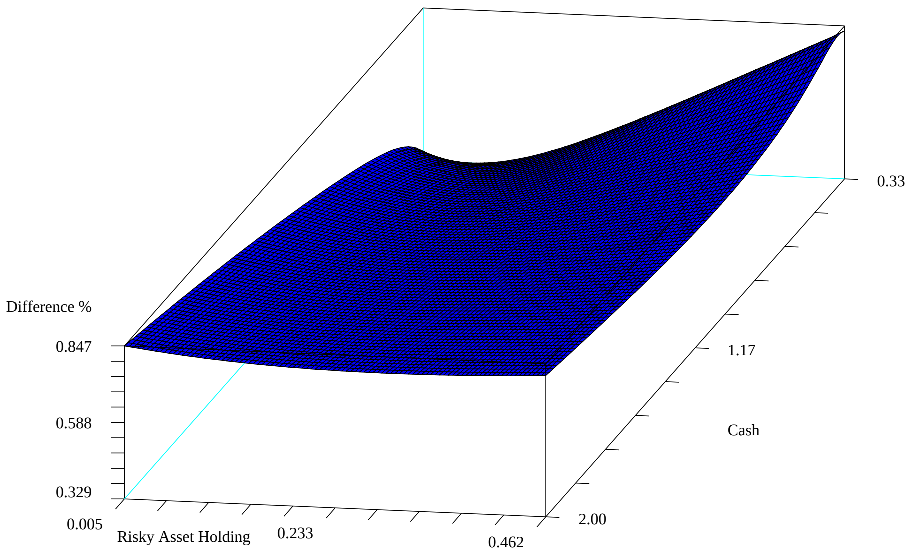


Figure 4: (Merton Soln-V)/Merton Soln n=81, m=96

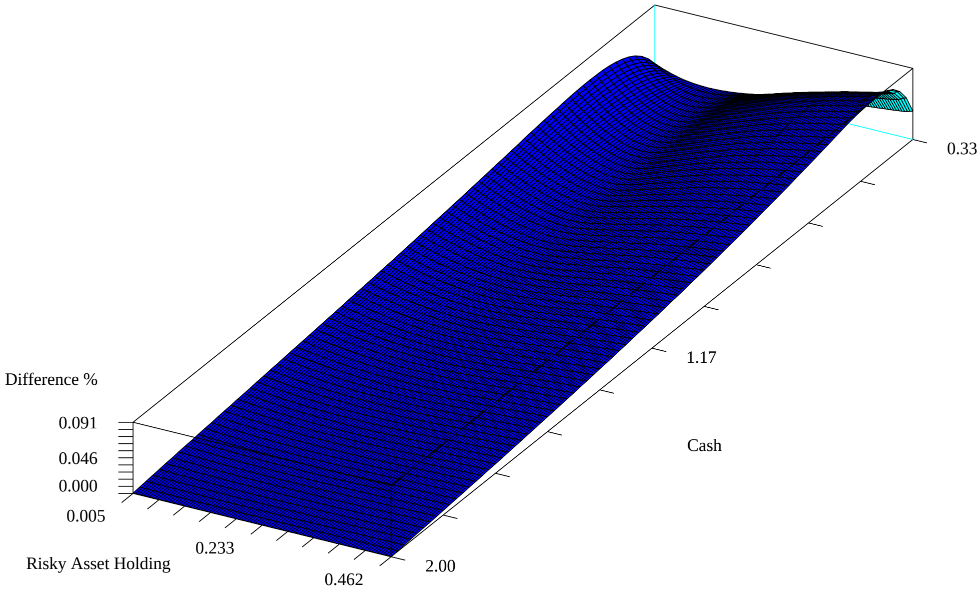


Figure 5: (Power Series-V)/Power Series Soln n=81, m=96