WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance

WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 240

Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk

by

Hayne Leland

November 1994

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Executive Publication Series is produced by the Research Program in Finance in the Walter A. Haas School of Business at the University of California. The Research Program in Finance has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The Executive reprint and working paper series were established in 1989 to supplement the Research Program in Finance's series of international Executive Seminars.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk

Hayne Leland

Haas School of Business
University of California, Berkeley

November 1994

Finance Working Paper #240

The author thanks William Keirstead, Stephen Schaefer, Richard Stanton, Matt Spiegel, and Klaus Toft for helpful comments.

Abstract

This paper examines the value of debt subject to default risk in a continuous time framework. By considering debt with regular principal repayments (e.g. through a sinking fund), we are able to examine bonds with arbitrary maturity while retaining a time-homogeneous environment. This extends Leland's [1994] earlier closed-form results to a much richer class of possible debt structures.

We examine the term structure of yield spreads and find that a rise in interest rates will reduce yield spreads of current debt issues. It may tilt the term structure as well. Duration is also affected by default risk. The traditional Macaulay duration measure overstates effective duration, which for "junk" bonds may even be negative. While short term debt does not exploit tax benefits as completely as does long term debt, it is more likely to provide incentive compatibility between debt holders and equity holders. The agency costs of "asset substitution" are minimized when firms use shorter term debt.

Optimal capital structure depends upon debt maturity. Optimal leverage ratios are smaller, and maximal firm values are less, when short term debt is used. The yield spread at the optimal leverage ratio increases with debt maturity.

CCD
12.10.95

BOND PRICES, YIELD SPREADS, AND OPTIMAL CAPITAL STRUCTURE WITH DEFAULT RISK

I. INTRODUCTION

A formula for the price of a coupon-paying bond with default risk and arbitrary maturity has proved elusive. The search for such an expression has been frustrated by the complexity of the debt instrument, whose cash flows include both coupon payments and a return of principal. When coupon payments are nonzero, Merton [1974] showed that bond values must satisfy a partial differential equation which has no known closed form solution for the general case.

A few special cases have admitted closed form solutions. Merton [1974] derived values for risky zero-coupon bonds. But coupon-paying bonds are far more common than zero-coupon bonds. Black and Cox [1979] examined coupon-paying debt, but only debt with infinite maturity. Neither Merton nor Black and Cox considered bankruptcy costs or the tax deductibility of interest payments. These are necessary ingredients for studying optimal capital structure as well as bond prices. Brennan and Schwartz [1978] used numerical methods to study debt value and optimal capital structure in a quite flexible setting. While instructive in providing examples, numerical techniques typically preclude the derivation of general results.

Leland [1994] derived closed form solutions for risky debt value and optimal capital structure in two environments consistent with time-homogeneous debt cash flows.[1] The first environment extends the infinite-maturity debt case examined by Black and Cox [1979] to include bankruptcy costs and tax deductibility of interest payments. The asset level which triggers bankruptcy is endogenously determined, and optimal capital structure can readily be calculated.[2]

Leland's second case also presumes a constant and perpetual coupon (unless default occurs). Unlike the first case, bankruptcy is triggered when the firm's net worth becomes negative. While this environment could be interpreted as infinite-life debt with a positive net worth covenant, Leland [1994] also likened it to rolling over very short term debt, or more exactly to a revolving line of credit, which is continuously renewed (at the same coupon rate) as long as assets are sufficient to repay debt principal.

While offering potential insights at both extremes of the maturity spectrum, Leland's analysis does not directly lend itself to analyzing debt with arbitrary maturities. Furthermore, the association of a positive net worth requirement with very short term debt begs the question of whether bankruptcy would be triggered *endogenously* if and only if net worth becomes

---

[1] Time homogeneity leads to ordinary differential equations for coupon-paying bond prices, which typically have closed form solutions. Using an alternative approach, Gennotte and Marsh [1992] develop an equilibrium model in which the term structure and risk premiums are endogenously determined. They estimate yield spreads on a generic risky (pure discount) bond. They do not examine optimal capital structure.

[2] A recent manuscript by Ross [1994] also examines capital structure in the context of a time-homogeneous model. The principal differences with Leland [1994] are Ross' assumptions that cash flow ("EBIT") follows a random walk whose drift may be arbitrary, and that bankruptcy is triggered whenever cash flow is less than the required bond coupon. Ross focuses on the "cost of capital" and on aspects of optimal recapitalization.

negative.

This paper derives a formula for the prices of a broad class of coupon-paying bonds with default risk and arbitrary maturity. We again focus on debt with time-homogeneous cash flows. But how is this possible, when finite-maturity debt cash flows seem time inhomogeneous by their very nature? The answer lies in examining debt with a time-homogeneous repayment of principal as well as a constant coupon. The real-world equivalent is a sinking fund provision. Sinking funds are quite common in corporate debt issues.[3] They require that a fraction of the principal value of debt be retired (or "amortized") on a regular basis.

In our model, a constant fraction of currently outstanding debt is retired annually, but replaced (except when bankruptcy occurs) by newly-issued debt.[4] The cash flow requirements for debt service in each period are therefore constant: a fixed coupon amount, and a fixed sinking fund requirement to retire current debt principal. The rate at which debt principal is retired serves as an inverse proxy for the average maturity of debt: The higher the debt retirement rate (or "rollover rate") m, the shorter is the average maturity $M = 1/m$ of debt. If $m = 0$, principal is never retired, and debt has infinite maturity as in

---

[3] See Smith and Warner [1979], for example.

[4] While some sinking funds allow the retirement of debt at its current market value, we assume (as is often the case) that debt is retired at its principal value.

Leland [1994]. As $m \to \infty$, the average maturity of debt approaches zero.[5]

Our results relate bond values and yield spreads to debt maturity, firm risk, leverage, tax rates, bankruptcy costs, and riskless interest rates.[6] The term structure of yield spreads may be increasing, humped, or (effectively) decreasing, depending on the degree of leverage of firms. These patterns confirm those derived for zero-coupon debt by Merton [1974] and Pitts and Selby [1983], and observed empirically by Sarig and Warga [1989].

The duration of debt with default risk exhibits some surprising properties. As risk increases, the sensitivity of risky debt to uniform shifts in interest rates (which we call *effective duration*) becomes significantly less than the traditional Macaulay [1938] measure of duration. For very risky debt, effective duration may even be negative. These findings are significant for fixed income portfolio managers attempting to immunize obligations by hedging against shifts in interest rate levels.

Duration also changes with the level of interest rates. For riskfree debt, a rise in interest rates shortens duration. This is reflected by "convexity": debt value is a convex function of

---

[5] For analytical convenience, we assume that a constant fraction of the *remaining* principal balance of each debt "vintage" is retired per unit time. This means (formally) that debt of each vintage has infinite life. However, by choosing m sufficiently large, the fraction of the principal outstanding at any future date (and the debt's duration) can be made arbitrarily small.

[6] Longstaff and Schwartz [1992] examine finite-life debt whose cash flows may be time inhomogeneous. But they assume the bankruptcy-triggering asset level (our V_B) is exogenously given and has constant present value. The endogeneity of bankruptcy is key to our results, and leads to many of the surprising comparative static results derived in Section III.

the interest rate. But risky debt behaves differently: it always exhibits less convexity than riskless debt, and its value may even be a *concave* function of the riskfree rate. Again, this has significant implications for hedging risky bond portfolios.

Optimal capital structure is shown to depend critically upon the maturity of debt. Optimal leverage will be considerably lower when shorter term debt is used. The term structure of yield spreads at optimal leverage levels is an upward sloping function of debt maturity. Firms should issue only high quality short term debt; longer term debt typically should not be as highly rated. Maximal firm value also increases with debt maturity. This last result poses an important question: why do firms ever choose to issue short-term debt?

An answer may be found in the greater potential agency costs of long-term debt. Agency theory suggests that stockholders will wish to increase risk--by "asset substitution"--in order to transfer value from debt to equity (see, for example, Jensen and Meckling [1976] and Harris and Raviv [1990]). This conclusion relies upon the analogy between equity of a levered firm and a call option suggested by Black and Scholes [1973]. However, the analogy is inexact except in Merton's [1974] case of zero-coupon bonds. With coupon-paying bonds, bankruptcy may be triggered by asset value falling to a critical level at any time, rather than only at maturity. This bankruptcy-triggering value is endogenous, and will change with the riskiness of the asset value process. And with tax benefits and bankruptcy costs, the value "pie" has participants beyond the stock and bondholders alone.

We show that the agency problems vanish with short and intermediate term debt, when bankruptcy is not imminent. Incentives are misaligned for risky long term debt. And they become misaligned for all debt maturities as bankruptcy is neared, or when bankruptcy costs and taxes are negligible.

Our generalizations confirm many but not all of Leland's [1994] conclusions. As m increases (and debt duration becomes shorter), the endogenously determined bankruptcy-triggering asset value V_B increases, but approaches a higher value than Leland [1994] predicted. And the fact that V_B is endogenously determined for short term debt also alters its behavior.

The analytical techniques of this paper also differ from earlier work. Merton [1974], Black and Cox [1979], and Leland [1994] solve (partial) differential equations to determine bond values. A martingale approach is used here to derive risk-neutral expected values.[7]

II. RISKY DEBT WITH ARBITRARY AVERAGE MATURITY

As in Merton [1974], Black and Cox [1976], and Brennan and Schwartz [1978], the firm has productive assets whose market value V follows a continuous diffusion process with constant proportional volatility:

[7] Such techniques, of course, have been widely used to study related problems since the work of Cox and Ross [1976] and Harrison and Kreps [1979]. Longstaff and Schwartz [1992] also use a martingale approach.

(1) $$\frac{dV}{V} = \mu(V,t)dt + \sigma dz$$

where dz is a standard Brownian motion. The process continues without time limit unless V falls to a bankruptcy-triggering value V_B, which is endogenously determined and depends upon the amount of debt issued.

We examine stationary debt policies. At each moment in time, the firm has debt with constant total principal P, paying a constant total coupon rate C. The firm continuously rolls over a fraction m of debt. That is, it continuously retires outstanding debt principal at the rate mP, and replaces it with new debt of equal coupon, principal, and seniority.[8]

At each instant τ, the firm issues new debt principal equal to a constant p, paying a coupon rate c.[9] Since mP is the amount of debt principal retired at each instant, it follows that

(2) $$p = mP$$

Let $p(\tau,t)$ denote the principal outstanding at time t of debt issued at time $\tau \leq t$. Note $p(\tau,\tau) = p$ for all τ. As t passes, the principal of the debt issued at any time $\tau \leq t$ is retired at a fractional rate m:

---

[8] Although the new debt is assumed to have the same coupon and principal (to preserve stationarity), the price at which the new debt can be sold depends upon current asset value V. Additional equity will have to be raised if the new debt selling price is lower than its principal value.

[9] Formally, p is a *rate* of principal issue; the actual amount of debt issued over the instant dt is pdt. However, in equations (2), (4), (5), (9), and (10) we shall drop the dt terms and speak of the "rate" as an "amount." The strict reader may wish to supply multiplicands of dt to both sides of these equations.

(3) $$\frac{\partial p(\tau,t)/\partial t}{p(\tau,t)} = -m$$

implying both the outstanding principal and coupon of debt of each vintage τ declines exponentially with time:

(4) $$p(\tau,t) = e^{-m(t-\tau)}p$$
$$c(\tau,t) = e^{-m(t-\tau)}c$$

Confirming equation (2), when $m > 0$, at any time t the total principal outstanding P and coupon rate C are:

(5) $$P = \int_{-\infty}^{t} p(\tau,t)d\tau = \int_{-\infty}^{t} e^{-m(t-\tau)}pd\tau = \frac{p}{m}$$

(6) $$C = \int_{-\infty}^{t} c(\tau,t)d\tau = \int_{-\infty}^{t} e^{-m(t-\tau)}cd\tau = \frac{c}{m}$$

Note that remaining units of debt from all prior issues have the same value per unit, since units of all vintages pay the same coupon, and the remaining units of all vintages will be retired at the same fractional rate m. Thus a unit of bonds issued five years ago will look exactly like (and will carry the same price) as a unit of bonds issued today, except that there will be fewer units of the older vintage bonds. All units have the same seniority.[10]

The inverse of the rollover rate m also serves as a parameter of the average maturity M of (riskfree) debt, and its duration Z. Let the current time be $\tau = 0$. The fraction of currently outstanding debt principal which is redeemed at time t in the future is me^{-mt}. The average maturity M of debt is therefore

(7) $$M = \int_0^{\infty} t(me^{-mt})dt = \frac{1}{m}$$

Let Z represent the Macaulay measure of duration. That is,

(8) $$Z = \frac{\int_{t=0}^{\infty} te^{-Rt}[e^{-mt}(c+mp)]dt}{\int_{t=0}^{\infty} e^{-Rt}[e^{-mt}(c+mp)]dt} = \frac{1}{m+R}$$

where R is the yield to maturity of the debt (equal to r when debt is riskfree), and $e^{-mt}(c + mp)$ is the cash flow (coupon plus amortized principal) accruing to currently-issued debt, at time t in the future. If $m = 0$, average maturity M is infinite and $Z = 1/R$, the duration of a consol with yield R. As $m \to \infty$, both average maturity M and duration Z approach zero.

---

[10] Because all outstanding debt units are homogeneous, we can treat the initial (at $t = 0$) total principal of debt P (and coupon C) as control variables, rather than simply the current flow p (and c). By assumption, however, once P and C are fixed they are expected to remain constant thereafter. The reader may note a similarity with the example of light bulbs whose longevity is exponentially distributed. At any moment looking forward, all light bulbs currently operating will have the same distribution of remaining life, regardless of how long they have already operated.

We now derive the value of debt as a function of m and V_B, where V_B, the asset value which will trigger bankruptcy if reached, is less than the current asset value V. Let $d(0)$ represent the value of the debt which is currently issued. Cash flows to these debtholders will include future coupon payments and fractional repayments of principal, in amounts $e^{-mt}(c + mp)$, unless bankruptcy occurs. Let α be the fraction of asset value lost in the event of bankruptcy. Then $(1-\alpha)V_B$ is the amount in total that bondholders receive if bankruptcy occurs. We presume absolute priority, in that bondholders receive all remaining asset value, and stockholders receive nothing, when the firm becomes bankrupt.

Using risk neutral valuation, and denoting the density of the first passage time t to V_B from V as $f(t;V,V_B)$, gives a value to currently issued debt

$$(9) \qquad d(0)=\int_0^\infty e^{-rt}e^{-mt}(c+mp)[1-F(t;V,V_B)]dt+\int_0^\infty e^{-rt}(e^{-mt}p/P)(1-\alpha)V_B f(t;V,V_B)dt$$

The first term in equation (9) represents the discounted expected value of the continuously declining coupon plus principal repayment (which will be paid with probability $(1-F)$, where F is the cumulative distribution function of the first passage time); the second term represents the expected present value of the fraction of the bankruptcy value of the firm which will go to owners of bonds issued at time zero, if bankruptcy occurs at time t. Recalling that $p/P = m$, integrating by parts, and simplifying gives

(10) $$d(0)=\frac{C+mP}{r+m}\left[1-\int_{t=0}^{\infty} e^{-(r+m)t}f(t;V,V_B)dt\right]+m(1-\alpha)V_B\left[\int_{t=0}^{\infty} e^{-(r+m)t}f(t;V,V_B)dt\right]$$

In Appendix A, it is shown that

(11) $$\int_{t=0}^{\infty} e^{-(r+m)t}f(t;V,V_B)dt=\left(\frac{V}{V_B}\right)^{-y}$$

where

(12) $$y=\frac{(r-\delta-.5\sigma^2)+[(r-\delta-.5\sigma^2)^2+2(m+r)\sigma^2]^{\frac{1}{2}}}{\sigma^2}$$

and δ is the (constant) proportional payout rate of asset value V by the firm.[11] Coupons are paid at rate C to bondholders; in addition, $p - d$ is the net cash outflow associated with redeeming a fraction m of the principal P, less the market value d of floating new debt of equal coupon and principal but whose value fluctuates with V. Thus $(\delta V - C - p + d)$ is the payout rate to stockholders. Note this payout rate declines as V falls and may become negative (i.e. new equity must be issued to meet bond requirements).[12]

---

[11] Following Brennan and Schwartz [1978] and Leland [1994], we assume that the underlying asset with value V is a traded asset, or is perfectly correlated with a traded asset. This implies that the drift $\mu(A,t)$ of the asset process in equation (1) equals the riskfree rate less the payout rate, $r - \delta$.

However, it might be argued that an unlevered firm inefficiently exploits tax benefits, and will not be traded. (The author thanks Fischer Black for raising this point). In this case, rather than returning a "fair" risk adjusted return of r-δ, the untraded asset may offer a lower risk adjusted return, e.g. r-δ-λ. Our payout rate δ can then be reinterpreted as the *sum* of an actual payout rate, plus an underperformance rate.

[12] Bankruptcy will occur at an endogenously-determined asset value, V_B, when equity value is no longer sufficient to cover the required bond coupons plus refundings. (See Section II(iii) below). One (but not the only) environment consistent with this description is that assets generate cash flows δV which are always paid out collectively to stock and bond holders. Note this does *not* mean that bankruptcy will occur when debt service payments exceed cash flow δV. Rather, bankruptcy is triggered when asset value has fallen to where equity can no longer be issued to meet debt service requirements. Stockholders will always want to avoid bankruptcy if they can by issuing stock to meet the current debt service--something (their diluted stock value) is worth more than nothing (their stock value if bankruptcy occurs).

The value of debt outstanding of generation τ, $\tau \leq 0$, is $e^{m\tau}d(0)$, since all outstanding units of debt sell for the same price, but there are fewer units outstanding of older debt vintages due to accumulated debt retirement. All units sell for the same price because all carry the same coupon, and the retirement of *remaining* units follows the same proportionally declining schedule. Integrating over $-\infty \leq \tau \leq 0$ gives $D = d(0)/m$, the total value of debt outstanding. Dividing equation (10) by m, and recalling that $P = p/m$ and $C = c/m$, gives

(13) $$D=\frac{C+mP}{r+m}[1-(\frac{V}{V_B})^{-y}]+(1-\alpha)V_B(\frac{V}{V_B})^{-y}$$

or, equivalently (since average maturity $M = 1/m$),

(13') $$D = \frac{MC+P}{Mr+1}[1-(\frac{V}{V_B})^{-y}]+(1-\alpha)V_B(\frac{V}{V_B})^{-y}$$

Note that as $m \to 0$, $y \to x$ (where x is given by equation (12) with $m = 0$), and $D \to C/r + [(1-\alpha)V_B - (C/r)](V/V_B)^{-x}$, the same as in Leland [1994; equation (7)]. As $m \to \infty$, we have $y \to \infty$, and $D \to P$.

Equation (13) is a closed-form solution for the value of debt with arbitrary rollover rate m, and therefore for debt with arbitrary average maturity $M = 1/m$. However, the bankruptcy-triggering value V_B remains to be determined. To find this value, we must invoke the smooth-pasting conditions for equity. But first we must determine the value of the firm and the value of equity.

(i) <u>The value of the firm v</u>

The total value of the firm, v, equals its asset value V, plus the value of tax benefits, less the value of bankruptcy costs: $v = V + TB - BC$, where (as in Leland [1994])

(14) $$TB = (\tau C/r)[1-(\frac{V}{V_B})^{-x}]$$

(15) $$BC = \alpha V_B(\frac{V}{V_B})^{-x}$$

implying

(16) $$v = V+(\frac{\tau C}{r})[1-(\frac{V}{V_B})^{-x}]-\alpha V_B(\frac{V}{V_B})^{-x}$$

where τ is the corporate tax rate.[13] This presumes that tax benefits are received whenever the firm is solvent, an assumption we modify later. It also assumes that tax benefits depend only upon the coupon, and not whether the debt is originally sold at a discount or premium

[13] As discussed in Miller [1977], in the presence of personal tax rates

$$\tau = 1 - (1-\tau_c)(1-\tau_s)/(1-\tau_d),$$

where τ is the effective tax advantage of debt, τ_c is the corporate tax rate, τ_s is the personal tax rate on equity income, and τ_d is the tax rate on debt income.

to principal value.[14] Observe that x (given by equation (12) with $m = 0$), not y, is the exponent in these equations. This is because total coupon C and principal P remain constant: total tax benefits and potential bankruptcy costs are *not* being reduced over time at rate m, in contrast with the outstanding amount of debt principal of each generation of debt. It can be shown that $(V/V_B)^{-x}$ is simply the present value of \$1 received at the first passage time of asset value to V_B, when commencing at V.

(ii) The value of equity E

The value of equity equals the firm less debt: $E = v - D$, where v is given by equation (16) and D by equation (13):

$$(17) \qquad E = V + \left(\frac{\tau C}{r}\right)\left[1 - \left(\frac{V}{V_B}\right)^{-x}\right] - \alpha V_B \left(\frac{V}{V_B}\right)^{-x} - \left(\frac{C + mP}{r + m}\right)\left[1 - \left(\frac{V}{V_B}\right)^{-y}\right] - (1 - \alpha) V_B \left(\frac{V}{V_B}\right)^{-y}$$

The dynamics of E, which follow directly from the dynamics of v and D, will be important in determining the value of options on the stock of leveraged firms. However, we shall not pursue these concerns in this paper.[15]

---

[14] Current U.S. tax rules require that any difference between initial selling price and principal value of bonds be considered interest. Much of our analysis examines the value of bonds which initially sell at par. However, subsequent debt issuances may occur at selling prices other than par, since by assumption c and p remain constant in each debt vintage, but V (and therefore debt value) may change. Our analysis ignores the potential changes in the value of tax deductions resulting from debt being sold at value other than par.

[15] Toft [1993] has provided a closed form solution for option prices when $m = 0$; his analysis can (in principle) be extended to price options on firms with debt of arbitrary duration using the dynamics of E.

(iii) The bankruptcy-triggering asset value V_B

Now consider the asset value V_B which, if reached, will trigger bankruptcy. Bankruptcy occurs when the asset value of the firm drops to a level such that the firm can no longer raise sufficient capital to retire the required amount of debt, plus pay the current total coupon.[16] Since over an interval dt the required amount of debt service is infinitesimal, it follows that the value of equity when asset value falls to V_B is zero: $E(V_B) = 0$. But in addition, to maximize equity value E, the smooth-pasting condition must be satisfied by E at $V = V_B$:

(18)
$$\frac{\partial E(V)}{\partial V}\Big|_{V=V_B} = 1+\left(\frac{\tau C x}{r V_B}\right)+\alpha x-\left(\frac{y}{V_B}\right)\left(\frac{C+mP}{r+m}\right)+(1-\alpha)y = 0$$

Solving for V_B gives

(19)
$$V_B = \frac{\left[\frac{(C+mP)y}{r+m}-\frac{\tau C x}{r}\right]}{1+\alpha x+(1-\alpha)y}$$

Using (19) to substitute for V_B in (13), (16), and (17) gives ***closed form solutions for the value of debt, the value of the firm, and the value of equity.***

---

[16] As in Brennan and Schwartz [1978] and Leland [1994], we assume that assets in place cannot be liquidated in order to raise money to pay debtholders. Equity must be raised to meet demands (dividends, interest, and net principal repayment) in excess of the cash flows paid out to investors, δV. See also footnote 12.

For all examples considered, E is strictly increasing and convex in V for $V > V_B$. In principle we can solve for V as a function of E from (17). Using this to replace V in (13) would create a debt value function in terms of equity value E rather than the (possibly difficult to observe) asset value V. We shall not pursue this approach, however.

As $m \to 0$, $y \to x$ and $V_B \to (1-\tau)(C/r)(x/(1+x))$, the same as in Leland's [1994] endogenous bankruptcy case. However, as $m \to \infty$, $y \to \infty$ and $V_B \to P/(1-\alpha)$. This is unlike Leland [1994], where it was argued that (very) short run debt could be associated with $V_B = P$, not $P/(1-\alpha)$. If $\alpha > 0$, bankruptcy will occur at a higher asset value than P when debt is very short term.

But if the firm has assets $V = V_B$ which exceed the bondholders' principal P, why must bankruptcy be declared? Recall that bankruptcy is triggered *not* because V falls beneath P, but rather because the firm cannot raise sufficient equity to pay the current coupon plus the net cost of retiring bond principal (the cost of retiring principal at par, less the revenue from selling--perhaps at less than par--the newly-issued bonds). In this event, bankruptcy occurs and productive assets must be liquidated (or reorganized). But assets are in place; liquidation or reorganization costs a fraction α of their value. As $m \to \infty$, debt becomes riskfree, since post-liquidation value, $(1-\alpha)V_B$, approaches the principal value P of debt. Thus an important implication of the model is that, as long as $P < V(1-\alpha)$, debt can be made essentially riskfree by making its average maturity sufficiently short.[17]

---

[17] We shall see in Section III that this limiting result is exactly that--a limiting result. In many cases we examine, debt with average maturity as short as 3 months still may carry a substantial yield premium.

If liquidation costs are zero, then for very short term debt the bankruptcy-triggering asset value V_B will equal the principal value P of the bonds, as in Leland [1994]. But if $\alpha > 0$, the bankruptcy triggering asset value (at which the firm cannot raise sufficient funds to pay debtholders their current coupon plus net return of bond principal) will exceed P.

Appendix B considers the optimal V_B in the case where tax deductibility is lost when V falls beneath $V_T \geq V_B$. It is shown there that

(20)
$$V_B = \frac{y(C+mP)rV_T}{(r+m)[rV_T[1+\alpha x+y(1-\alpha)]+\tau Cx]}$$

When $V_T > V_B$, V_B in (20) will exceed V_B in (19). Thus the loss of tax deductibility raises the asset value at which bankruptcy will be declared. This in turn will lower the value of debt and equity, as can be seen from equations (13) and (17).

III. APPLICATIONS

(i) Debt Value and Debt Capacity

From equation (19) or (20), V_B can be substituted into equation (13) to yield a closed form solution for total debt value, given coupon C and principal value P. When debt is first issued, however, there is typically a further constraint on relating market value, coupon, and principal: the coupon is set so market value D equals principal value P. If V_0 is the asset value when the debt is first issued, this constraint requires that C be the smallest solution to the equation

$$D(V_0;C,P) = P \tag{21}$$

Using (13) and (19) or (20), it is simple to find solutions to (21) numerically.[18]

Figure 1 plots the value of newly-issued debt D as a function of *leverage (D/v)* for different maturities $M = 1/m$, given that principal P and debt value D coincide at current value $V_0 = 100$. Our base-case example assumes $r = .075$, $\sigma = .20$, $\alpha = .50$, and $\tau = .35$.[19] We further assume that $V_T = 50 + 2.5C$. This implies that, at the initial asset value, the coupon rate can be as high as 20% of asset value before the tax advantage of deducting coupons is lost. All ability to tax shelter coupon payments is lost if V falls to 50 or less; implicitly, the firm is generating no profits to shield interest payments at that low asset value.[20]

---

[18] Our analysis in Section II considers the case where at each moment, including the present, only a small amount of debt ($mPdt$) is issued. However, since all outstanding units from previous vintages are identical in value, we can equally well assume the total debt is issued at the current moment, and thereafter rolled over at rate m. However, while C may be chosen so that principal equals debt value when debt is originally issued, it will not generally equal debt value thereafter. This is because we require C and P to remain at the same level, once they are originally set. Subsequent issues of debt will sell above par (principal) value if $V > V_0$, and below par when $V < V_0$.

[19] These parameters were chosen to reflect the current U.S. environment, with the possible exception of the bankruptcy parameter α. We require substantial bankruptcy costs for short term debt to have "reasonable" yield spreads. Actual bankruptcy costs are difficult to ascertain, although many studies indicate they may be smaller than 50% (e.g. Altman [1989]). Perhaps yield spreads will ultimately prove to be the data we use to impute bankruptcy costs, rather than vice-versa.

[20] Without a V_T which increases in the coupon paid, debt capacity using shorter term debt may be unboundedly large: D is monotonically increasing in C when bonds are sold at par. But clearly such a result is absurd: at some level, coupons will exceed profits and tax deductibility will be limited.

Finally, we assume that the firm pays out an amount δV which covers the initial coupon at $V = V_0 = 100$, plus a 3 percent dividend on the initial value of equity E_0.[21] That is, the payout by the firm satisfies $\delta V_0 = C + .03E_0$, or $\delta = .01C + .0003E_0$

In Figure 1, debt capacity (D_{max}) is the maximal value of the debt value curve. Note that the debt capacity is smaller for shorter maturities. It can be shown that debt capacity falls as volatility σ and/or bankruptcy costs α rise. Maximal debt value tends to occur at higher coupon levels (denoted C_{max}) for shorter term debt, but at approximately the same leverage (about 75%-80%) for debt of different maturities.

For any given maturity, as σ increases, debt value falls when $C/V < C_{max}/V_0$, but increases with σ when C is very large relative to V and the bond is "junk".[22] This surprising result, which we revisit in Section V (and Figure 8), results from the endogeneity of bankruptcy. It is more pronounced with longer term debt.

A rise in the riskfree rate can also increase debt value when the firm is near bankruptcy. Leland [1994] derived similar results for "junk" bonds, but only for the case of long term

---

[21] This is in contrast with the assumption of Brennan and Schwartz [1978], who presume that coupons and the cost of debt retirement are entirely paid by newly-issued equity. Shareholders receive δV - C, less the net debt retirement costs (the market value of newly issued debt less the principal value of retired debt). As V falls, this net dividend received by shareholders falls, and eventually becomes negative (additional equity must be issued). At $V = V_B$, the contributions required are no longer met by shareholders with limited liability, and bankruptcy occurs.

[22] If debt is optimally issued originally, the coupon rate will exceed C_{max} only if the bond is a "fallen angel"--V has fallen beneath V_0. Firms would never initially offer debt carrying so large a coupon.

debt; short term debt (which in his model had an exogenously specified V_B) exhibited no such anomalies. In this model, anomalies can occur even with short term debt.

(ii) The Term Structure of Yield Spreads

Figure 2 examines yield spreads ($C/D - r$) of newly-issued debt as a function of maturity, for alternative leverage ratios. (Figure 3 presents the same data in a 3-dimensional format, using a log scale.) For high leverage levels, yield spreads are high, but decrease as maturity increases beyond 0.50 years. For moderate leverage levels, we find that yield spreads are distinctly "humped": intermediate term debt offers higher yields than either very short or very long term debt. In Figure 2, at 50% leverage, the yield spread for short term debt (maturity M = 3 months) is 50 basis points, rising to 122 basis points for debt with maturity 5 years, and falling to 110 basis points for 20-year debt. Finally, for firms that have low leverage, yield spreads are low but increase with debt duration. Interestingly, these patterns are also predicted by Merton's [1974] model of zero coupon debt (without taxes or bankruptcy costs), which in turn have been verified empirically by Sarig and Warga [1989].[23]

We turn now to the behavior of yield spreads as bankruptcy costs, asset risk, and riskfree

---

[23] With high leverage, the term structure of yield spreads decreases from a peak at very short maturity, but it still is of a "humped" shape. This is because as $M \to 0$ debt becomes riskless, since in the cases we examine $V > V_B \to P/(1-\alpha)$, and the firm is solvent even in the limit. In Merton [1974], the yield spread approaches infinity as maturity approaches zero in the case where $P/V > 1$. But this case implies the firm is insolvent in the limit as $M \to 0$, since $V_B \to P > V$.

rates change. There are two sets of comparative statics to consider. First, we ask how the yield spreads of current debt (with coupon and principal equal to those providing 50% leverage in the base case) change.[24] Then, we examine how yield spreads of newly-issued debt (with coupon and principal providing leverage of 50%, and bonds selling at par in the changed environment) vary.[25]

Yield spreads of both current and newly-issued debt are quite sensitive to the volatility σ of underlying assets. The yield spreads of current debt with 3-month, 5-year, and 20-year maturity rise to 133, 208, and 177 basis points, respectively, if volatility rises to 25%, and fall to 7, 49, and 50 basis points, respectively, if volatility falls to 15%. This compares with yield spreads of 50, 122, and 110 basis points in the base case with 20% volatility.[26] For newly-issued debt, yield spreads rise to 98, 217, and 198 basis points, respectively, if volatility rises to 25%, and fall to 25, 52, 45 basis points, respectively, when volatility falls to 15%.

The effect of changes in the level of riskfree interest rates on yield spreads is more surprising. For current debt, a rise in riskfree rates *reduces* yield spreads. If the riskfree rate increases from 7.5% to 10%, yield spreads on 3-month, 5-, and 20-year maturities fall

---

[24] This must be an unexpected change, since we have not allowed for the possibility of randomly changing parameters. Since environmental changes mean that the bonds no longer sell at par, we have computed a yield spread to maturity M.

[25] Since in this case the change takes place before debt is issued, it is possible that the parameters will remain fixed thereafter, as our model presumes.

[26] Note that for the longer maturities, the yield spread at a 20% volatility is approximately the average of yield spreads for 15% and 25% volatilities. The yield spread for shorter term maturities seems to exhibit greater convexity in volatility, implying an average yields spread across volatilities which exceeds the yield spread at the average volatility.

to 18, 45, and 34 basis points, respectively, from 50, 122, and 110 basis points. A fall in the riskfree rate to 5% raises yield spreads to 103, 257, and 247 basis points.

When leverage is kept at 50% by issuing new debt at par, the yield curve rotates clockwise as interest rates rise. The yield spread on 3-month debt rises to 57 basis points and falls to 39 basis points as the riskfree rate rises to 10% or falls to 5%, respectively. But for 5-year (20-year) debt, yield spreads decrease to 98 (85) basis points when rates rise, but increase to 150 (143) basis points when rates fall. This rotation of the yield spread structure for newly-issued debt is observed at leverages between 40% and 60% as well. At 60% leverage, yield spreads on newly-issued debt swing from 173/262 bps (shortest/longest maturity debt) to 281/207 bps, as riskfree interest rates (with a flat term structure) rise from 5% to 10%. The net swing of 163 basis points in long/short yield spreads actually shifts the term structure for these corporate bonds from upward to downward sloping. Of course, these results assume asset value V remains fixed. It is likely that shifting interest rates may also change V, and the net effect on yield spreads must reflect this effect as well.

Yield spreads are quite also sensitive to bankruptcy costs, particularly for short term debt. For example, with bankruptcy costs falling to 25%, the yield spreads on current debt fall to 1, 63, and 78 basis points, for debt with maturities 3 months, 5 years, and 20 years, respectively; the figures for newly-issued debt are 4, 67, and 77 basis points, respectively, when leverage is maintained at 50%. This compares with 50, 122, and 110 basis points when bankruptcy costs are 50%.[27]

(iii) The Duration and Convexity of Risky Debt

The Macaulay [1938] measure of duration is an accurate description of the percent change of a bond price in response to a uniform change in the level of interest rates--for bonds with no default risk. A critical question follows: How sensitive are correct measures of duration to the presence of default risk? By "correct" measure of duration, we simply mean an expression which correctly predicts the percentage change in the (risky) bond value in response to a change in the riskfree rate.

In our framework, the Macaulay duration of risky bonds is $1/(m+R)$, where $R = C/D$ when bonds sell at par (i.e. $D = P$ at $V = 100$). Using the base example, we compute the change in the value of current debt for a 1% change in the riskfree interest rate. The true or effective duration of risky debt, plotted as a function of the Macaulay duration, is given in Figure 4, for different degrees of leverage. Figure 5 also plots effective duration vs. Macaulay duration, but for different levels of yield spreads rather than leverage. When leverage is 30% or less, effective duration is slightly smaller than Macaulay duration. As

[27] Although 77 basis points is the average spread for investment-grade corporate debt over Treasury bonds as reported in Kim, Ramaswamy, and Sundaresan [1993], two important aspects may limit the realism of such a comparison. First, corporate bonds tend to be less liquid than governments. A more realistic comparison might be the corporate yield spread relative to off-the-run Treasury bonds. Second, the typical corporate bond is callable, which tends to increase its spread relative to the noncallable government bonds.

leverage becomes larger, effective duration becomes much shorter than Macaulay duration: with leverage of 60%, debt of 20-year maturity has Macaulay duration of about 6.7 years but effective duration is only 2.1 years.

When leverage exceeds 75%, effective duration of short term debt becomes negative: as the riskfree rate rises, so do bond prices. At 80% leverage, duration for all maturities is negative. This reflects the importance of the endogenous bankruptcy value V_B. As previously observed, high risk ("junk") bonds may rise in value with the riskfree rate because the bankruptcy value V_B is lower, implying bankruptcy is less imminent.

Such dramatic differences in effective duration vs. Macaulay duration suggest that immunization and related techniques using corporate bonds must explicitly reflect actual bond risk, and not rely on traditional duration-matching methods.

Riskless debt value is a convex function of the interest rate r. Convexity is critical for managing a duration-matching strategy. A dynamic strategy must be followed, because duration increases as interest rates fall. However, if debt values were concave rather than convex in the riskfree interest rate, the *opposite* kind of dynamic hedging would be required.

We find that the riskiness of debt, as well as its maturity, affects convexity. Figure 6 examines debt value as a function of the riskfree interest rate r. Panels 6a - 6c examine long term (20-year maturity) debt carrying yield spreads of 0, 50, and 200 basis points when r =

7.5%. Panels 6d - 6f examine intermediate-term debt (5-year maturity), for coupon levels also consistent with yield spreads of 0, 50, and 200 basis points at r = 7.5%.

The differences in convexity as yield spreads (and therefore bond risks) increase are dramatic, particularly for long term debt. As debt becomes increasingly risky, convexity is reduced and ultimately turns to concavity. The degree of concavity is most pronounced at lower interest rates. Again, this suggests that the hedging of risky debt portfolios requires quite different actions than the hedging of riskfree debt.

(iv) Bankruptcy Rates and Bond Ratings

Figure 7 considers the cumulative probability of bankruptcy over a 20 year period, for debt with average maturities of 3 months, 5 years, and 20 years. Panels 7a - 7c reflect debt at each duration bearing a 100 bp yield spread over the riskfree rate. The probability of bankruptcy over a period T is given by

$$N\left(\frac{-b-\lambda T}{\sigma\sqrt{T}}\right)+e^{-2\lambda b/\sigma^2}N\left(\frac{-b+\lambda T}{\sigma\sqrt{T}}\right) \qquad (21)$$

where $b = Log(V/V_B)$, $\lambda = \mu - \delta - .5\sigma^2$, μ (the rate of return to the asset V, including payouts) is assumed to be 15% per year, and other parameters are as in the base case.

For long term debt, the cumulative probability of bankruptcy is negligible over the first two years, and eventually rises to 1.5%. Thus, approximately 1.5% of 20-year debt which pays

100 bps over the riskfree rate will default.[28] Very short term debt carrying the same yield spread implies a considerably higher probability of default over any fixed time period. This follows because the bankruptcy-triggering value V_B is larger for shorter maturities.[29] Reflecting the high probability of default, optimal use of short term debt dictates much lower coupon levels than those which generate a 100 bp yield spread.

Panels 7d - 7f chart bankruptcy probabilities at the optimal coupon--that is, the coupon that maximizes firm value--for debt with maturity 0.25, 5, and 20 years. Optimal leverage (considered in Section IV below) is 22.1 percent for debt with maturity 3 months, 39.9 percent for debt with maturity 5.0 years, and 49.2 percent for debt with maturity 20 years. Associated yield spreads are 0 bps, 45 bps, and 103 bps. In contrast with panels 7a - 7c, which assumed a 100 bp yield spread for all durations, panels 7d - 7f indicate that (optimal) debt with the shortest duration gives the smallest probabilities of bankruptcy, about 0.35% over a 20-year horizon). Clearly this implies that it is optimal to issue only the highest quality short term debt, as seems to characterize the commercial paper market.[30] Optimal long term debt, carrying a 103 bps annual yield spread, has a probability of default of about 1.5% over a 20-year horizon. Such debt might receive Moody's bond ratings in the range

---

[28] The long-term limiting probability of default is highly dependent on the drift μ assumed for the asset process V. For example, if $\mu = .125$ (rather than .15), the probability of default of 20-year debt would be close to 4% rather than 1.5%. The limiting probability of default for 0.25 year debt would rise to 30% from 20%.

[29] Short term yield spreads are low despite the higher cumulative probabilities of bankruptcy because the bulk of short term principal will be repaid before the cumulative probabilities of bankruptcy become sizable.

[30] However, commercial paper spreads are larger than predicted here, despite their high credit ratings. This may reflect the tax and liquidity advantages of short-term Tbills rather than a default premium. Interestingly, defaults in the commercial paper market have been negligible--consistent with our findings for optimal short term debt.

A to AA, reflecting a moderate yield spread.

In principle, our techniques could be used to produce bond ratings themselves. An important question is "what are we trying to measure?" with a bond rating. Is it probability of default during the debt's life, or yield spreads? The two are related but the relation is complex; the probability of default also requires the actual drift of the asset process. Yield spreads seem the more important variable to predict, since they are intimately related to market valuation. Yield spreads of newly-issued debt are given by

$$C/D - r = \frac{C(r+m)}{(C+mP)[1-(V/V_B)^{-y}]+(r+m)(1-\alpha)V_B(V/V_B)^{-y}} - r \tag{22}$$

where V_B is given by equation (19) or (20).

Bond ratings based on predicted yield spread ranges will reflect current asset value, risk, debt maturity, bankruptcy costs, the riskfree interest rate, and the (total) bond coupon and principal. In fact, examination of equation (22) suggests that the ratio $(V/V_B)^{-y}$ is the critical statistic for determining yield spreads, and therefore bond ratings. An examination of the comparative statics of this ratio are critical to the behavior of "our" bond ratings.

Current bond rating methodologies take many of the variables listed above into consideration, although exactly how they are combined is somewhat murky. Popular rating methodologies also focus on flow measures, such as interest coverage ratios, which are not directly evident in our approach. (However, coverage ratios may be indirectly reflected in the ratio V/V_B, since V_B depends upon C through equation (19) or (20), and V will reflect cash flow variables such as earnings before taxes and interest. Finally, note that the specification of V_T may also reflect cash flow considerations.)

It is important to recall why cash flows (and therefore coverage ratios), are *not* directly key to our analysis. Bankruptcy in our model is caused by a shortfall of equity *value* to raise the funds needed to service debt. Current cash flows could be negative, but if equity value remains, the firm need not be forced into bankruptcy.[31] Of course, asset value V--which is crucial in our analysis--will reflect past and projected cash flows.

IV. OPTIMAL LEVERAGE

We now examine the leverage ratio which maximizes firm value for alternative choices of debt maturity. If there are no limits on tax deductibility (i.e. arbitrarily large coupons can create further tax benefits), there may be no limit to the optimal amount of debt issued, for shorter duration debt. Ever larger coupon payments provide ever larger tax benefits; these large coupons do not provoke bankruptcy "quickly enough" for bankruptcy costs (and

[31] For example, if $m = 0$ (infinite maturity debt) carrying a yield spread of 200 bps, then for our example we find from (20) that $V_B = \$40.36$. We compute $\delta = .0793$ in this case (covering the \$6.55 coupon plus a 3% stock dividend, when $V = \$100$). If δ is associated with proportional cash flow (EBIT), then cash flow just before bankruptcy is .0793(40.36) = \$3.20. Clearly, equity financing is making up the difference between the \$6.55 coupon and the cash flow generated by the firm. This confirms that our bankruptcy-triggering condition is quite different than Ross' [1994] condition that bankruptcy occurs whenever cash flow falls beneath the required coupon payment. Of course, one could always find a δ such that the two conditions coincided; however, recall that V_B itself depends upon δ.

the loss of tax benefits) to offset the tax gains. This implies that there must be a limit to potential tax benefits. We model this by assuming that tax benefits are lost at an asset level V_T which increases with the coupon C. Following our earlier example, we assume

$V_T = 50 + 2.5C.$

With the base case parameters of our earlier examples, we can relate firm value v to the leverage ratios, for debt maturities from 0.25 years to 20 years. Figure 9 plots this relationship. It is assumed that the current asset value $V = 100$, and that the principal of debt equals its market value. Observe that the leverage ratio which maximizes firm value is larger for longer duration debt. The maximal firm value is also greater. Table I reports the actual values of variables at the optimal leverage, including the volatility of equity and debt.

TABLE I

Maturity	Coupon	Firm Value	Leverage	Yield Spread	σ_{Equity}	σ_{Debt}
0.25 yrs.	1.75	105.7	22.1%	0 bps	25.7%	0.0%
1.0 yrs.	2.25	107.1	27.9%	2 bps	27.8%	0.0%
5.0 yrs.	3.50	110.4	39.9%	45 bps	34.4%	1.5%
10.0 yrs.	4.25	112.0	45.7%	81 bps	35.3%	2.5%
20.0 yrs.	4.75	113.3	49.2%	103 bps	36.4%	3.7%
Infinity	5.60	115.4	54.6%	139 bps	37.6%	5.7%

Given these results, why would firms ever issue short term debt? At least one answer to this may be the differing agency costs associated with different debt maturities. We now show that shorter debt provides fewer incentives for increasing firm risk, and thus minimizes potential agency costs.

V. AGENCY EFFECTS: Debt Maturity and Asset Substitution

Since Black and Scholes [1973] and Jensen and Meckling [1976], it has been a tenet of financial economics that, after debt is issued, stockholders will wish to increase the riskiness of the firm's activities. This is presumed to transfer value from debt to equity--the problem of "asset substitution." The presumption follows from regarding equity as a call option on the underlying firm value, as indeed is the case when debt has no coupon, and taxes and bankruptcy costs are ignored--the case studied by Merton [1974].

Equity in our model, however, is not precisely analogous to an ordinary call option. First, there is no obvious "expiration date." Bankruptcy may occur at any time, when assets fall to the value V_B. Second, V_B itself will change with the risk of the firm's activities, as can be seen by equation (19). Finally, the existence of tax benefits (and their potential loss in bankruptcy) implies that debt and equity holders are not splitting a claim whose value depends only on the underlying asset value.[32]

---

[32] It has been brought to my attention that many of these reasons were anticipated by Long [1974].

Figure 8 examines the behavior of the derivative of equity value, $dE/d\sigma^2$, and the derivative of debt value, $dD/d\sigma^2$, as the underlying value V changes. If debt is riskfree (as will occur with all situations, when $V \to \infty$), $dE/d\sigma^2 \to 0$. But when debt is risky, the behavior of $dE/d\sigma^2$ is somewhat complex.

Panels 8a - 8d plot the two derivatives as V varies, for optimal debt levels (at $V = 100$: see Section IV) at maturities of 0.25, 5, and 20 years, plus a consol ($M = \infty$). The dotted line maps $dE/d\sigma^2$; the solid line maps $dD/d\sigma^2$. As $V \to \infty$, debt becomes risk free and the derivative of debt with respect to risk σ^2 approaches zero from below. Observe that

(i) For either short-term or intermediate-term debt, increasing risk will not benefit bondholders *or* shareholders, except as bankruptcy is imminent ($V \to V_B$).

(ii) The incentives for increasing risk are much more pronounced for longer term debt.[33] For very long term debt, $dE/d\sigma^2 > 0$ for all asset levels.

(iii) At all maturities, the incentives for increasing risk become positive for both stockholders and bondholders, as bankruptcy V_B is approached. However, incentives to increase risk become positive for stockholders before they become positive for bondholders.

---

[33] More exactly, the price measure of asset values for which debtholders and equityholders are in conflict about raising risk is larger, the longer the maturity of debt. Bankruptcy (and resultant distortions) may occur at higher asset values for short duration debt, but the "window" for incentive incompatibility is relatively small.

The incentive compatibility problem exists only for the range of V for which $dE/d\sigma^2 > 0$, and $dD/d\sigma^2 < 0$. For very short term debt ($M = 0.25$ years), this range is minuscule: $35.5 < V < 36$. For intermediate term debt ($M = 5.0$ years), the range is $45 < V < 51$. The range extends to $47 < V < 74$ with 20 year debt; for very long term debt long term debt the asset substitution problem exists whenever $V > 48$.

The extent of conflict between stockholders and bondholders increases when tax rates τ and bankruptcy costs α decline. This is because the outside parties have less claim on firm values, and the "game" between bondholders and stockholders approaches a zero sum game. Figures 8e and 8f illustrate the effect of increased risk on stock and bond values when both $\alpha = 0$ and $\tau = 0$, and the firm has 50% leverage. Here there is direct conflict between bondholders and stockholders, although the problem is much reduced (except near bankruptcy) when short term debt is used.

These results suggest that the "general" asset substitution problem has been overstated, except when debt is very long term, or when taxes and bankruptcy costs are minimal. The results do illustrate that incentive incompatibilities will arise as bankruptcy is approached. On the very brink of bankruptcy, incentive compatibility is again restored: both stock and bondholders want to raise risks to avoid bankruptcy costs and preserve the potential of tax shelters for debt.

VI. MULTIPLE CLASSES AND SENIORITIES OF DEBT

Multiple debt securities, with differing maturity and seniority, can be brought within the framework developed here. Each different debt issue will be valued according to equation (13), with the amount received by each class in bankruptcy reflecting its seniority.[34] The asset value V_B which triggers bankruptcy must satisfy the smooth-pasting conditions for equity, which will reflect the total value of the firm from equation (16) less the *sum* of debt values. Detailed results will be pursued in a future paper.

In principle, alternative sinking fund schedules could also be used (e.g., a constant fraction of initial rather than remaining principle is repaid each moment). But now units of debt issued at different times $\tau \leq t$ would no longer be identical at time t. Each vintage would have to be valued separately.

VII. CONCLUSIONS

This paper has developed a model of risky corporate bond prices of arbitrary maturity. Key to the model is a constant rollover rate of outstanding debt. Average debt maturity is the reciprocal of the rollover rate. By considering rollover rates from zero to infinity, we can study debt of any average maturity. The time homogeneity of a constant repayment of

---

[34] The amount each debt issue receives in bankruptcy will reflect its appropriate claim on asset value after bankruptcy costs. If all debt is of equal seniority, its proportion of bankruptcy value will equal its outstanding principal value relative to the total principal of all extant debt.

principal as well as coupon allows closed form solutions for pricing bonds of any maturity. We relate bond values to firm risk, leverage, bankruptcy costs, tax rates, dividends, and the riskfree interest rate.

Our results show that risky debt behaves very differently from riskless debt. Effective duration may be far shorter than Macaulay duration--and even become negative. Convexity can become concavity. This suggests that the proper hedging of fixed income portfolios must consider potential default risks.

The "term structure of yield spreads" -- the relationship between yield spread and maturity, for a given leverage -- exhibits patterns similar to those which have been observed empirically in a related context by Sarig and Warga [1989]. For low risk (low leverage) debt, yield spreads increase with maturity. This is reversed when leverage (and therefore risk) is very large. At intermediate leverage levels, yield spreads are humped, reaching a maximum at intermediate durations.

Yield spreads of current debt decrease as riskless rates rise. For newly-issued debt, a rise in riskless rates will to "tilt" yield spreads negatively: a rise in riskless rates will increase yield spreads of short term debt, but decrease spreads of long term debt. Our techniques also allow computation of the probabilities of default occurring over any horizon. This requires knowledge of the actual (not risk-neutral) drift of the asset value process, however.

Optimal capital structure depends upon debt maturity. Short term debt requires lower leverage ratios than those which are optimal for long term debt. Yield spreads increase markedly with maturity, at the optimal leverage ratio. An important conclusion is that firms should issue higher-rated short term debt than long term debt.

The fact that longer term debt generates higher firm value poses the interesting challenge of why firms issue short term debt. A possible answer to this question lies in the lessening of agency problems, specifically the problem of asset substitution. In contrast with conventional wisdom, our results show that stockholders of firms issuing short-term debt generally will not have an incentive to raise firm risk.

Our model has the virtue of simplicity. In several dimensions it might be thought simplistic. We ignore the possibility that riskfree rates may vary stochastically. And the firm always replaces retired debt which with the same amount of new debt--the same coupon, and the same principal. Extensions to include randomly varying riskfree rates and optimal dynamic adjustments remain for the future.[35]

---

[35] First steps in this direction have been taken by Fischer, Heinkel and Zechner [1989].

APPENDIX A[36]

Let f(t) be the density of the first passage time for a Brownian motion reaching a barrier from above. Let the barrier be at zero, and the starting point $k = ln(V/V_B)$.

Let $\mu = r - \delta - \sigma^2/2$. Then

(23)
$$f(t) = \frac{k}{\sigma\sqrt{2\pi t^3}} e^{-\frac{1}{2}[\frac{k+\mu t}{\sigma\sqrt{t}}]^2} dt$$

Let $K = k/\sigma(2\pi t^3)^{1/2}$, and $\lambda = [\mu^2 + 2(r+m)\sigma^2]^{1/2}/\sigma^2$. Then

(24)
$$\int_{t=0}^{\infty} e^{-(r+m)t} f(t)dt = \int_{t=0}^{\infty} Ke^{-(r+m)t} e^{-\frac{1}{2}(\frac{k+\mu t}{\sigma\sqrt{t}})^2} dt$$

$$= \int_{t=0}^{\infty} Ke^{-\frac{1}{2}(\frac{k^2}{\sigma^2 t} + \frac{(\mu^2+2(r+m)\sigma^2)t}{\sigma^2} + \frac{2k\mu}{\sigma^2})} dt$$

$$= \int_{t=0}^{\infty} Ke^{-\frac{1}{2}(\frac{k^2}{\sigma^2 t} + (\frac{\sqrt{\mu^2+2(r+m)\sigma^2}}{\sigma^2})^2(\frac{\sigma^4 t^2}{\sigma^2 t}) + \frac{2k\mu}{\sigma^2})} dt$$

$$= \int_{t=0}^{\infty} Ke^{-\frac{1}{2}(\frac{k+\lambda\sigma^2 t}{\sigma\sqrt{t}})^2} e^{-\frac{k}{2}(\frac{2\mu}{\sigma^2} - 2\lambda)} dt$$

$$= (\frac{V}{V_B})^{(\lambda - \frac{\mu}{\sigma^2})} [\int_{t=0}^{\infty} g(t)dt]$$

where g(t) is the first passage time of a process starting at k with volatility σ and drift $\lambda\sigma^2$.

---

[36] The author thanks Klaus Toft for assistance in this derivation.

For such a process, it is well known that the limiting probability of absorption (as $T \to \infty$) at V_B, when starting at V, is

(25)
$$\int_{t=0}^{\infty} g(t)dt = e^{\frac{-2\lambda\sigma^2 k}{\sigma^2}} = \left(\frac{V}{V_B}\right)^{-2\lambda}$$

Substituting this into the last line of equation (24) gives

(26)
$$\int_{t=0}^{\infty} e^{-(r+m)t} f(t)dt = \left(\frac{V}{V_B}\right)^{-(\lambda+\frac{\mu}{\sigma^2})} = \left(\frac{V}{V_B}\right)^{-y}$$

where

(27)
$$y = \lambda + \frac{\mu}{\sigma^2} = \frac{(r-\delta-.5\sigma^2)+[(r-\delta-.5\sigma^2)+2(r+m)\sigma^2]^{\frac{1}{2}}}{\sigma^2}$$

APPENDIX B

Following the Appendix of Leland [1994], the value v of the firm when $V_B \leq V \leq V_T$ is

(28) $$v = V + A_1 V + A_2 V^{-x} - \alpha V_B \left(\frac{V}{V_B}\right)^{-x}$$

where from equations (49) and (50) in Leland [1994] we have

(29) $$A_1 = \left(\frac{\tau C}{r}\right)\left(\frac{x}{x+1}\right)\left(\frac{1}{V_T}\right)$$

(30) $$A_2 = -\left(\frac{\tau C}{r}\right)\left(\frac{x}{x+1}\right)\left(\frac{V_B^{x+1}}{V_T}\right)$$

Equity value $E = v - D$, where D is given in equation (13). Taking the derivative of E with respect to asset value V gives

(31) $$\frac{\partial E}{\partial V} = 1 + A_1 - x A_2 V^{-x-1} + \alpha x \left(\frac{V}{V_B}\right)^{-x-1}$$

$$-\left(\frac{y}{V_B}\right)\left[\frac{C+mP}{r+m} - (1-\alpha)V_B\right]\left(\frac{V}{V_B}\right)^{-y-1}$$

Evaluating the derivative at $V = V_B$ and invoking the smooth-pasting condition gives

(32) $$1 + A_1 - x A_2 V_B^{-x-1} + \alpha x - \left(\frac{y}{V_B}\right)\left(\frac{C+mP}{r+m} - (1-\alpha)V_B\right) = 0$$

Substituting for A_1 and A_2 from (29) and (30) and solving for V_B gives the desired result

(33) $$V_B = \frac{y(C+mP)rV_T}{(r+m)[rV_T[1+\alpha x+(1-\alpha)y]+\tau Cx]}$$

Following Leland's [1994] Appendix A, it can then be shown that, when $V > V_T$,

(34)
$$v = V + \tau C/r + B_2 V^{-x} - \alpha V_B \left(\frac{V}{V_B}\right)^{-x}$$

where

(35)
$$B_2 = -\left(\frac{\tau C}{r}\right)\left(\frac{x}{x+1}\right)\left(\frac{1}{V_T}\right)\left(V_B^{x+1} + \frac{V_T^{x+1}}{x}\right)$$

REFERENCES

Black, F., and J. Cox, 1976, Valuing corporate securities: Some effects of bond indenture provisions, *Journal of Finance* 31, 351-367.

Black, F., and M. Scholes, 1973, The pricing of options and corporate liabilities, *Journal of Political Economy* 81, 637-654.

Brennan, M. and E. Schwartz, 1978, Corporate income taxes, valuation, and the problem of optimal capital structure, *Journal of Business* 51, 103-114.

Fischer, E., R. Heinkel, and J. Zechner, 1989, Dynamic capital structure choice: Theory and tests, *Journal of Finance* 44, 19-40.

Harris, M. and A. Raviv, 1991, The theory of capital structure, *Journal of Finance* 44, 297-355.

Jensen, M. and W. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs, and ownership structure, *Journal of Financial Economics* 4, 305-360.

Kim, I., K. Ramaswamy, and S. Sundaresan, 1993, Does default risk in coupons affect the valuation of corporate bonds?: A contingent claims model, *Financial Management*, 117-131.

Leland, H., [1994], Corporate Debt Value, Bond Covenants, and Optimal Capital Structure, *Journal of Finance*, September.

Long, J., 1974, Discussion, *Journal of Finance* 29, 485-488.

Longstaff, F., and E. Schwartz, 1992, Valuing risky debt: A new approach, Working Paper, UCLA (March; revised September 1993).

Macaulay, F. [1938], *Some Theoretical Problems Suggested by the Movements of Interest Rates, Bond Yields, and Stock Prices in the U.S. Since 1856*, National Bureau of Economic Research: New York.

Merton, R., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449-469.

Miller, M., 1977, Debt and taxes, *Journal of Finance* 32, 261-275.

Pitts, E., and Selby, M., 1983, The Pricing of Corporate Debt: A Further Note, *Journal of Finance* 38, 1311-1314.

Ross, S., 1994, "Capital Structure and the Cost of Capital," Financial Economics Working Paper #43, Series F, Yale University, August.

Sarig, O., and A. Warga, 1989, Some empirical estimates of the risk structure of interest rates, *Journal of Finance* 44, 1351-1360.

Smith, C. and J. Warner, 1979, On financial contracting: An analysis of bond covenants, *Journal of Financial Economics* 7, 117-161.

Toft, K., 1993, Options on leveraged equity with default risk, Working Paper, Haas School of Business, University of California, Berkeley.

FIGURE 1: Debt Value

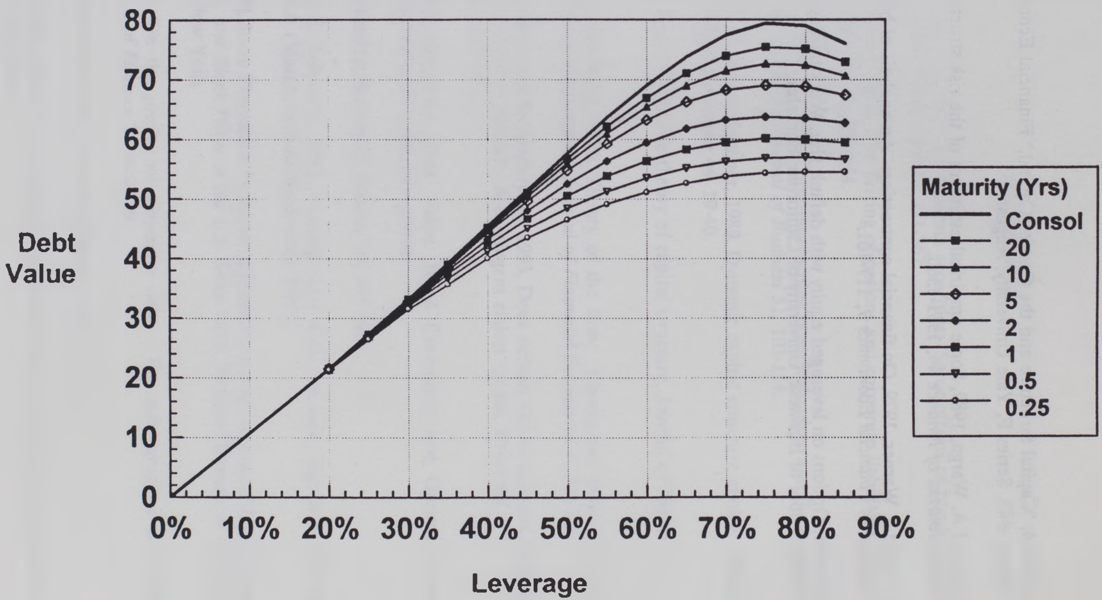


FIGURE 2: Term Structure of Yield Spreads

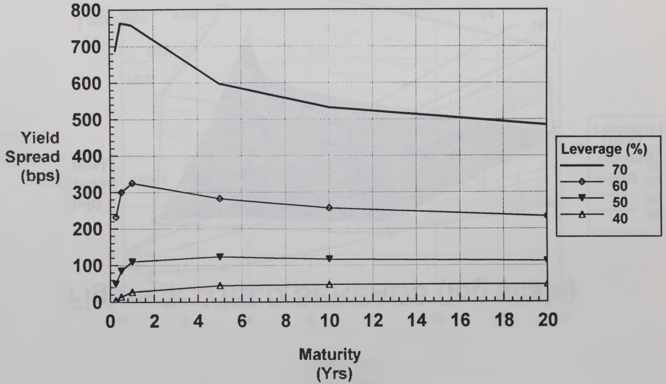


Figure 3: YIELD SPREADS (log scale)

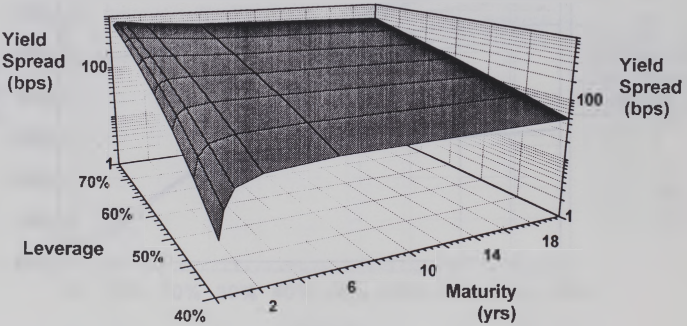


Figure 4: DURATION (given Leverage)

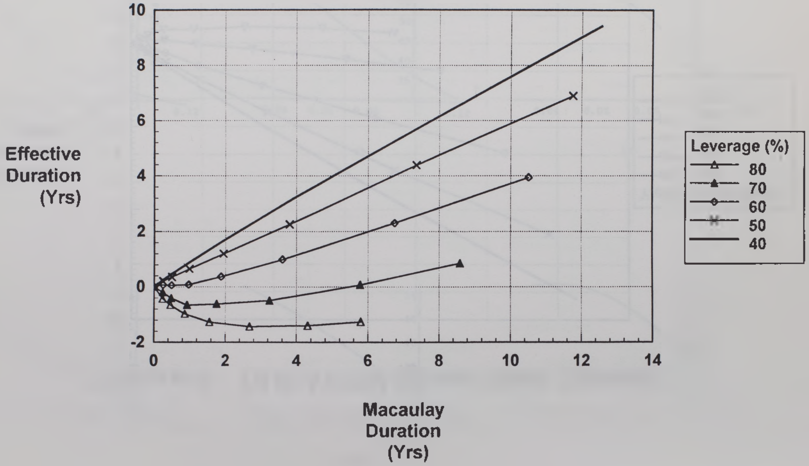


Figure 5: DURATION (given Yield Spread)

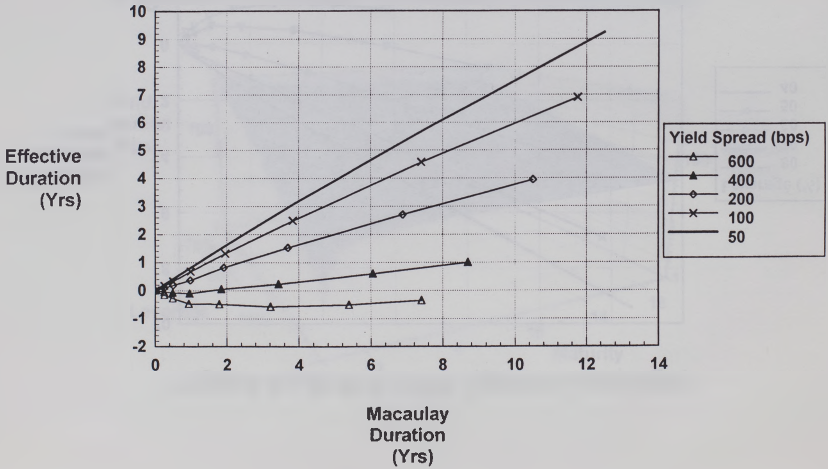


Figure 6

Convexity of Debt Values

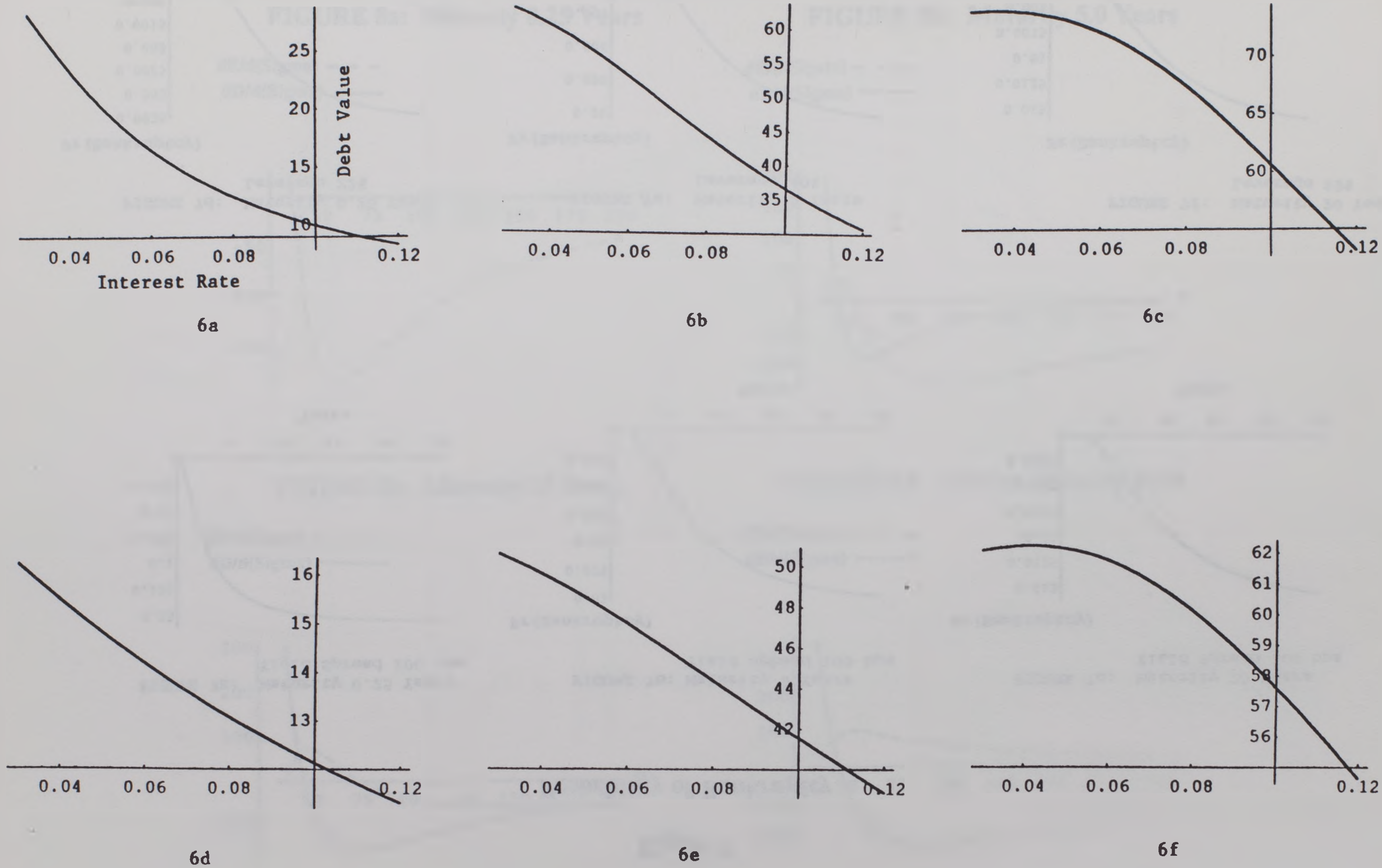


Figure 7

Probability of Bankruptcy

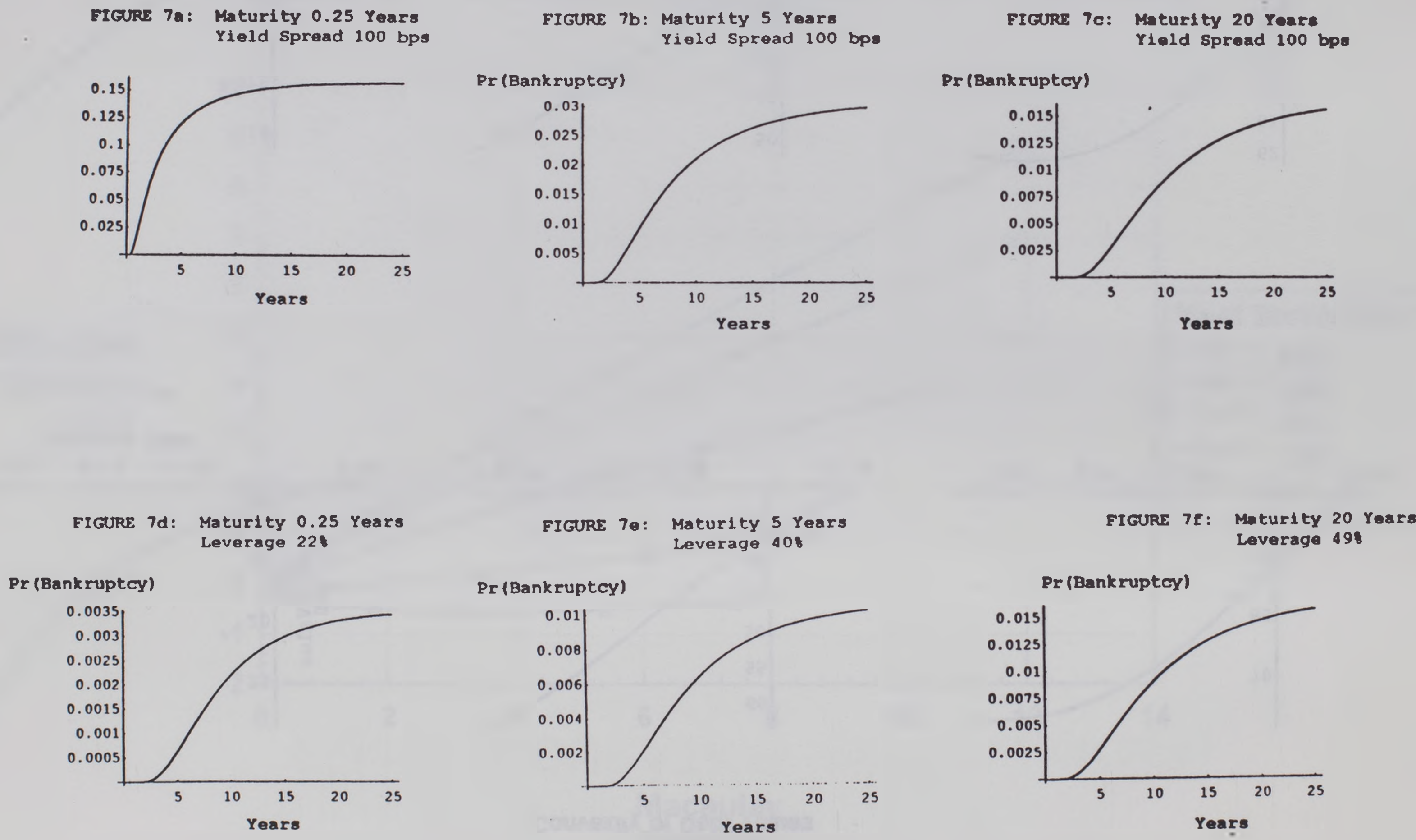


Figure 8: Effect of Increase in Risk σ on Bond and Equity Values

FIGURE 8a: Maturity 0.25 Years

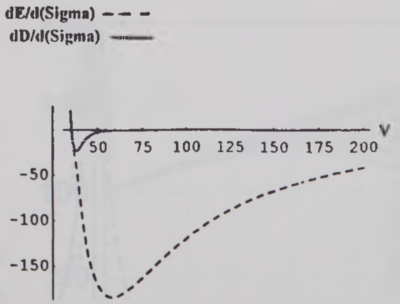


FIGURE 8b: Maturity 5.0 Years

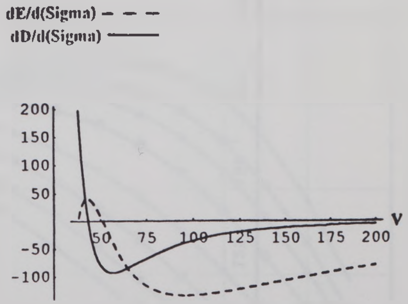


FIGURE 8c: Maturity 20 Years

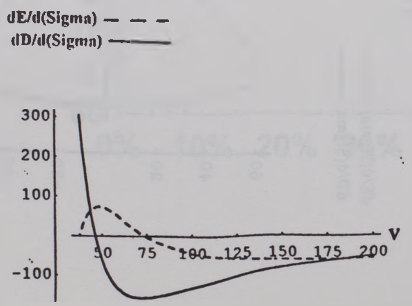


FIGURE 8d: Infinite Maturity Debt

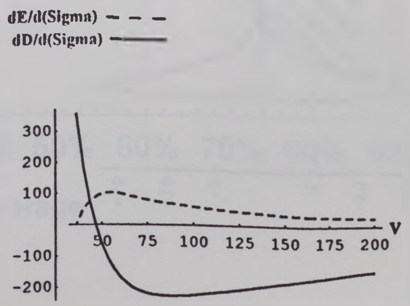


FIGURE 8e: Maturity 0.25 Years

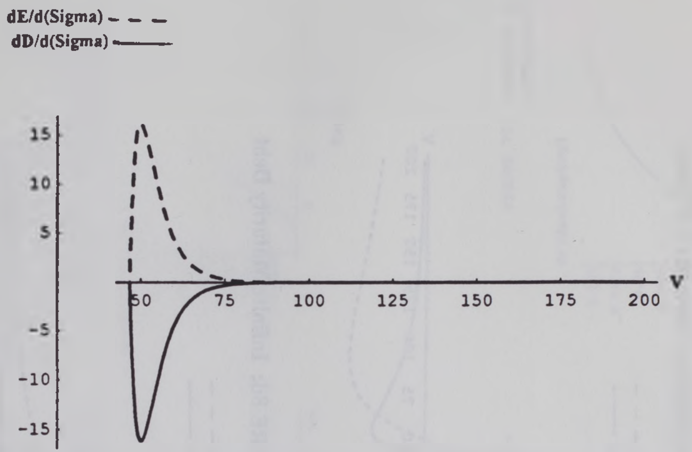


FIGURE 8f: Maturity 5 Years

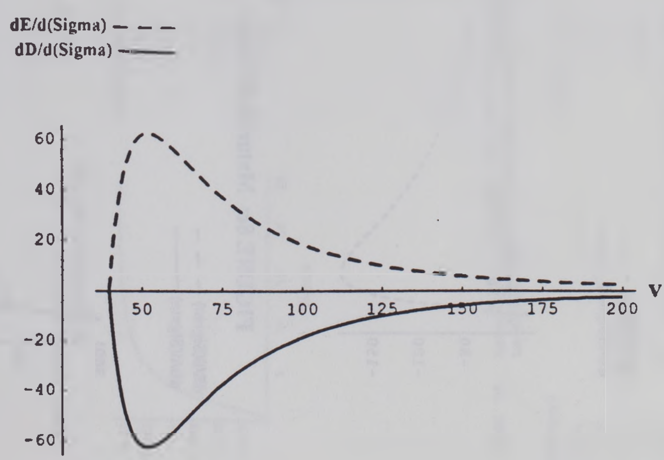
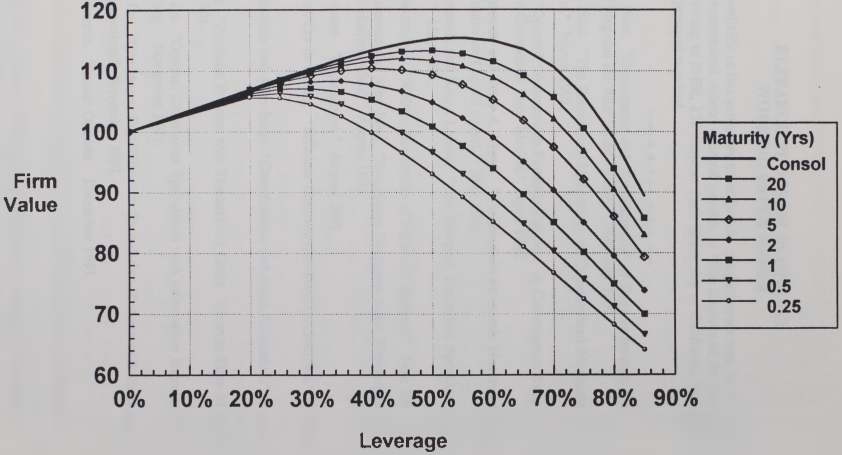


FIGURE 9: Firm Value



RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" --by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

207. Steven E. Plaut. "Reinsurance and 'Securitization' of Deposit Insurance; A Workable Proposal for Risk-Based Pricing." February 1991.

208. Steven E. Plaut. "The Prepayment Uncertainty of Collateralized Mortgage Obligations." March 1991.

209. Hua He. "Optimal Consumption-Portfolio Policies: A Convergence from Discrete to Continuous Time Models." April 1991.

210. Gerard Gennotte and Terry A. Marsh. "Variations in Economic Uncertainty and Risk Premiums on Capital Assets." May 1991.

211. Gerard Gennotte and Hayne Leland. "Low Margins, Derivative Securities, and Volatility." May 1991.

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993..

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance

WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 241

On the Accounting Valuation of Employee Stock Options

by

Mark Rubinstein

December 1994

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE
WALTER A. HAAS SCHOOL OF BUSINESS,
UNIVERSITY OF CALIFORNIA AT BERKELEY

The Executive Publication Series is produced by the Research Program in Finance in the Walter A. Haas School of Business at the University of California. The Research Program in Finance has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The Executive Publication Series was established in 1989.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH
Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

On the Accounting Valuation of Employee Stock Options

Mark Rubinstein

Haas School of Business
University of California, Berkeley

December 1994

Finance Working Paper #241

Mark Rubinstein is a professor of finance at the University of California, Berkeley. This paper arose out of a consulting project for Intel Corporation. The author thanks Robert Sprouse for his accounting courses at Stanford, Jim Ohlson for instructive conversations on accounting over many years, and Stephen Penman for assistance with employee stock options.

Abstract

In its exposure draft, "Accounting for Stock-based Compensation," FASB proposes that either the Black-Scholes or binomial option pricing model be used to expense employee stock options, and that the value of these options be measured on their grant date with typically modest ex-post adjustment. This brings the accounting profession squarely up against the Scylla of imposing too narrow a set of rules that will force many firms to misstate considerably the value of their stock options and the Charybdis of granting considerable latitude which will increase non-comparability across financial statements of otherwise similar firms. This, of course, is a common tradeoff afflicting many rules for external financial accounting.

It is not my intention to take a position on this issue, but merely to point out the inherent dangers in navigating between these twin perils. To examine this question, this paper develops a binomial valuation model which simultaneously takes into consideration the most significant differences between standard call options and employee stock options: longer maturity, delayed vesting, forfeiture, non-transferability, dilution, and taxes. The final model requires 16 input variables: stock price on grant date, stock volatility, stock payout rate, stock expected return, interest rate, option striking price, option years-to-expiration, option years-to-vesting, expected employee forfeiture rate, minimum and maximum forfeiture rate multipliers, employee's non-option wealth per owned option, employee's risk aversion, employee's tax rate, percentage dilution, and number of steps in the binomial tree. Many of these variables are difficult to estimate. Indeed, a firm seeking to overvalue its option might report values almost double those reported by an otherwise similar firm seeking to undervalue its options.

The alternatives of expensing minimum (zero-volatility) option values, whether at grant or vesting date, can easily be gamed by slightly redefining employee stock option contracts, and therefore would not accomplish FASB's goals.

As an alternative, FASB could give more careful consideration to exercise date accounting, under which an expense is recognized at the time of exercise equal to the exercise value of the option. This would achieve the long sought external accounting goal of realizing stock options as compensation, while at the same time minimizing the potential for the revised accounting rules to motivate gaming behavior or non-comparable statements.

On the Accounting Valuation of Employee Stock Options

Employee stock options are call options given by employing firms to their employees in compensation for labor services. Typically, at the time an option is granted, its striking price is set equal the firm's concurrent stock price. Usually, during the first portion of its life (the vesting period), the employee cannot exercise his options and in fact must forfeit them should he be fired or voluntarily resign. After the vesting date, typically three years after the grant date, the employee can exercise his options at any time until maturity (usually seven years after the vesting date) but cannot sell or otherwise transfer them. Indeed, if he leaves the firm during this period, he is usually forced to choose between forfeiting or exercising his options within a short time after his departure.

A survey by Coopers & Lybrand indicates that "long-term incentive executive compensation" for U.S. corporations grew from 20% of total compensation in 1982 to 31% in 1992.[1] About 40% of corporations with revenues less than $100 million have long-term incentive plans, and 78% of those with revenues above $10 billion have such plans. Non-qualified stock options, the subject of this paper, are by far the most popular method of long-term compensation.

Currently, in the United States, such options granted at-the-money, even though they are granted in lieu of cash compensation for labor services, are not considered an expense under generally accepted accounting principles. That is, they are not charged against earnings at grant, at vesting date, upon exercise, or at any other time.

For example, compare two otherwise identical firms, one which uses only cash compensation and the other which substitutes stock options for half its compensation. Under current rules, the second firm will report less compensation expense and therefore greater aggregate earnings and, at least initially, greater earnings per share. This situation clearly violates a key objective of the Financial Accounting Standards Board (FASB): nearly identical firms should report nearly identical earnings. Perhaps, the chief reason FASB has not corrected this situation earlier has been the difficulty of measuring the expense. More recently, persuaded by advances in option pricing methods, in the Exposure Draft "Accounting for Stock-based Compensation," FASB proposes that a modified version of either the Black-Scholes or binomial option pricing model[2] be used to value employee stock options and that this value be recognized as an expense on the grant date.[3]

---

[1] Coopers & Lybrand, *Stock Options: Accounting, Valuation and Management Issues*, New York (1993).

[2] See Fischer Black and Myron Scholes, "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy* (May/June 1973) and John Cox, Stephen Ross and Mark Rubinstein, "Option Pricing: A Simplified Approach," *Journal of Financial Economics* (September 1979).

[3] Financial Accounting Standards Board, "Accounting for Stock-based Compensation," Exposure Draft, #127-C (June 30, 1993). FASB's confidence in modern option valuation techniques is indicated by the following quotation from the Exposure Draft:

> "Trading of options in the financial markets has increased significantly in the last 20 years. During that time, mathematical models to estimate fair value of options have been developed to meet the needs of investors. Software available for personal computers reduces the application of those models to a fill-in-the-blank exercise."

The public reaction to FASB's proposal was extraordinary. Several groups representing corporate executives and boards of directors, institutional investors, all of the big six accounting firms,[4] and Secretary of the Treasury Bensten vociferously lobbied FASB, the U.S. Congress, and the SEC to drop the proposal. Responsive to this pressure, FASB held public forums as well as an academic roundtable in April, 1994 (which I attended) to reconsider the question. On May 3, the United States Senate for the first time in its history conducted a debate over external (not tax) accounting standards. It passed a non-binding resolution, 88 to 9, expressing opposition to FASB's proposal.[5] In June, as a result of this and further analysis, FASB decided to postpone implementation of its proposal and to restudy the question of expensing employee stock options.

What could have caused such an unprecedented reaction? If FASB's proposal were adopted, many firms, particularly in high-tech areas, would report substantial reductions on the order of 25% in earnings per share.[6] It is feared that such reductions would be translated into commensurately reduced stock prices. Note that it is not the disclosure of the estimated option values that has met with objection, but rather the recognition of these values in income statements and balance sheets. Additionally, many firms favoring stock options as a means of top management compensation may not want the high levels of this compensation to become transparent. It is also argued that incentives provided by stock options have been the engine of growth in successful newly developed U.S. industries, and that discouraging the use of these options through required expense recognition would deprive some of the country's most important corporations of a management tool crucial to success against foreign competition.

These arguments are all seriously flawed. While reported earnings per share would certainly fall to permanently lower levels for many firms, the claim that this will lead to lower stock prices presumes either that the revised earnings supplies new information to the market or that the market is quite inefficient at digesting available information in security prices. Since there seems to be little objection to disclosure, the presumption behind the argument must be extreme inefficiency, which in the light of most academic empirical evidence -- relating both to previous accounting changes such as the shift from FIFO to LIFO accounting and recognition of pension obligations, as well as to many other studies of market efficiency -- seems highly unlikely. Moreover, if stock prices decline, one could easily argue that the recognition of the expense simply increases market efficiency and improves resource allocation in the economy. After all, stock prices can be inefficiently priced too high as well as too low. It is also possible that recognition of

---

[4] To quote from a letter signed by all the big six accounting firms to the FASB dated July 15, 1994:

> "... we believe that the best solution is to withdraw the proposal to change the accounting and, instead, expand disclosures. ... If the Exposure Draft proceeds to a final standard, many companies have indicated that their stock-compensation plans will have to be curtailed or otherwise modified to manage an expense charge that they do not accept as either meaningful or representationally faithful."

[5] Senator Joseph Lieberman went so far as to co-sponsor a bill that would, if passed, have overruled any final FASB decision to change accounting for stock options.

[6] The Coopers & Lybrand study, from a sample of 27 firms, reports that the estimated average reduction in earnings after the phase-in period required in the Exposure Draft is 3.4% for "mature" firms and 26.5% for "emerging" firms.

stock option compensation may improve the allocation of resources within firms by forcing them to come to grips with true cost of their compensation plans. In any event, it is not the intended role of FASB to concern itself with the consequences of accounting rules for resource allocation; rather its role is to provide a framework in which the relevant corporate information is made cheaply available for all investors, permitting them to make informed investment decisions, whatever they may be.[7]

A more serious and sophisticated objection, and one which I will argue has merit, is that adoption of FASB's proposal in its current or reasonably modified form could lead to even greater non-comparability of accounting statements than we have in the current situation where most stock option plans are valued at zero. In the fields of finance and economics, the primary interest lies in how assets and securities are valued. But in the field of accounting, knowledge of valuation is not sufficient; *in addition, firms need to be induced to report correct values*. That is one reason why GAAP do not value inventories and plant and equipment at market. Too often market prices are not directly observable, and attempted marking-to-market would give firms free reign to make highly subjective estimates which may make external accounting statements less comparable.

I will argue that employee stock options differ from standard call options in significant ways. Nonetheless, for the most part, these can be incorporated into a generalized binomial model. Unfortunately, it seems that reasonable individuals can easily make different estimates of critical inputs which can lead to substantially different values. In addition, recent empirical work has questioned the validity of either the Black-Scholes or standard binomial model, even as it is applied to short-term exchange-traded options.

I. Problems in Applying Standard Option Pricing Techniques to Exchange-Traded Options

Assuming the Black-Scholes or standard binomial model is correct for valuing short-term exchange-traded options, there still remains the difficult task of estimating volatility. Commonly used historical estimates of volatility can vary over a significant range depending on the length of the historical period and the sampling frequency selected during the period. For example, selecting a period at random, estimating volatility for the S&P 500 index on September 30, 1986 from recent past historical index changes produces the following estimates:

---

[7] A caveat: FASB must also deal with the difficult trade off between providing relevant information and requiring firms to release information which could damage their competitive position in their industry.

Table I
Sensitivity of Historical Volatility to Sampling Period and Frequency

sampling period	sampling frequency: 5 minutes	1 day	2 days
1 day	31%		
1 week	25%	19%	
1 month	34%	28%	26%
2 months	28%	22%	21%
3 months	26%	22%	21%

Choice of the sampling period and frequency is currently an art, not a science. As a result practitioners use a wide variety of procedures, including complications related to differential measurement of intraday, overnight, weekend, and holiday volatility, and, in more sophisticated approaches, explicit methods for measuring volatility in the presence of acknowledged non-stationarity of historical time-series. For example, consider a benchmark standard European call at-the-money with underlying stock price and striking price of $100, time-to-expiration of one-year, an annualized dividend yield of 3.5% and an interest rate of 8%: near the extremes of volatility shown above, 21% and 34%, such an option would have a Black-Scholes value of $10.09 or $14.88, respectively.

II. Differences Between Exchange-Traded Options and Employee Stock Options

Complicating these issues further, apart from accounting treatment, employee stock options differ from exchange-traded options in seven important respects:

(1) Maturity: their maturity is much longer, typically 10 years;

(2) Delayed Vesting: through delayed vesting, exercise is usually not permitted for a period after grant, typically 3 years;

(3) Forfeiture: employees will lose unvested options when they leave their jobs and may be forced to exercise prematurely then unexercised but vested options;

(4) Non-Transferability: employees are usually not permitted to sell their options; so that the value of an option to the employee and his optimal exercise strategy is affected by his personal aversion to bearing risk, by his personal probability beliefs concerning his employer's future stock price, by the nature of his labor income, and by any other options or assets he may be holding;

(6) Taxes: non-qualified employee stock options[8] granted at-the-money are not taxed at grant, but are taxed at exercise at the employee's ordinary income tax

[8] Most employee stock options granted since the Tax Reform Act of 1986 are non-qualified (NQO). In contrast, the profits of incentive stock options (ISO) are not taxed to the employee until the stock acquired though exercise is sold, and then the tax is assessed at the capital gains tax rate. However, this advantage is usually more than offset by the fact the employing firm receives no tax deduction for this form of compensation.

rate based on the difference between the firm's stock price at that time and the striking price, and simultaneously give rise to an offsetting taxable expense for the firm;[9]

(6) Capital Structure Effects: the exercise of the options causes the associated firm to issue new shares of common stock and to receive the striking price in cash upon exercise, which increases both the number of outstanding shares and the total level of funds in the firm; in addition, instead of paying for the options in cash, employees pay with their labor services, which leaves additional cash in the firm which can be used for other purposes;[10]

(7) Operating Income Effects: compensation in the form of options can have the effect of increasing revenues, reducing expenses, or increasing risk-taking through altered work incentives.

These differences significantly complicate the problem of valuing these options even if the Black-Scholes or standard binomial approach is used. FASB's Exposure Draft describes corrections to these approaches which attempt to deal with differences (3) and (4) only. To handle difference (3), for options valued with either the Black-Scholes formula or binomial trees, the resulting option value is adjusted downward by multiplying the value that would otherwise have obtained by one minus the probability of forfeiture through the vesting date. To handle difference (4), users of the Black-Scholes formula are to value an option by replacing the time-to-expiration of the option with its expected time-to-exercise or expiration, whichever comes first.

Below we consider the efficacy of these modifications in the light of a more complete model of employee stock option valuation which takes account of differences (1)-(6).

Difference (1) Maturity: The basic inputs into either the Black-Scholes or standard binomial option valuation approach are the underlying asset price, volatility and payout rate, the interest rate, and the option striking price and time-to-expiration. Particularly over long periods of time, it becomes difficult to estimate underlying asset volatility and payout, and even slight errors in payout measurements (which over shorter periods would not have been as important) can radically change calculated option values. For example, consider our benchmark standard European call at-the-money with stock price on the grant date and striking price of $100, annualized stock volatility of 30%, and interest rate of 8%. The following table shows how a long time-to-expiration of the call can make its Black-Scholes value very sensitive to the assumed dividend yield:

---

[9] If the option is granted in-the-money, compensation expense to the firm and income to the employee equal to the in-the-money amount may be required to be recognized at the time of grant.

[10] In contrast, stock appreciation rights are satisfied by a cash payment from the firm to its employees equal to the difference between the stock price and striking price on the exercise date. In this respect, they are similar to cash-settled exchange-traded index options.

Table II

Sensitivity of Black-Scholes Option Values to Dividend Yield

annualized dividend yield	years-to-expiration 1	10
2.5%	$13.99	$41.61
3.5%	$13.41	$35.59
4.5%	$12.84	$30.33

Options are European and at-the-money, with underlying stock price and striking price equal to $100, annualized stock volatility of 30% and interest rate of 8%. The options are valued using the Black-Scholes formula.

While an error of 1% in projected payout creates only about a 4% error in the calculated value of options maturing in one year, it creates a 15%-17% error for options maturing in ten years.

Estimation of dividend yield, while usually quite reliable over a single year, can be quite difficult over longer periods. Corporations that are currently growing rapidly and currently pay little or no dividends should be able to make a persuasive case that dividends could well increase markedly after about five years as the corporation matures and its growth rate diminishes. But such a forecast, while possibly accurate, is subject to considerable uncertainty and manipulation.

Errors resulting from volatility estimation, while not as sensitive to maturity, can nonetheless be quite substantial. For example, under the above situation with a dividend yield of 3.5%:

Table III
Sensitivity of Black-Scholes Option Values to Volatility

annualized volatility	years-to-expiration 1	10
25%	$11.56	$32.67
30%	$13.41	$35.59
35%	$15.25	$38.49

Options are European and at-the-money, with underlying stock price and striking price equal to $100, annualized dividend yield of 3.5% and interest rate of 8%. The options are valued using the Black-Scholes formula.

Here too, corporations that are currently growing rapidly can reasonably argue that volatility should gradually decline as the corporation's market matures and it becomes increasingly diversified across product lines, so that after 10 years volatility may reach much lower levels. Using the Black-Scholes formula, one should input the average volatility to be experienced during the life of an option, but in this case, this is likely to be considerably lower than the current volatility possibly implied in the market prices of its exchange-traded stock options.

A recent study, submitted by the firm Thermo Electron to FASB, examines over-the-counter warrants with lives of 5 to 10 years.[11] Of the roughly 300 existing warrant issues, 20 were of the right maturity and near-the-money at the time of the study. Using simple historical estimates of dividends and volatility, the study compares the standard binomial values of the warrants to their market prices. Of the 16 warrants with a history of zero dividends, 15 were overvalued by the model, using either 100-day or 3-year historical volatility. The average overvaluation of all 16 warrants was about 100%, and 13 of the 16 were overvalued by at least 30%. Interestingly, all 4 warrants with a positive history of dividends were *undervalued* by about 23%. This study is very suggestive of the naivety of estimating inputs to option models under the presumption that history is expected to repeat.[12]

FASB's Exposure Draft allows two alternative valuation approaches: Black-Scholes and binomial, and requires that the Black-Scholes approach use the expected life of the option in place of its time-to-expiration. Unfortunately, this can lead to exactly the wrong correction in many circumstances. Binomial trees are widely used for exchange-traded options, principally because -- unlike the Black-Scholes formula -- they explicitly take account of optimal early exercise permitted for American-style options. Since employee stock options can also be exercised early, binomial models should provide more accurate values. However, since other things equal, American exchange-traded option values are higher than Black-Scholes values and reducing the time-to-expiration in the Black-Scholes

---

[11] See "Valuation of Employee Stock Options," position paper presented at April 18, 1994 roundtable discussion of the Financial Accounting Standards Board, Thermo Electron Corporation.

[12] In "Pricing Warrants: An Empirical Study of the Black-Scholes Model and Its Alternatives" (*Journal of Finance*, September 1990), Beni Lauterbach and Paul Schultz also present evidence of difficulties of applying standard option models to long-term options.

formula reduces the values of calls, FASB's modification may tend to move computed values of employee stock options in the wrong direction. To get an idea of the magnitude of this bias, using the benchmark option, we can use a binomial tree to calculate the (risk-neutral) expected life of the option, known in the trade as the option "fugit".[13] For our benchmark option, the fugit is 9.14 years. Below we use this in the Black-Scholes formula to value a European option assumed to expire at that expected life.

Table IV
Sensitivity of Option Values to Exercise Assumption

exercise assumption	exercise	option value
Binomial (optimal early exercise)	optimal	$37.81
Black-Scholes (exercise only at expiration)	10 years	$35.59
Black-Scholes (exercise at expected option life)	9.14 years	$34.98

Options are at-the-money with time-to-expiration of 10 years, underlying stock price and striking price equal to $100, annualized stock volatility of 30%, annualized dividend yield of 3.5% and interest rate of 8%. The binomial calculations use a tree size of 200 steps.

Clearly, in this case, FASB's amended procedure has driven the option value even further than the naive Black-Scholes model from the optimal early exercise binomial value. For firms with dividend yields closer to the interest rate, since early exercise is even more desirable and therefore the fugit is smaller, this bias will be even larger. For example, in an otherwise identical situation, if the dividend yield were 4.5% instead of 3.5%, the fugit is 8.81 years and the Black-Scholes option value with this time-to-expiration is $34.70.

Difference (2) Delayed Vesting: Most option plans do not permit employees to exercise their granted options until after a predefined period of time has elapsed. The options then are neither European (can only be exercised at expiration) nor American (can be exercised at any time), but rather some hybrid which some have termed "Bermudian" (being between the United States and Europe). Fortunately, this difficulty can be easily handled by appropriately modifying the standard binomial model. Working backwards from the end of the tree, provided exercise is possible, at each node substitute the current early exercise value of the option for its current holding value if the former is greater. Then, as one continues to work backwards and enters the region where exercise is not possible, only use the current holding value at each node. However, this complication requires use of a modified binomial model. To see what effect early exercise can have on the value of an option, consider the same situation as above:

[13] Mark Garman, in his article, "Semper Tempus Fugit," *RISK* (May 1989), shows how to use binomial trees to calculate the risk-neutral expected life of an option by working backwards recursively from the end of the tree.

Table V
Sensitivity of Option Values to Delayed Vesting Method

delayed vesting method	option value
European (Black-Scholes at fugit)	$34.99
Bermudian (modified binomial)	$37.78
American (standard binomial)	$37.81

Options are at-the-money with time-to-expiration of 10 years, underlying asset price and striking price equal to $100, a volatility of 30%, a dividend yield of 3.5% and an interest rate of 8%. The Black-Scholes formula uses as the time-to-expiration the fugit of the Bermudian case of 9.16. The Bermudian and American option values are calculated using a 200 step binomial tree, and the modified binomial assumes that vesting occurs after the end of the third year in the life of the option.

Fortunately, the effect of delayed exercise is small in this case because it will usually not pay to exercise a ten-year option early in its life.

Difference (3) Forfeiture: The current value of granted options must be adjusted downward to account for the probability that an employee will be fired or voluntarily resign. As suggested in the Exposure Draft, this probability can be estimated actuarially across a large pool of employees. The value of the options is then simply adjusted downward by multiplying the value that would otherwise have been obtained by one minus the probability of forfeiture through the vesting date.

The anticipated forfeiture rate is another variable, like payout and volatility, that will have to be estimated. In many cases, it could be reasonably argued that history is a poor guide to the future because employment conditions have changed, and even if history is useful there are questions about how far back forfeiture rates should be averaged. Using past experience to estimate the termination rate is not easy, since past results are no doubt influenced by the degree of past success of the firm. For example, realized forfeiture rates are likely to be lower than ex-ante expectations during times when the stock price has risen rapidly.

The following table indicates how sensitive calculated option values are to this variable:

Table VI
Sensitivity of Bermudian Option Values to Forfeiture Rate

annualized forfeiture rate	option value
3.5%	$33.95
5.0%	$32.39
6.5%	$30.88

Options are Bermudian and at-the-money with time-to-expiration of 10 years, underlying asset price and striking price equal to $100, a volatility of 30%, a dividend yield of 3.5% and an interest rate of 8%. The Bermudian option values are calculated using a 200 step binomial tree with vesting occurring after the end of the third year in the life of the option. Forfeiture is considered by following FASB's procedure and multiplying the value of the option $37.78 by one minus the annualized forfeiture rate raised to the third power.

Even if the forfeiture rate can be measured exactly, there are several reasons why FASB's amended procedure is flawed.

First, the possibility of forfeiture continues to affect the values of most employee stock options even after the vesting date. Should an employee leave his job after his options have vested but before their expiration date, he is usually forced to exercise the options shortly after his departure. Since American call options are normally worth more alive than dead, this reduces the value of the options even further.

Second, FASB's approach ignores that the probability of forfeiture is no doubt negatively correlated with the success of the corporation. In particular, if the underlying stock price rises over the life of the options and perforce the options become quite valuable, employees are probably less likely to be fired or leave their jobs voluntarily. This means that to this extent the suggested approach will overstate the effect of forfeiture on the value of the options. If some firms account for this dependence and others do not, their external financial statements will not be comparable.

Third, the probability of forfeiture may be positively correlated with the time remaining to the vesting date, other things equal. The less time remaining, the less likely an employee will voluntarily resign and the less likely the employee will be fired since the employee has had additional time to prove his value to the firm. Therefore, the suggested approach to handling forfeiture needs to be revised to account for the changing average time to the vesting date of the actuarial pool of employees.

Fourth, simply multiplying by one minus the probability of forfeiture, either as proposed by FASB or as outlined above, presupposes that the market discounts the uncertainty associated with forfeiture as if it were risk-neutral toward this risk. This follows from a basic idea of modern financial economics that calculating the present value of uncertain income by discounting its future expected value by the interest rate is only justified if the risk of this income can be diversified away by holding a well-diversified portfolio. In fact, since for the reasons given above, this risk is likely to be negatively correlated with the underlying stock price, which, in turn, is likely to be positively correlated with the value of a well-diversified portfolio, its effect on valuation should be handled using risk-adjusted discounting -- a serious complication about which the theory of finance has no easy answers.

To get an idea of the significance of some of these flaws in FASB's approach, consider the following revised binomial tree. First, to address complication (1), suppose the annualized probability of forfeiture is a constant 5% and we are using a 200 step binomial tree to value an option maturing in 10 years. Then the probability of retention at any node in the tree is $(1-.05)^{10/200} = .99744$. Suppose at a given node the value of the option unexercised is A and its value exercised is B. As we work backwards in the tree, revise the calculated value of the option at each node as follows:

> if the option is out-of-the money or the node is before the vesting date, replace the value of the option at that node with $.99744 \times A$;
>
> if the option is in-the-money and the node is after the vesting date, replace the value of the option at that node with $(.99744 \times \max[A,B]) + (1-.99744) \times B$;

and continue to work backwards in the tree using these values. In our benchmark example, the value of a Bermudian option with 3-year delayed vesting before considering potential forfeiture is \$37.78. Under FASB's proposal, the value after forfeiture would be $\$37.78 \times .95^3 = \$32.39$. Using the above revised binomial tree, the value would instead be lower at \$30.75.

To address the second complication, suppose we use the value of an employee's options themselves to predict the probability of forfeiture. Presumably, other things equal, the higher the value of these options, the less likely he will be terminated. At very low values, assume he is about twice as likely to be terminated and at very high values assume he is half as likely to be terminated. In between, at step i, node j, assume the probability of being terminated is inversely proportional to $(\log C_{ij})/\Sigma_j P_{ij}(\log C_{ij})$, where C_{ij} is the value of his option at step i, node j, and P_{ij} is the probability of ending up at node j at step i, estimated at the beginning of the tree over all possible nodes at step i so that $\Sigma_j P_{ij} = 1$. Thus, roughly speaking, the higher the value of the option at step i, node j, relative to its expected value at step i, the lower the probability of being terminated at step i, node j. Without this adjustment we would have assumed that the probability of forfeiture at step i, node j was $1-.99744 = .00256$. This adjustment gives rise to probabilities of forfeiture $(.00256 \times .5) < \pi'_{ij} < (.00256 \times 2)$ which are negatively correlated with the option value at that step-node. Finally, to be consistent with an overall probability of forfeiture at that step of .00256, these probabilities must be scaled so that the final probabilities π_{ij} satisfy $\Sigma_j P_{ij} \pi_{ij} = .00256$. The following table shows this sensitivity:

Table VII
Sensitivity of Bermudian Option Values to Forfeiture

forfeiture assumption	option value
(FASB method)	$32.39
(revised binomial, constant rate)	$30.75
(revised binomial, correlated rate)	$31.63

Options are Bermudian and at-the-money with time-to-expiration of 10 years, underlying asset price and striking price equal to $100, a volatility of 30%, a dividend yield of 3.5% and an interest rate of 8%. The option values are calculated using a 200 step binomial tree, modified to allow vesting after the end of the third year in the life of the option. The average annualized forfeiture rate is 5%. For the second option, the binomial tree is modified to incorporate a constant 5% annualized forfeiture rate throughout the life of the option. For the third option, the tree is modified to include an expected annualized forfeiture rate of 5% with a realization which is negatively correlated as outlined above with the remaining option value.

Difference (4) Non-Transferability: Unlike exchange-traded options, employee stock options are not traded in a secondary market. Therefore, the only way an employee can liquidate her position is to exercise the options and then sell the stock she receives in the secondary market.[14] Since the wealth of many employees is poorly diversified and heavily tied by way of continued employment, cash bonuses and stock options to the performance of their employing firm (the very intention of a stock option program), employees may not value their stock options at as high a level as the Black-Scholes model or standard binomial model would suggest.

Since the option has two values (and the second a highly personal one depending on the preferences and financial circumstances of each employee), one might ask which should be used by the corporation in its external financial statements for the purpose of communicating with stockholders. Fortunately, the answer is clearly that the corporation should value the option according to the effect the existence of the option, other things equal, has on the value of its stock -- not value the option from the employee's point of view -- a position correctly taken in FASB's Exposure Draft. In addition, the argument below shows that since this "compensating differential" can only arise during the vesting period, it is not likely to be a large amount.[15]

---

[14] As an alternative, an employee could consider short-selling his employer's stock. Aside from the usual problems faced by most investors from the loss of the interest on the proceeds of short sale, an employee must face the reputational difficulties short-selling might entail from this circumvention of the incentives intent of the stock options. In addition, for officers and directors, Section 16-b of the 1934 Securities Act requires that any profits generated by short selling an employer's stock that occur within a six month period following the short sale, whether or not they are actually realized during that time, must be returned to the firm. As a result of these constraints, I suspect that short sales of employer's stock are quite rare.

[15] FASB's proposal advocates amortizing the value of the options over the vesting period. This would be a reasonable procedure if employees could sell their options in the secondary market immediately after vesting. However, because they can not, vested options continue to provide work incentives for employees until the options are exercised and the stock is sold. After vesting, the employee faces a dilemma: on the one hand, he would like leave his options unexercised because of their remaining time value, but on the other, he would like to exercise them to increase his diversification. Of course, if he could sell his options he would

Even so this difference in the way diversified investors and employees look at the options creates problems in determining the exercise strategy assumed in the valuation. The standard binomial model, implicitly presupposing a secondary market for the option, assumes that it would be optimal to exercise an option whenever its discounted risk-neutral expected value is less than its current exercisable value. However, it is likely that pressures to diversify her source of income may cause an employee to exercise her options much earlier than would be optimal for a well-diversified investor. As long as this potential for premature exercised is considered when evaluating an option, except for the exercise prohibition during the vesting period, there will be no difference between the value of the option to the employee and the cost to the firm since the employee forces its value to her to equal its cost to the firm by following the exercise strategy which is in the employee's best interest.[16]

To get an idea of how much this cause of premature exercise can affect the value of an option, we will superimpose upon our current model a highly simplified exercise strategy specially designed to preserve the single state-variable binomial approach.[17] Assume that for each of **N** granted stock options, an employee has a total of **A** dollars of non-option wealth, all currently invested in riskless assets at interest return over a single binomial move **r**. The value to the employee of his entire portfolio provided he holds the options to expiration is:

$$W(j;n) = N \times \{ \max[0, u^j d^{n-j} S - K] + Ar^n \}$$

where **j** is the number of up moves with capital gain return **u** and **n-j** is the number of down moves with capital gain return **d** out of a total of **n** steps in the binomial tree, **S** represents the stock price on the grant date, and **K** is the striking price of the options.

Assume furthermore that the employee's utility function is in the class of myopic functions:

$$U(j;n) = W(j;n)^{1-b}/(1-b) \quad \text{for} \quad 0 < b$$ [18]

where the greater **b**, the more risk aversion. In this case, since utility is unique up to an

---

probably do so, but this alternative is not open. As long as he retains his options, the "forced" concentration of his wealth in his employing firm may cause him to work harder. This argues that correct matching of revenues with expenses requires that only part of the option cost be amortized during the vesting period, and that the remainder be amortized from the end of the vesting period to the date of exercise or expiration, whichever comes first.

[16] Inaccurate handling of the exercise strategies of employees would not be as significant if there were a way to correct these errors retroactively based on realized behavior. Indeed, FASB proposes that after options have either been exercised or expired, the options be revalued using the realized life of the options in place of their expected life and the financial statements trued-up accordingly. Unfortunately, this retroactive procedure does not make sense. To see this, options which end up in-the-money are likely to be exercised early and therefore lead to subsequent downward adjustment in their values. On the other hand, options which remain out-of-the-money, will never be exercised, leading either to no adjustment or subsequent upward adjustment in their values. So we have the embarrassing situation where options which turn out to provide high payoffs to employees will, in the end, after the proposed ex-post correction, be valued much lower than options which turn out to be worthless.

[17] The model used here of the effects of non-transferability on the employee's exercise decision is adapted from Alan Marcus and Nalin Kutalilaka, "Valuing Employee Stock Options," *Financial Analysts Journal* (November/December 1994).

[18] If b equals 1, then the utility function is its limit as b approaches 1, which is log(W).

increasing linear transformation, the employee's utility will be independent of the scale of his wealth N (so henceforth we will ignore N).

Let **E** be the investor's own subjective annualized expected return of the underlying stock. Assume also that the investor believes that the stock rate of return follows a stationary random walk. In a binomial model, this implies that at each node in the tree, the expected stock return over the next move is:

$$E^h = qu\delta + (1-q)d\delta$$

where **h** ≡ t/n (the ratio of the years-to-expiration of the option divided by the number of steps in the tree), **δ** is one plus the dividend yield over the next move, and **q** is the subjective probability of an up move. Thus, taking E as given, we can derive q as:

$$q = ((E^h/\delta) - d)/(u-d).$$

The employee can calculate his expected utility and exercise strategy recursively by using the following procedure. For an earlier period **k**, conditional on not exercising his options during this period, his expected utility is:

$$E_H[U(j;k)] = qE[U(j+1;k+1)] + (1-q)E[U(j;k+1)]$$

on the other hand, conditional on exercising his options, his expected utility is:

$$E_X[U(j;k)] = \{ (\max[0, u^j d^{k-j} S - K] + Ar^k)r^{n-k} \}^{1-b}/(1-b)$$

His actual expected utility will be:

$$E[U(j;k)] = \max\{ E_H[U(j;k)], E_X[U(j;k)] \}$$

This model of early exercise makes three highly simplifying assumptions:

(1) the only assets the employee holds are his non-transferable stock options and cash;

(2) at each date after vesting, the employee either exercises none or all of his options;

(3) upon exercise, the employee immediately sells his stock and reinvests the proceeds in cash and remains 100% invested in cash through the expiration date.

Thus, in this simplified model, in addition to the information required before, knowledge only of the investor's initial non-option wealth A, his subjective stock expected return E, and his risk aversion b is enough to determine the employee's optimal exercise strategy. Each of these variables is quite difficult to estimate. Non-option wealth not only includes the employee's holdings of real estate and securities outside his employing firm, but also includes some fraction of the present value of his human capital which is not solely dependent on the fortunes of his currently employing firm. Not only are expected returns subjective but they are also notoriously difficult to estimate from historically observed returns.[19] For the U.S. population as a whole, various academic studies have estimated risk aversion b in the range of 1 to 10, and many start-up or high-tech firms may self-select employees with even lower risk aversion.

The following table indicates how sensitive calculated option values are to these variables:

Table VIII
Sensitivity of Bermudian Option Values to Non-Transferability Variables

	risk aversion (b)								
	.5			2			4		
non-option	expected return			expected return			expected return		
wealth (A)	10%	15%	20%	10%	15%	20%	10%	15%	20%
30.00	37.60	35.56	35.60	32.82	35.66	37.77	29.12	31.46	32.97
60.00	37.77	35.56	35.56	34.69	37.27	36.56	31.67	33.77	36.00
120.00	37.76	35.56	35.56	36.36	37.75	35.79	33.82	36.14	37.70

Options are Bermudian and at-the-money with time-to-expiration of 10 years, underlying asset price and striking price equal to $100, a volatility of 30%, a dividend yield of 3.5%, an interest rate of 8%, and vesting occurs after the end of the third year in the life of the option. The Bermudian options are calculated with a 200 step binomial tree. The employee is assumed to base his exercise strategy on a myopic utility function of wealth at option maturity with risk aversion b; the only assets the employee holds are his non-transferable stock options and cash (equal to A on the grant date); at each date after vesting, the employee either exercises none or all of his options; and upon exercise, the employee immediately sells his stock and reinvests the proceeds in cash.

The numbers in this table, which do not reflect the possibility of forfeiture, should be compared to $37.78 from Table V. This is an upper bound on the values in Table VIII since restrictions on non-transferability (which lead to non-optimal exercise behavior from the point of view of an investor with access to a secondary market) should only serve to decrease option values.[20]

---

[19] For example, assume the stock return follows a stationary random walk with standard deviation 30%. Even after an historical sample covering 25 years, the standard deviation of the historically sampled mean is $30\%/\sqrt{25} = 6\%$. Even worse, since we don't inhabit a stationary random walk world, this should be regarded as a lower bound.

[20] One might have expected that the option values in this table should have been increasing in non-option wealth and expected return, but decreasing in risk aversion. Indeed, had the dividend yield been zero, such would have been the case. With positive dividends, had the options been traded in a secondary market, to maximize their market value it would pay to exercise them early under some circumstances. However, in the absence of a secondary market, increases in non-option wealth or expected return, or decreases in risk aversion, may cause an employee to postpone this exercise, thereby reducing the value of the option, not to him, but to the issuing firm.

Here is another curious anomaly. Other things equal, standard options are more valuable the greater the volatility of their underlying asset. In the case of employee stock options, however, increased volatility could lead a poorly-diversified employee to exercise his options even earlier, thereby reducing the value of the options.

Difference (5) Taxes: Taxes may have many effects on the values of options. Here we only consider the effect of taxation on the early exercise strategy. Since the compensation or profit from options granted at-the-money is only taxed upon exercise, this taxation will delay exercise in an attempt to postpone the tax. Typically, this delay will cause the option values to increase. Assuming a 25% tax on the exercisable value paid on the exercise date causes the option values in Table VIII to deviate from the values reported there in a range of $-0.44 to $1.50.

Difference (6) Capital Structure Effects: Unlike exchange-traded calls which are typically obligations of parties unassociated with the underlying firm, employee stock options are obligations of the underlying firm itself. As a result, like warrants, they give rise to additional capital for current investment (in lieu of immediate employee compensation), potentially newly issued shares in the future, and the receipt of the striking price upon exercise. To analyze this difference, we need to make some assumption about the effects of the granting and exercise of stock options on the investment activities of the firm. To separate cleanly capital structure from investment issues[21], we will assume that the stochastic process of the portfolio total market value of the firm's stock and stock options is unaffected by the granting or exercise of options, and that it is this value that is the underlying variable in our binomial tree. In particular, this means that the total value of this portfolio V^* on the exercise date of the options will be unaffected by the proportional division of this portfolio between stock and options.[22]

In that case, if the firm has n shares of outstanding common stock and has granted stock options each with striking price K, which if all exercised would give rise to a total of m newly issued shares of stock, the value of an option at exercise would be:

$$(V^* + mK)/(m+n) - K = (V^* - nK)/(n+m)$$

Letting $S^* \equiv V^*/n$ (the value -- inclusive of granted options -- per share) and $\lambda \equiv m/n$ (the dilution factor), then the payoff of a single option can be rewritten as:

$$\max[0, S^* - K]/(1 + \lambda)$$

If we assume that if exercised, all the stock options are exercised at once, then we need only modify the previous analysis by continuing to model the stationary binomial movement of S^* with volatility σ (now interpreted to include any value of the granted options), and to calculate the proceeds at exercise by the above formula instead of the usual $\max[0, S^* - K]$.

One final adjustment reflects the effect of forfeiture on the dilution factor. An approximate way to incorporate this is to use $\lambda(1-p)^\tau$, where p is the expected annualized probability of forfeiture and τ is the years-to-vesting, in place of λ.

---

[21] Effects of stock options, through alterations in the operating characteristics of the firm, are considered separately in our taxonomy by Difference (7).

[22] This is the same approach taken by John Cox and Mark Rubinstein in *Options Markets* (Prentice-Hall, 1985).

Joint Effects of Differences (1)-(6): Our full binomial model of employee stock options requires the following sixteen inputs:

(1) stock price on grant date [$100]
(2) stock volatility [25% - 35%]
(3) stock payout rate [2.5% - 4.5%]
(4) stock expected return [10% - 20%]
(5) interest rate [8%]
(6) option striking price [$100]
(7) option years-to-expiration [10]
(8) option years-to-vesting [3]
(9) expected forfeiture rate [3.5% - 6.5%]
(10) minimum forfeiture rate multiplier [.25 - 1.00]
(11) maximum forfeiture rate multiplier [1 - 4]
(12) employee's non-option wealth per owned option [$30 - $120]
(13) employee's risk aversion [0.5 - 4.0]
(14) employee's tax rate [25%]
(15) percentage dilution [10%]
(16) number of steps in binomial tree [200]

The joint effect of many of these alternative assumptions is examined in the three cases below. In each case, the stock price on the date of grant is $100, the option striking price is $100, the option time-to-expiration is 10 years, time-to-vesting is 3 years, the interest rate is 8%, time to vesting is 3 years, the employee's tax rate is 25%, the percentage dilution (before considering forfeiture prior to vesting) is 10%, and the binomial tree size is 200 steps.

Table IX
Joint Sensitivity of Option Values to Valuation Assumptions
under Grant Date Accounting

	normal case	understated case	overstated case
stock volatility	30%	25%	35%
stock payout rate	3.5%	4.5%	2.5%
stock expected return	15%	10%	20%
expected forfeiture rate	5.0%	6.5%	3.5%
minimum forfeit rate multiplier	.5	1.0	.25
maximum forfeit rate multiplier	2.0	1.0	4.0
employee's non-option wealth	$60	$30	$120
employee's risk aversion	2.0	4.0	.5
option value	$29.10	$18.68	$36.32

Options are at-the-money with time-to-expiration of 10 years, underlying asset price and striking price equal to $100, an interest rate of 8%, and vesting occurs after the end of the third year in the life of the option. The employee's tax rate is 25%, the percentage dilution is 10%, and the binomial tree has 200 steps.

Here the cumulative effects of these different assumptions is to undervalue the option by 36% or to overvalue the option by 25%. In this way, a firm seeking to overvalue its options might report values almost double those reported by an otherwise similar firm seeking to undervalue its options.

III. New Approaches to Option Valuation

It can be argued that the Black-Scholes formula is likely to work best in the market for index options; and yet in recent years the formula has worked very poorly, to the point where most professionals do not really use it.[23] A basic prediction of this formula is that all options on the same underlying asset with the same time-to-expiration (but different striking prices) must have the same implied volatility. While more or less true during the early years of this market and for the early years of the market for equity options, this is far from true today. For example, during early 1990, it was quite common to find six-month index calls which are 9% out-of-the-money with implied volatilities of 13%, while otherwise similar options which are 9% in-the-money have implied volatilities of 23%. This implies that relative to the valuations of Black-Scholes one of these options must have a percentage pricing error of at least 15% or an absolute pricing error of at least $4.00. While the exact implied volatilities are different today, the percentage and dollar errors are no doubt comparable.[24] It may be surmised that the stock market crash of 1987 has permanently changed the way index and equity options are valued so that the Black-Scholes approach is no longer adequate even as a rough approximation.

A generalized binomial model along the lines recently suggest by Bruno Dupire, Emanual Derman and Iraj Kani, or Mark Rubinstein,[25] is likely to become the preferred way used by professionals to handle the above problems. While the Black-Scholes or standard binomial model presupposes that the underlying asset at option expiration has a risk-neutral lognormal distribution (so the only variable really in contention is its volatility), these newer approaches allow the user to input a completely arbitrary terminal distribution (as well as an assumed payout history that can depend on the future stock price and time). This means that corporations using this improved model can not only easily justify its use by pointing to the market failure of the Black-Scholes model, but may easily be able to justify using whatever terminal distribution suits their purposes -- since at the current state of knowledge, this is more an art than a science.

The following table gives an indication of the sensitivity of option values to assumptions about the "shape" of the risk-neutral probability distribution of the underlying asset price

---

[23] Professionals assign a different implied volatility to each option – clearly a kluge to deal with the inadequacies of the Black-Scholes formula since there is no obviously superior candidate to replace it.

[24] These minimum errors from Black-Scholes values have been recently documented in Mark Rubinstein, "Implied Binomial Trees," *Journal of Finance* (July 1994).

[25] See the "The Supermodel Comes of Age," *RISK*, p.6 (January 1994). For specific papers, see Bruno Dupire, "Pricing with a Smile," *RISK* (January 1994), Emanual Derman and Iraj Kani, "Riding on the Smile," *RISK* (February 1994), and Mark Rubinstein, "As Simple as One-Two-Three!" RISK (January 1995).

at the option expiration date. In all these cases, the volatility through the expiration date is fixed at 30%:

Table X
Sensitivity of American Option Values to Shape of Probability Distribution

skewness	kurtosis	option value
.00	2.99	$37.82
-.95	3.93	$34.51
+.91	4.00	$45.51

Options are American and at-the-money with time-to-expiration of 10 years, underlying stock price and striking price equal to $100, annualized stock volatility of 30%, annualized dividend yield of 3.5% and interest rate of 8%. The calculations are based on generalized binomial trees with 200 steps.

IV. Minimum Value as an Alternative

Even with these large potential percentage differences in option values, it may still be argued that some positive valuation is better than zero, which is the current practice. But this is not obvious. For example, consider the case of two otherwise identical firms with options that should properly be valued at $29.10, but where one firm deliberately undervalues them at $18.68 and the other deliberately overvalues them at $36.32 (see Table IX). Before implementing FASB's proposal, both firms would have reported the same profits since the options would have been valued at zero. After implementing the proposal, they would report different profits and their accounting statements would no longer be comparable. Of course, it can be argued that since both $18.68 and $36.32 are closer to the correct $29.10 than they are to zero, in an absolute sense both firms are now, after FASB's implementation, reporting profits closer to their true amounts. However, where before the firms had comparable accounting statements, now they do not. It is not clear the former benefit outweighs the latter drawback.

This line of reasoning seems to suggest that if comparability can be maintained and firms could report a value for their employee stock options that, while incorrect, at least brings their reported profitability closer to its true amount, then such a procedure should be adopted. It might seem that the alternative minimum option value technique discussed in the Exposure Draft might have these advantages. Minimum value accounting would require firms to value options on the date of grant at current stock price (adjusted downward for expected dividends) minus the present value of the striking price, provided this were greater than zero, or zero otherwise. To remove any chance for non-comparability to arise from misestimation of the expected life of the options, one could simply value the options as if they would be exercised at the first available opportunity (the vesting date). In addition, to account for forfeiture, one would multiply this value by one minus the probability of forfeiture. To see this concretely, suppose that **p** is the annualized probability of forfeiture, **S** the current value of the underlying stock, δ the annualized one plus dividend yield, **K** the striking price, **r** the annualized interest return, and τ the time-to-vesting, then the value of an option would be:

$$(1\text{-}p)^{\tau} \times \max[0, S\delta^{-\tau} - Kr^{-\tau}]$$

Provided comparable firms estimated p and δ the same, both firms would value their options the same. Not only does this value place a lower bound on the value of the options,[26] it also is quite easy to implement. For example, in the benchmark situation described above where $p = .05$, $\tau = 3$, $S = K = 100$, $\delta = 1.035$ and $r = 1.08$, the option would be worth $9.27. This, of course, is much lower than the true value of $29.10, but at least it moves the financial statements in the right direction, that is, away from zero.

Unfortunately, even this approach has potentially serious problems for three reasons.

First, the $9.27 value is much lower than the true value of $29.10. So the intent of FASB's proposal would only be very partially realized.

Second, there still remains room for significant non-comparability as the table below indicates:

Table XI
Joint Sensitivity of Minimum Values to Valuation Assumptions

	normal case	understated case	overstated case
payout rate	3.5%	4.5%	2.5%
forfeiture rate	5.0%	6.5%	3.5%
option value	$9.27	$6.74	$12.11

Options are at-the-money with time-to-vesting of 3 years, underlying asset price and striking price are equal to $100 with an interest rate of 8%.

However, here the likely effects of non-comparability may be overstated since over the shorter 3 year rather than 10 year period required for the minimum value calculation, it will be more difficult for comparable firms to justify such large differences in assumed payout and forfeiture rates.

Third, and by far the most important, firms can easily circumvent the intention of the minimum value approach by changing the terms of their options. While this could be accomplished in a number of ways, here is a particularly elegant method: change the striking price so that it is increased by the ratio of the interest return divided by the payout return through the vesting date. In the example above, the striking price $K = 100$ would instead be replaced by $K(r/\delta)^{\tau} = 100(1.08/1.035)^3 = \$113.61$. In this case the minimum

[26] This value is equivalent to the Black-Scholes value obtained with a time-to-expiration of 3 years and a volatility of 0%.

option value would be 0. Since these options would be granted with a higher striking price, employees would, of course, receive less value per option; nonetheless the total size of the compensation package could be maintained by granting more options. Not only would employee incentives be maintained (indeed, possibly enhanced), but stockholders might well agree that this was a superior compensation plan since employees would only be rewarded by the options if the stock price plus dividends were to grow faster than the interest rate -- an alternative easily available to the stockholders without investing in the stock. In the end, what would FASB have achieved by adopting the minimum value approach? Many firms would continue to report their employee stock options at zero value, but would have perhaps altered their plans solely for accounting, not economic, purposes.

V. Effects on Earnings Per Share

For the purpose of calculating primary earnings per share, the number of shares is set equal to the actual number of outstanding shares plus the number of additional shares that would need to be issued with just sufficient proceeds to buy back outstanding options at their currently <u>exercisable values</u>. For this calculation, the Exposure Draft would have firms only consider the number options that are expected to survive the vesting period.[27] A problem with this approach is that currently out-of-the-money options create no reported dilution even though they can be expected to create at least some dilution in the future (since there is a positive probability the options may end up in-the-money and be exercised). The correct way to handle this, given that a reliable method can be found to value the options, is to add to the number of outstanding shares, the number of additional shares that would need to be issued to buy back outstanding options at their current <u>values</u>. In particular, since out-of-the-money options have positive values, to that extent they would increase the number of assumed shares for the purposes of calculating EPS.

VI. Vesting Date Accounting

In response to the storm of protest over its Exposure Draft, FASB is considering measuring the option expense on the vesting date rather than on the grant date.[28] In particular, the stock option is valued as a standard call using the stock price on the vesting date, a time-to-expiration equal to the expected time to expiration or exercise remaining after the vesting date, and the actual number of options vested. This revision has three estimation advantages:

(1) because maturity is nearer on the vesting date, the problems of

---

[27] In contrast, fully diluted earnings per share would count all outstanding options, whether or not they are likely to vest.

[28] A member of FASB's staff stated that FASB was considering the vesting date as an alternative, not because it was conceptually superior to grant date, although a reasonable conceptual case can be made for it, but primarily because it solves several problems related to grant date accounting.

estimating dividends and volatility are reduced (mitigation of Difference (1));

(2) the option model need not be revised for delayed vesting (elimination of Difference (2));

(3) there is no need to adjust the value of the options downward to account for the probability of forfeiture through the vesting date (mitigation of Difference (3));

In addition, realizing the inconsistency of its previously proposed ex-post adjustment for the realized life of the option, FASB seems to be dropping this adjustment.

Two conceptual arguments help justify vesting date accounting: the view that contingent contracts are not liabilities and that the proper measure of the actual service rendered is the increase in the stock price between the grant and the vesting date. It can be argued that as long as there is the precondition of continued employment before exercise is possible and as long as the employee has not agreed to anything, an employee does not really have an option.

To get an idea of how much switching from grant to vesting date accounting will reduce opportunities for non-comparable financial statements, reconsider the comparison made in Table IX. Suppose that the stock price and option striking price are both $100, the interest rate is 8%, time-to-expiration is 7 years, the option vests immediately, the tax rate is 25%, percentage dilution is 10% × $(1-.05)^3$, and the binomial tree size is 200 steps.

Table XII
Joint Sensitivity of Option Values to Valuation Assumptions under Vesting Date Accounting

	normal case	understated case	overstated case
stock volatility	30%	27%	33%
stock payout rate	3.5%	4.0%	3.0%
stock expected return	15%	10%	20%
expected forfeiture rate	5.0%	6.5%	3.5%
minimum forfeit rate multiplier	.5	1.0	.25
maximum forfeit rate multiplier	2.0	1.0	4.0
employee's non-option wealth	$60	$30	$120
employee's risk aversion	2.0	4.0	.5
option value	$27.86	$17.49	$31.61

Options are at-the-money with time-to-expiration of 7 years, underlying asset price and striking price equal to $100, an interest rate of 8%, and vesting is immediate. The employee's tax rate is 25%, the percentage dilution is 10%×$(1-.05)^3$, and the binomial tree has 200 steps.

This situation is only somewhat improved over grant date accounting. This could have been anticipated since, as Table V indicated, eliminating delayed vesting from the calculation should have had little impact on values.

In the spirit of vesting date accounting, FASB has more recently considered a relaxed minimum value approach under which an option is expensed at approximately its minimum value calculated based on the stock price measured on the vesting date. The specific proposal is to expense the option valued as if it were a standard call on the vesting date with a 90-day maturity. Again this approach is likely to engender another slightly more sophisticated, but almost as efficacious, form of gaming. As before, as a response to minimum value at the grant date, firms might grant options out-of-the-money, but set a floating vesting date such that the option automatically vests on the day the stock price first hits the striking price.[29] Conceivably, this might actually improve the incentive effects of stock options while at the same time leading to a very small accounting cost of compensation.

YES!

VII. Exercise Date Accounting

The Exposure Draft advocates expensing options based on their grant date values with ex-post truing-up for the realized forfeiture rate during the vesting period and the realized life of the options. Note that errors in other model inputs such as volatility and dividends are not to be trued-up. This means that the cumulative balance sheet retained earnings figure will never be corrected over the life of the corporation. This may be unlike any other form of accounting treatment. Accounting depreciation, for example, while it may be very different than actual market value depreciation during the life of plant and equipment, will nonetheless be trued-up to actual market value transactions when the plant and equipment is finally sold or decommissioned.

As an alternative, FASB gives brief consideration to expensing options based on their realized payoffs at exercise or expiration. This is known as "exercise date accounting". Under this approach, options would still be expensed when granted based on some pricing model, but as their expiration date approached this estimate would be periodically retroactively adjusted for the changed value of the options. A final model-free adjustment would be made upon exercise, setting the option value equal to its ex-post realized exercise date payoff or upon expiration setting its value to zero. This extreme form of truing-up to actual transactions minimizes the damage created by inaccurate valuation during the life of the options, since eventually model- and estimate-free truth will out. Errors in volatility, payout, and forfeiture rate estimates, incorrect modeling of the employee's exercise strategy, and use of an incorrect option pricing formula or algorithm, are all eventually corrected under exercise date accounting. Not only does this reduce the incentives for firms to misvalue their stock options to manage earnings or to game the accounting rules by revising the terms of their options, but it also substantially reduces

---

[29] This is an example of an up-and-in barrier call where the barrier equals the striking price. Black-Scholes type formulas for barrier options can be found in Mark Rubinstein and Eric Reiner, "Breaking Down the Barriers," RISK (September 1991). However, in this case, ignoring possible forfeiture, with the barrier equal to the striking price, the option is equivalent to a simple out-of-the-money call.

the informational damage to the market from doing so or even from unintentional errors.

So why does FASB balk at requiring exercise date accounting? Unfortunately, it would force it to reconsider some fundamental issues in accounting, notably, the very definitions of liabilities and equities. Exercise date accounting effectively treats employee stock options, not as equity, but as a liability of the firm. To be consistent, other securities such as warrants would also need be reclassified as liabilities. But, given the proliferation of corporate securities, like convertible bonds that have some equity and some liability features under current definitions, it may be time to do so. Let me suggest that employee stock options, warrants, preferred stock, etc. be lumped together as a third as yet unnamed class of securities, and reserve the term "equity" to refer only to the last residual claim on assets -- common stock. From the perspective of preexisting common stock holders, these securities are clearly not equity, and just as the stock holders are interested ultimately in the realized return, rather than the expected return, of an investment, so too they are interested ultimately in the realized exercise date cost of an employee stock option, rather than its expected cost as estimated on the grant date.

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" --by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

208. Steven E. Plaut. "The Prepayment Uncertainty of Collateralized Mortgage Obligations." March 1991.

209. Hua He. "Optimal Consumption-Portfolio Policies: A Convergence from Discrete to Continuous Time Models." April 1991.

210. Gerard Gennotte and Terry A. Marsh. "Variations in Economic Uncertainty and Risk Premiums on Capital Assets." May 1991.

211. Gerard Gennotte and Hayne Leland. "Low Margins, Derivative Securities, and Volatility." May 1991.

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993..

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program
in Finance
WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 242

Explaining Forward Exchange Bias. . .Intraday

by

Richard K. Lyons

and

Andrew K. Rose

January 1995

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Explaining Forward Exchange Bias. . .Intraday

Richard K. Lyons and Andrew K. Rose

Haas School of Business
University of California, Berkeley
and
NBER

January 1995

Finance Working Paper #242

We thank the following for helpful comments: an anonymous referee, Richard Meese, and workshop participants at LSE and the University of Washington. We also thank the San Francisco Federal Reserve Bank for hospitality and assistance with data. Lyons gratefully acknowledges financial assistance from the National Science Foundation and the Berkeley Program in Finance.

Abstract

Intraday interest rates are zero. Consequently, a foreign exchange dealer can short a vulnerable currency in the morning, close this position in the afternoon, and never face an interest cost. This tactic might seem especially attractive in times of crisis, since it suggests an immunity to the central bank's interest rate defense. In equilibrium, however, buyers of the vulnerable currency must be compensated on average with an intraday capital gain as long as no devaluation occurs. That is, currencies under attack should typically *appreciate* intraday. Using data on intraday exchange rate changes within the EMS, we find this prediction is borne out.

Richard K. Lyons and Andrew K. Rose
Haas School of Business
University of California, Berkeley
Berkeley, CA 94720-1900
and NBER

Explaining Forward Exchange Bias . . . Intraday

This paper examines implications of the fact that interest rates are zero intraday.[1] In particular, we focus on the foreign exchange (FX) market, and ask whether trading strategies might be affected. The answer to this question is of greater import than might first appear. For example, in times of crisis central banks typically employ an interest rate defense, raising domestic rates to attract a capital inflow and punish short-sellers. But, if dealers are immune to this defense — at least on an intraday basis — then perhaps the viability of fixed rate regimes is undermined. (Goldstein et al (1993) provide an overview of how central banks defended their currencies during the 1992 currency crisis.)

A simple example helps. With intraday interest rates of zero, a dealer can short a high interest rate currency in the morning, close her position in the afternoon, and never face an interest cost. If there is any likelihood of an intraday devaluation, this appears to be an attractive strategy, other things equal, since the dealer is immune to the interest cost of an overnight short position.

Other things should not be equal in equilibrium, however. Buyers of the vulnerable currency must be compensated on average with an intraday capital gain, as long as no devaluation occurs. That is, devaluation risk is offset by systematic *appreciation*. Further, the greater the probability and size of the devaluation, the greater the implied appreciation. Thus, the absence of a role for the interest differential in equating expected returns across currencies implies that the exchange rate itself takes up the slack.

In a regression of intraday exchange rate changes on interest differentials we find this prediction is borne out: the higher the weak currency's interest rate, the

---

[1] For details regarding settlement, see Stigum (1990), particularly pages 893–901.

more that currency *appreciates* intraday. The same finding elsewhere in the literature is referred to as "forward rate bias". Though the longer–horizon findings — that high interest rate currencies tend to appreciate — remain unexplained, our's does not: intraday, the expected cost of shorting a currency in crisis offsets the expected gains from devaluation.

The paper is organized as follows: Section 2 presents a model of intraday trading in times of crisis; Section 3 describes the data; Section 4 presents our results; and Section 5 concludes.

I. A Model of Intraday Trading

Consider a single asset that is tradable in a single market at any time over a span divided into n periods, each of length T. In order to abstract from portfolio balance issues, we assume the asset is in zero net supply (we discuss risk premia in our comments on intervention below). Let S_t denote the price of the asset at time t. For concreteness, we associate S_t with the nominal exchange rate in French Francs per Deutschemark, or FF/DM. Further, let R_t^{FF} and R_t^{DM} denote the per–period nominal interest rates, in FF and DM respectively, applying to open positions. Our core assumptions are the following:

(A1) Settlement–FX: all FX trades effected within a period are settled at period close.

(A2) Settlement–Interest: open positions in FX involve interest on a per–period basis, but only if open positions are carried across period close. If carried across a period close and offset in the subsequent period, open positions accrue a full period of interest, regardless of how far into the subsequent period the position is maintained.

(A3) Uncovered interest parity (UIP) holds.

Assumption (A1) is realistic since spot FX is traded over periods within which settlement time is unchanging (in reality, settlement typically occurs two days forward rather than at the day's "close"). Assumption (A2) captures the fact that daily interest is a discrete variable: if one opens a position and closes it five minutes later, but settlement of the second trade is one day later than that of the first trade, then one full day's interest will accrue. Assumption (A3) — though rejected empirically over monthly and quarterly horizons — allows us to focus attention on the expected return consequences of intraday trading. To our knowledge, UIP has not been tested at this horizon (Hodrick (1987)). Henceforth, we work with with a log–linear approximation of UIP (the negligible size of intraday cross terms is demonstrated below).[2]

The above assumptions imply that:

$$\mathrm{E}[s_{t+\tau}-s_t|\Omega_t] = \mathrm{D}_{t+\tau}(R_t^{FF} - R_t^{DM}), \qquad \tau < T. \tag{1}$$

Here, $s_t=\log(S_t)$ and Ω_t denotes the representative agent's information set at time t. $\mathrm{D}_{t+\tau}$ is an indicator variable equal to 1 if $t+\tau$ is in the period subsequent to that containing t, and equal to 0 if $t+\tau$ is in the same period as containing t. Thus, when $\mathrm{D}_{t+\tau}=0$, the expected change in the log of the exchange rate must also be zero.

The expected dynamics implied by equation (1) are presented in Figure 1. Implicit in the figure is the assumption that R_t^{FF} and R_t^{DM} are constant, with $R^{FF}>R^{DM}$. The most distinctive feature is that this model generates expected discontinuities in the exchange rate at the settlement points.[3]

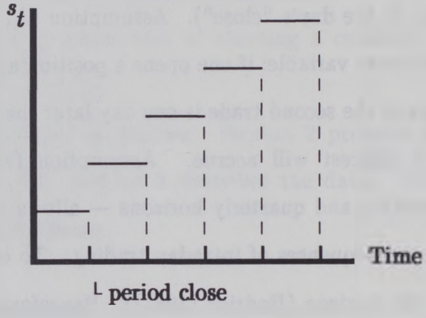
---

[2] We note that any terms arising from Jensen's inequality are absorbed into the constant of our estimating equation as long as second moments are time–invariant.

[3] In reality, spot FX is settled the second business day after the transaction, so there is a distinction between the time the settlement date (value date) advances one day, and the time payments are actually made on the day of settlement. For our purposes, what matters is the

Figure 1

Expected Exchange Rate Behavior Over Time



Times of Crisis

We turn to implications of the model in times of crisis. Times of crisis are interesting because interest differentials become first–order relevant even at horizons of one day. For example, during March, 1983 the value of $R^{FF}-R^{DM}$ topped 80% on an annual basis (30–day eurorates). It is within this extreme context that policy–makers must evaluate the effectiveness of the interest rate defense.

In order to gauge the size of interest rate differentials on a daily basis, Table I presents some statistics. The numbers in the columns on the right represent the size of the periodic exchange rate discontinuities illustrated in Figure 1. (Note that the columns are the same up to the precision reported. Hence, the cross terms that distinguish the linear version of uncovered interest parity from the exact version are quite small at this horizon.)

---

time the value date advances one day. For the currencies we consider below, the worldwide standard for advancing the value date has varied between 9 PM and 10:30 PM London time (GMT) over the EMS period (sources: bank dealers and Reuters).

Table I

Annual Interest Differentials on a Daily Basis*

Quoted $R^{FF}-R^{DM}$ Annual Basis	$(R^{FF}-R^{DM})/360$ Daily Basis Points	$(1+R^{FF}/360)/(1+R^{DM}/360) - 1$ Daily Basis Points
10%	2.8	2.8
15%	4.2	4.2
20%	5.6	5.6

* Daily–basis values are expressed in 0.01%, or basis points. Eurorates (other than Sterling) are quoted on a 360–day basis so that gross yield over t days equals $1+R(t/360)$ where R is the quoted rate [see Stigum (1981), pages 175–178]. The DM rate used is the median quote in our sample.

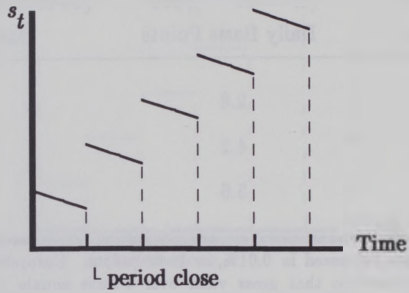
The question we want to answer is this: does the lack of intraday interest rates provide agents with a costless means of speculating against vulnerable currencies? Our analysis follows directly from equation (1), as before, except that now we must determine the implications of our assumptions under a positive probability of devaluation. Clearly, the *total* expected change in the exchange rate must still be zero. Accordingly, for intra–period open positions we can write:

$$\mathrm{E}[s_{t+\tau}-s_t|\Omega_t] = p\mathrm{E}[s_{t+\tau}-s_t|\text{deval.}] + (1-p)\mathrm{E}[s_{t+\tau}-s_t|\text{no deval.}] = 0 \qquad (2)$$

where p denotes the exogenous probability that a devaluation will occur between t and $t+\tau$. With $\mathrm{E}[s_{t+\tau}-s_t|\text{deval.}]>0$, this implies that $\mathrm{E}[s_{t+\tau}-s_t|\text{no deval.}]<0$. That is, *conditional on no devaluation, the weak currency should appreciate* on average within the period. Figure 2 provides a qualitative illustration:

<u>Figure 2</u>

Ex-Post Exchange Rate Behavior Conditional on No Devaluation



A testable implication of our model as applied to crises is presented in the following proposition:

Proposition 1: Intraday, if a higher weak–currency interest rate reflects greater expected devaluation then — conditional on no devaluation — a higher weak–currency interest rate implies greater expected appreciation, ceteris paribus.

Proof: We know from equation (2) that intra–period $E[s_{t+\tau}-s_t|\Omega_t] = pE[s_{t+\tau}-s_t|\text{deval.}] + (1-p)E[s_{t+\tau}-s_t|\text{no deval.}] = 0$. But, if an increase in $(R_t^{FF}-R_t^{DM})$ => an increase in $pE[s_{t+\tau}-s_t|\text{deval.}]$, then $pE[s_{t+\tau}-s_t|\text{no deval.}]$ must be lower.

This is the implication we test in the data. That is, we estimate the following regression:

$$\Delta s_{t+\tau} = \beta_0 + \beta_1(R_t^{FF}-R_t^{DM}) + \epsilon_{t+\tau} \tag{3}$$

where: $\Delta s_{t+\tau}$ in the intra–day change in the log of the spot rate; $R_t^{FF}-R_t^{DM}$ is the interest differential (daily basis); and $\epsilon_{t+\tau}$ is a stationary expectational error. (Since $\epsilon_{t+\tau}$ represents news it is orthogonal to available information such as interest rates; hence, least squares is a consistent estimator for (3).) Proposition 1 implies that if a higher interest differential reflects higher expected devaluation, then β_1 should be negative *so long as there are no intraday devaluations in the sample.* (In the sample we consider, none occurred. That said, it is important that devaluation *can* occur intraday. Sweden provides an example: the November 19, 1992 devaluation occurred during business hours. Further, the devaluation was news: the Prime Minister was apprised just ten minutes before flotation (see the *Financial Times*, 11/20/92)).[4]

Note that under covered interest parity our regression is exactly the canonical regression of $\Delta s_{t+\tau}$ on the forward discount. The estimated coefficients in the literature are consistently negative for intermediate horizons, in violation of uncovered interest parity (see Hodrick (1987) for a comprehensive discussion of forward discount bias). In contrast, our model, derived from UIP, *predicts* a negative coefficient — for intraday horizons.

II. Data and Related Issues

Our empirical implementation uses two currencies within the Exchange Rate Mechanism (ERM) of the European Monetary System (EMS): the French Franc and the Italian Lira, both relative to the German DM, anchor of the ERM. A number of factors are relevant for our choice of data. First, we need fixed exchange rates to get the devaluation possibility that drives the model. Second, we need high expected devaluation — proxied by high interest differentials — otherwise the

[4] Note that if a large sample were available — i.e., one that includes a representative number of intraday devaluations — one would expect a value of zero for β_1 under UIP.

implied intraday drift is too small to detect. Third, the ERM dominates both Bretton Woods and developing–country possibilities: more crises were defended with high interest rates than under Bretton Woods, and institutional issues are not the problem they would be in the developing–country context (e.g., capital controls, thin markets, etc.). Finally, within the ERM, the French Franc and Italian Lira account for the lion's share of high interest differential observations. Indeed, there are still relatively few attacks of the magnitude we require; hence, we pool our data across countries. (See also Svensson (1993) for further evidence regarding the intrinsic appeal of the ERM as a target of analysis.)

Our sample runs from 3/13/79 to 10/26/92, which includes a total of 3555 weekdays. We construct FF/DM and IL/DM rates (IL denotes Italian Lira) using dollar quotes, i.e., the FF/DM rate equals (FF/$)($/DM). Our end–of–period rates are the daily London close quotes (midpoints) from the *Financial Times*, which over this period were recorded at 5 PM London time. Our beginning–of–period rates are European Currency Unit (ECU) fix rates recorded at 2:15 Swiss time (1:15 London time) by the Bank of International Settlements (BIS). There is no spread for the ECU fix series, since fixings are auctions. Finally, these fix series are the earliest consistent series available for London trading hours, to our knowledge.

Our interest rate data for the FF, IL, and DM are the 30–day euro–currency rates recorded at 10AM Swiss time (9 AM London time) by the BIS. As euro–rates, they are virtually free of political risk. The 30–day market is deeply traded; we also use 2–day rates for a robustness check.[5]

We need to determine a definition of a crisis in terms of interest differentials since our model's non–zero drift prediction is only relevant during times of crisis.

[5] Note that the interest cost of an *overnight* short position is tied to a *forward* interest rate, from t+2 days to t+3 days, since spot deliveries are typically two days forward. This has no bearing, however, on the fact that the interest cost of an intraday short position is zero.

The larger the cutoff interest differential, the larger the implied drift, but the cost is lower statistical power since the available sample shrinks rapidly. Our preferred cutoff is a ten percent interest differential (annual basis), $R^{FF}-R^{DM}$ or $R^{IL}-R^{DM}$, although we present results for different thresholds. This preference is based on three factors. First, a ten percent differential is large enough to be a strong signal of crisis. Second, on a daily basis, a ten percent differential is large enough to imply a drift that is not dominated by typical spreads (Lyons (1993b) finds a 2 basis point median spread in DM/$ transactions data; note that Reuters' indicative quotes overstate inter–dealer spreads by a factor of 2 or 3). Third, ten percent is not so large as to limit severely our sample size.

Parenthetically, though intervention often takes place during crises, this does not vitiate our results. Unsterilized intervention — the more important for the FX market — has effects that are captured by the interest rates in our model. One could argue that *intraday* unsterilized intervention is not reflected in the morning interest rates, and creates bias in our regression since it systematically goes in the support direction. This argument is flawed, however: it neglects the fact that only innovations in intervention should impact the exchange rate; what matters is departures from expected intervention, not just the direction. In addition, we view the case for sizeable portfolio–balance effects from sterilized intervention as weak, especially given the point about innovations above (see Edison (1992)). Irrespective of these arguments, though, if the data generate a significant negative β_1 in equation (3) then there is a cost to shorting vulnerable currencies intraday, whether the source is intervention or not. Of course, if central banks are the only buyers earlier in the day, then perhaps they do not require the expected appreciation that maintains UIP. This possibility makes a finding of a significant negative β_1 all the more striking.

III. Estimation Results

Table II presents our OLS results. To get a sense of the sensitivity of our sample size and results to the interest differential, we provide estimates for three different cutoffs: 10%, 15%, and 20% on an annual basis. To provide more interpretable coefficients, we translate the annualized interest differentials to a daily basis [using the Table 1 formula $(R_t - R_t^{DM})/360$, where R_t and R_t^{DM} are annual basis quotes, and R_t denotes either R_t^{FF} or R_t^{IL} as appropriate.].

Table II

The Intraday Returns Relationship*

$$\Delta s_{t+\tau} = \beta_0 + \beta_1(R_t - R_t^{DM}) + \epsilon_{t+\tau}$$

	$\hat{\beta}_0$	$\hat{\beta}_1$	OBS
30–Day Interest Diff. (annualized)			
$\geq$10%	0.0002 (1.05)	–0.90 (–2.36)	842
$\geq$15%	0.0007 (1.91)	–1.48 (–2.71)	261
$\geq$20%	0.0009 (1.43)	–1.74 (–2.30)	105

* $\Delta s_{t+\tau}$ is the change in the log of the exchange rate over the intraday holding period, in FF/DM or IL/DM as appropriate. $R_t - R_t^{DM}$ is the nominal interest differential, daily basis, where R_t denotes either R_t^{FF} or R_t^{IL} as appropriate. OBS denotes number of observations meeting the interest differential cutoff criterion. The criterion $\geq$10% denotes observations for which the own–currency interest rate is at least 10% higher than the DM interest rate on an annual basis. Similarly for the other criteria. Estimated using OLS. Sample: 3/13/79 to 10/26/92. T–statistics in parentheses.

The results are clear: the greater the interest differential, the more the vulnerable currency appreciates intraday.[6] The implications of our model are apparently borne out in the data.

We can go further and interpret the β_1 magnitudes, but this introduces the knotty problem of translating trading hours into trading days. With some simplifying assumptions, it is easy to show that UIP predicts $\beta_1 = -1$.[7] Again, the prediction works well: while $\beta_1 = 0$ can be rejected at conventional levels of statistical significance, the hypothesis that $\beta_1 = -1$ cannot.

IV. Conclusions

Our first result derives from analysis of our model: intraday interest rates of zero do not imply that agents have a costless means of speculating against vulnerable currencies within the day. On the contrary, if the interest differential cannot do its work then exchange rate dynamics have to take up the slack. Further, if expected returns are to be equated, then the larger the expected devaluation, the more the vulnerable currency is expected to *appreciate* within any day in which a devaluation does not occur.

Our second result is empirical: our analytical results are borne out in the data.

---

[6] We conduct three types of sensitivity analysis: (1) we use 2–day interest rates instead of 30–day rates, (2) we split the data by country, and (3) we bootstrap the standard errors. The 2–day interest rates produce a negative and highly significant β_1. The country results are weaker for Italy: though France alone still generates a significantly negative β_1, Italy does not. Finally, bootstrapped standard errors are roughly twice as large as conventionally–calculated standard errors, but are conditional on independence of the residuals over time, a strong assumption in this context. The reported t–statistics use conventionally–calculated standard errors.

[7] The assumptions are: (i) per proposition 1, the daily–basis $R_t - R_t^{DM} = p\mathrm{E}[s_{t+\tau} - s_t | \text{deval.}]$ where τ is the length of a trading day, (ii) our empirical measure of $s_{t+\tau} - s_t$ corresponds to one trading day, and (iii) p is small, so that $(1-p)$ is close to 1. To see that UIP predicts $\beta_1 = -1$, note that equation (2) implies $\mathrm{E}[s_{t+\tau} - s_t | \text{no deval.}] = -(1-p)^{-1}(p\mathrm{E}[s_{t+\tau} - s_t | \text{deval.}]) \approx -(R_t - R_t^{DM})$.

The larger the expected devaluation — proxied by the interest differential— the more the *vulnerable* currency *appreciates* intraday. Hence, dealers are not immune to the central bank's interest rate defense within the day. That said, the implied intraday drifts are not large. This kind of an effect is irrelevant for all but the lowest transaction–cost participants at times of substantial devaluation risk.

References

Edison, Hali, 1992, The effectiveness of central bank intervention: A survey of the post–1982 literature," Working paper, Federal Reserve Board of Governors, Washington, D.C.

Goldstein, Morris, David Folkerts–Landau, Peter Garber, Liliana Rojas–Suarez, and Michael Spencer, 1993, *International Capital Markets, Parts I and II*, The International Monetary Fund, Washington D.C.

Hodrick, Robert J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets* (Harwood, London).

Lyons, Richard K., 1993, Tests of microstructural hypotheses in the foreign exchange market, NBER Working Paper #4471, forthcoming in the *Journal of Financial Economics.*

Stigum, Marcia, 1981, *Money Market Calculations: Yields, Break–Evens, and Arbitrage* (Dow Jones–Irwin, Homewood, IL).

Stigum, Marcia, 1990, *The Money Market* (Dow Jones–Irwin, Homewood, IL).

Svensson, Lars E.O., 1993, Assessing target zone credibility: Mean reversion and devaluation expectations in the ERM, 1979–1992, *European Economic Review,* 37, 763–793.

Edison, Hali J., 1992, The effectiveness of central bank intervention: A survey of the post-1982 literature, Working paper, Federal Reserve Board of Governors, Washington, D.C.

Goldstein, Morris, David Folkerts-Landau, Peter Garber, Liliana Rojas-Suarez, and Michael Spencer, 1993, *International Capital Markets, Parts I and II*, The International Monetary Fund, Washington D.C.

Hodrick, Robert J., 1987, *The Empirical Evidence on the Efficiency of Forward and Futures Foreign Exchange Markets* (Harwood, London).

Lyons, Richard K., 1993, Tests of microstructural hypotheses in the foreign exchange market, NBER Working Paper #4471, forthcoming in the *Journal of Financial Economics*.

Stigum, Marcia, 1981, *Money Market Calculations: Yields, Break-Evens, and Arbitrage* (Dow Jones-Irwin, Homewood, Il).

Stigum, Marcia, 1990, *The Money Market* (Dow Jones-Irwin, Homewood, Il).

Svensson, Lars E.O., 1993, Assessing target zone credibility: Mean reversion and devaluation expectations in the ERM, 1979-1992, *European Economic Review* 37, 763-793.

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" --by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

210. Gerard Gennotte and Terry A. Marsh. "Variations in Economic Uncertainty and Risk Premiums on Capital Assets." May 1991.

211. Gerard Gennotte and Hayne Leland. "Low Margins, Derivative Securities, and Volatility." May 1991.

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance

WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 243

Foreign Exchange Volume: Sound and Fury Signifying Nothing?

by

Richard K. Lyons

January 1995

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Foreign Exchange Volume:
Sound and Fury Signifying Nothing?

Richard K. Lyons

Haas School of Business
University of California, Berkeley
and
NBER

January 1995

Finance Working Paper #243

I thank the following for helpful comments: George Constantinides, Mark Flood, Jeff Frankel, Antonio Melo, Carol Osler, and seminar participants at Berkeley, LSE, Northwestern, NYU, UBC, the IMF, MIT, and *The Microstructure of Foreign Exchange Markets* conference. I also thank Jeff Bohn for valuable research assistance, and Merrill Lynch and Lasser Marshall for access to dealers/brokers while trading. Financial assistance from the National Science Foundation and the Berkeley Program in Finance is gratefully acknowledged. Any errors are my own.

Abstract

This paper examines whether currency trading volume is informative, and under what circumstances. Specifically, we use transactions data to test whether trades occurring when trading intensity is high are more informative -- dollar for dollar -- than trades occurring when intensity is low. Theory admits both possibilities, depending primarily on the posited information structure. We present what we call a hot-potato model of currency trading, which explains why low-intensity trades might be more informative. In the model, the wave of inventory-management trading among dealers following innovations in order flow generates an inverse relationship between intensity and information content. Empirically, low-intensity trades are more informative, supporting the hot-potato hypothesis.

Richard K. Lyons
Haas School of Business
University of California, Berkeley,
Berkeley, CA 94720-1900
and NBER

Foreign Exchange Volume: Sound And Fury Signifying Nothing?

1. Introduction

Volume in the spot foreign exchange (FX) market dwarfs that in any other financial market. But is all this trading informative? This paper provides some empirical evidence. At the broadest level, our results help to clarify why trading volume in this market is extraordinarily high. At a narrower level, we provide some sharp results regarding the relationship between the intensity of trading and the informativeness of trades.

Specifically, we provide results that discriminate between polar views of trading intensity, which we refer to as (1) the event–uncertainty view and (2) the hot–potato view. The event–uncertainty view holds that trades are more informative when trading intensity is high; the hot–potato view holds that trades are more informative when trading intensity is low. In general, theory admits both possibilities, depending primarily on the posited information structure.

To understand the event–uncertainty view, consider the work of Easley and O'Hara (1992). In contrast to earlier models where new information is known to exist, in Easley and O'Hara (1992) new information may not exist. That is, there is some probability, say p, of new information, and probability (1–p) of no new information. In the event of new information, there is some probability, say q, that an informed trader has received good news, and probability (1–q) of having received bad news. They demonstrate that if there is no trade at time t then a rational dealer raises the probability she attaches to the no–information event, and lowers the probability of news having occurred. Put differently, if trading intensity is low, an incoming trade of a given size induces a smaller update in beliefs since it is less likely to be signaling news. On the flipside, trades occurring when intensity is high should induce a larger update in beliefs.

To understand our term "the hot–potato view" — that trades are more informative when trading intensity is low — consider the ideas of Admati and Pfleiderer (1988). Key to their model is the presence of discretionary liquidity traders: in order to minimize the losses they suffer to informed traders, rational liquidity traders clump together in their trading. (The reason informed traders cannot fully offset this advantage to clumping is that information is short–lived.) Due to this clumping of liquidity traders, trades occurring when intensity is high tend to be less informative.

The metaphor of the hot–potato offers a link between discretionary liquidity trading and FX trading. FX dealers use the metaphor in referring to the repeated passage of idiosyncratic inventory imbalances from dealer to dealer following an innovation in customer order flow. These inter–dealer liquidity trades are clearly discretionary as to timing, hence the connection between discretionary liquidity trading and the hot–potato view of order–flow information. To clarify the hot–potato process, consider the following crude, but not unrealistic example. (Keep in mind that roughly 90% of FX trading is inter–dealer, a much higher share than in other multiple–dealer markets.) Suppose there are 10 dealers, all of whom are risk averse, and each currently with a zero net position. A customer sale of $10 million worth of DM is accommodated by one of the dealers. Not wanting to carry the open position, the dealer calculates his share of this inventory imbalance — or 1/10th of $10 million — calls another dealer, and unloads $9 million worth of DM. The dealer receiving this trade then calculates his share of this inventory imbalance — or 1/10th of $9 million — calls another dealer, and unloads $8.1 million worth of DM. The hot–potato process continues. In the limit, the total inter–dealer volume generated from the $10 million customer trade is: $9m/(1–0.9)= $90 million. The resulting share of wholesale trading that is inter–dealer: 90%, roughly matching the empirical share.

Here are two possible reactions to the example above, neither of which vitiates its message. Reaction one: shouldn't the multiplier be infinite since risk–averse dealers would not choose to retain any of the imbalance? The answer to this query is this: in equilibrium, price will adjust to induce dealers to hold some of the perceived excess supply. [The 10% rule of the example is a crude approximation of a much richer short–run clearing mechanism; see Lyons (1994) for an optimizing model in which hot–potato trading arises between dealers.] Reaction two: inter–dealer trades can reduce idiosyncratic inventory imbalances — which reduces idiosyncratic risk rather than simply bouncing it — and this will mute the multiplier. This is true, particularly if the trades are brokered. It is therefore more reasonable to think about the example in terms of *net* customer orders, rather than gross.

The role of time in the empirical microstructure literature has only recently emerged. Two important contributions are Hasbrouck (1991) and Hausman, Lo, and MacKinlay (1992). Hasbrouck decomposes the variance of stock price changes into trade–correlated and trade–uncorrelated components, and finds trades are more informative at the beginning of the trading day. Also working with stocks, Hausman et al test for exogeneity of the length of time between transactions, which they reject at conventional significance levels. However, they argue that their estimates do not change when endogeneity is accounted for via instrumental variables. On the basis of this, they forge ahead with the assumption of exogenous inter–transaction times.

Empirical microstructure work in FX has been constrained until recently by a lack of transaction–level data. The paper most closely related to the analysis here is Lyons (1993a), which introduces a transactions dataset that is a subset of the data used here (namely, dealt quotes only). That paper presents evidence supporting both of the two main branches of microstructure theory: the asymmetric–information branch and the inventory–control branch. Though many papers have provided evidence supporting the asymmetric–information branch, little or no direct evidence had previously been found in support of the inventory–control branch [see for example Madhavan and Smidt (1991), Manaster and Mann (1993), and the overview in O'Hara (1994)]. The fact that they are both present provides further impetus for the application of microstructural models to the FX market. The application here extends previous work by addressing the informational subtleties of order flow.

The paper is organized as follows: Section 2 presents a model of transaction prices that includes a relationship between trading intensity and the information content of trades; Section 3 describes the data; Section 4 presents our results; and Section 5 concludes.

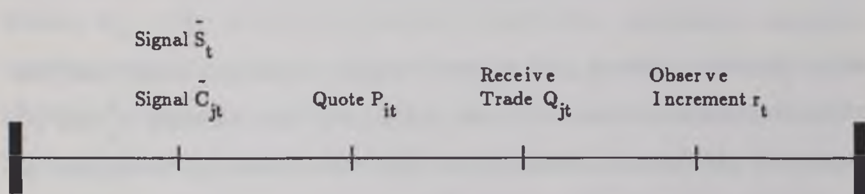
2. A Model in Which Time Matters

The following model extends the model of Madhavan and Smidt (1991) by incorporating a role for inter–transaction time. As they do, we will exploit the model's ability to disentangle the information effects of trades from the inventory control effects. The result is a richer characterization of market depth.

There are two assets in a pure exchange economy: one riskless (the numeraire) and one with a stochastic liquidation value — representing FX. The FX market is organized as a decentralized dealership market with n dealers. Here, we focus on the pricing behavior of a representative dealer, denoted dealer i. A period is defined by a transaction effected against dealer i's quote, with periods running from t=1,2,...,T. Let j denote the dealer requesting i's quote in any period. Figure 1 summarizes the timing in each period:

Figure 1

Sequencing in each period



* Definitions: $\bar{S}_t$ is a public signal of the full information value V_t; $\bar{C}_{jt}$ is dealer j's private signal of V_t, where j denotes the dealer requesting the quote from dealer i; P_{it} is dealer i's bilateral quote to dealer j, a schedule matching each transaction quantity with a price; Q_{jt} is the signed quantity traded, positive for dealer j purchases, negative for sales; and r_t is the period t increment to V_t.

2.1. The Information Environment

The full information price of FX at time T is denoted by $\tilde{V}$, which is composed of a series of increments — e.g. interest differentials — so that $\tilde{V} = \sum_{i=0}^{T} \tilde{r}_i$, where r_0 is a known constant. The increments are i.i.d. mean zero. Each increment r_t is realized immediately after trading in period t. Realizations of the increments can be thought to represent the flow of public information over time. The value of FX at t is thus defined as $V_t = \sum_{i=0}^{t} r_i$. At the time of quoting and trading in period t, i.e. before $\tilde{r}_t$ is realized, $\tilde{V}_t$ is a random variable. In a market without private information or transaction costs the quoted price of FX at time t, denoted P_t, would be equal to V_{t-1}, which is the expected value of the asset price conditional on public information available at t.

The following two signals define each period's information environment prior to dealer i's quote to dealer j:

$$\tilde{S}_t = V_t + \tilde{\eta}_t \tag{1}$$

$$\tilde{C}_{jt} = V_t + \tilde{\omega}_{jt} \tag{2}$$

where the noise terms η_t and ω_{jt} are normally distributed about zero, are independent of one another and across periods, and have variances σ_η^2 and σ_ω^2 respectively. At the outset of each period t, all dealers receive a public signal S_t of the full-information value V_t. Also at the outset of each period t, dealer j — the dealer requesting a quote — receives a private signal C_{jt} of V_t. In the FX market, one potential source of private signals at the dealer level is order flow from non-dealer customers; because each dealer has sole knowledge of his own-customer order flow, to the extent this flow conveys information it is private information, which can be exploited in inter-dealer trading [see, for example, Goodhart (1988) page 456, and Lyons (1994)].

Dealer i conditions on S_t, and then quotes his schedule as a function of Q_{jt}. The schedule's sensitivity to Q_{jt} insures that any realization of Q_{jt} will be regret–free for the quoting dealer, in the sense of Glosten and Milgrom (1985). That is, the quote takes account of the adverse selection arising from dealer j's additional information $\bar{C}_{jt}$. Of course, the realization of Q_{jt} still provides dealer i a signal of $\bar{C}_{jt}$. As is standard, the signed quantity dealer j chooses to trade is linearly related to the deviation between dealer j's expectation and the transaction price, plus a quantity representing liquidity demand X_{jt} that is uncorrelated with V_t:

$$Q_{jt} = \theta(\mu_{jt} - P_{it}) + X_{jt} \tag{3}$$

where μ_{jt} is the expectation of V_t conditional on information available to dealer j at t, and the value of X_{jt} is known only to dealer j. [The demand function that supports equation (3) requires either exponential utility defined over a single period, or mean–variance optimization over multiple periods.]

We introduce a role for time in the model via equation (3) and the liquidity demand X_{jt}. The hot–potato hypothesis of order–flow information associates liquidity demand X_{jt} with inventory–adjustment trading. In FX — according to the hypothesis — innovations in non–dealer order flow spark repeated inter–dealer trading of idiosyncratic inventory imbalances. This rapid passing of the hot–potato generates a relatively large role for liquidity trades in periods of short inter–transaction times. The event–uncertainty hypothesis, in contrast, associates short inter–transaction times with a relatively large role for informative trading: in the presence of event uncertainty, intense trading is a signal that an information event has occurred. To summarize, for given precisions of the signals C_{jt} and S_t, we can characterize these views as:

Hot–potato hypothesis: $\sigma^2_{Xj} \begin{cases} \text{high when inter–transaction times are short} \\ \text{low when inter–transaction times are long} \end{cases}$

Event–uncertainty hypothesis: $\sigma^2_{Xj} \begin{cases} \text{low when inter–transaction times are short} \\ \text{high when inter–transaction times are long} \end{cases}$

This change in the relative intensity of liquidity trading will alter the signal extraction problem faced by the quoting dealer, to which we now turn.

2.2. The Formation of Expectations

Dealer *i*'s quotes depend on his conditional expectation of V_t at the time of quoting, which we denote μ_{it}. This expectation, in turn, is a function of the variables described above: S_t and Q_{jt}; the third variable described above, $\bar{C}_{jt}$, is communicated (noisily) to dealer *i* via Q_{jt}.

We now address the determination of this expectation μ_{it}. Dealer *i*'s prior belief regarding V_t is summarized by the public signal S_t. Dealer i then considers the "what if" of various possible Q_{jt}'s. In particular, from any Q_{jt} dealer *i* can form the statistic Z_{jt} (see appendix):

$$Z_{jt} \equiv \frac{Q_{jt}/\theta + P_{it} - \lambda S_t}{1-\lambda} = V_t + \bar{\omega}_{jt} + [1/\theta(1-\lambda)]\bar{X}_{jt} \qquad (4)$$

where $\lambda \equiv \sigma^2_{\omega}/(\sigma^2_{\eta}+\sigma^2_{\omega})$. This statistic is normally distributed, with mean V_t and variance equal to the variance of the last two terms, both of which are orthogonal to V_t. Via X_{jt}, the variance of the second of these two terms is a function of inter–transaction times, per above. Let σ^2_{Zs} denote the variance of the statistic Z_{jt} when inter–transaction times are short, and let σ^2_{Zl} denote the variance of Z_{jt} when inter–transaction times are long.

Since Z_{jt} is statistically independent of S_t, dealer i's posterior μ_{it}, expressed as a function of any Q_{jt}, takes the form of a weighted average of S_t and Z_{jt}:

$$\mu_{it} = \kappa_k S_t + (1-\kappa_k) Z_{jt} \qquad k = s,l \tag{5}$$

where $\kappa_s \equiv \sigma^2_{Zs}/(\sigma^2_{Zs}+\sigma^2_\eta)$ and $\kappa_l \equiv \sigma^2_{Zl}/(\sigma^2_{Zl}+\sigma^2_\eta)$. This expectation plays a central role in determining dealer i's quote. Note that $\kappa_s > \kappa_l$ if $\sigma^2_{Zs} > \sigma^2_{Zl}$, i.e., if liquidity trading is relatively important when inter-transaction times are short.

2.3. The Determination of Bid/Offer Quotes

Consider the following prototypical inventory-control model. Here, price is linearly related to the dealer's current inventory — a specification that is optimal in a number of inventory control models:

$$P_{it} = \mu_{it} - \alpha(I_{it} - I^*_i) + \gamma D_t \tag{6}$$

where μ_{it} is the expectation of V_t conditional on information available to dealer i at t, I_{it} is dealer i's current inventory position, and I^*_i is i's desired position. The inventory-control effect, governed by α, will in general be a function of relative interest rates, firm capital, and carrying costs. The variable D_t is a direction-indicator variable with a value of 1 when a buyer-initiated trade occurs, and a value of −1 when a seller-initiated trade occurs. The term γD_t then picks up (half of) the effective spread: if dealer j is a buyer then the realized transaction price P_{it} will be on the offer side, and therefore a little higher, ceteris paribus. This term can be interpreted as compensation resulting from execution costs, price discreteness, or rents.

Consistent with the regret-free property of quotes, we substitute dealer i's expectation conditional on possible Q_{jt}'s — equation (5) — into equation (6), yielding:

$$P_{it}=\kappa_k S_t + (1-\kappa_k)Z_{jt} - \alpha(I_{it}-I_i^*) + \gamma D_t \qquad k= s,l \tag{7}$$

which is equivalent to (see appendix):

$$P_{it} = S_t + \left[\frac{1-\phi_k}{\phi_k \theta}\right]Q_{jt} - \left[\frac{\alpha}{\phi_k}\right](I_{it}-I_i^*) + \left[\frac{\gamma}{\phi_k}\right]D_t \tag{8}$$

where $\phi_k \equiv (\kappa_k-\lambda)/(1-\lambda)$ and $0<\phi_k<1$ since $0<\kappa_k<1$, $0<\lambda<1$, and $\kappa_k>\lambda$.

2.4. An Estimable Equation

Equation (8) is not directly estimable because S_t is not observable to the econometrician. Our assumptions about the signals available and the evolution of V_t allow us to express the period *t* prior S_t as equal to the period *t−1* posterior from equation (6) lagged one period, plus an expectational error term ϵ_{it}:

$$S_t= \mu_{it-1}+ \epsilon_{it} = P_{it-1}+ \alpha(I_{it-1}-I_i^*) - \gamma D_{t-1} + \epsilon_{it}. \tag{9}$$

Substituting this expression for S_t into equation (8) yields:

$$P_{it} = \left[P_{it-1}+\alpha(I_{it-1}-I_i^*)-\gamma D_{t-1}+\epsilon_{it}\right] + \left[\frac{1-\phi_k}{\phi_k \theta}\right]Q_{jt} - \left[\frac{\alpha}{\phi_k}\right](I_{it}-I_i^*) + \left[\frac{\gamma}{\phi_k}\right]D_t.$$

which implies:

$$\Delta P_{it}= \left[\frac{\alpha}{\phi_k} - \alpha\right]I_i^* + \left[\frac{1-\phi_k}{\phi_k \theta}\right]Q_{jt} - \left[\frac{\alpha}{\phi_k}\right]I_{it} + \alpha I_{it-1} + \left[\frac{\gamma}{\phi_k}\right]D_t - \gamma D_{t-1}+ \epsilon_{it}. \tag{10}$$

This corresponds to a reduced form estimating equation of:

$$\Delta P_{it} = \beta_0 + \beta_1 Q_{jt} + \beta_2 I_{it} + \beta_3 I_{it-1} + \beta_4 D_t + \beta_5 D_{t-1} + \epsilon_{it}. \tag{11}$$

Thus, the change in the transaction price from t–1 to t is linearly related to: (i) the signed incoming order at t, (ii) the inventory level at t, (iii) the inventory level at t–1, (iv) whether P_{it} is at the bid or offer, and (v) whether P_{it-1} is at the bid or offer. Note that the last two regressors — the indicator variables D_t and D_{t-1} — are accounting for bid–offer bounce. The model predicts that $\{\beta_1,\beta_3,\beta_4\}>0$, $\{\beta_2,\beta_5\}<0$, $|\beta_2|>\beta_3$, and $\beta_4>|\beta_5|$, irrespective of the inter–transaction time. (The latter inequalities derive from the fact that $0<\phi_k<1$.) These more general predictions are borne out in the data, and are presented in Lyons (1993a). Here, our focus is on the information in order flow measured by β_1, which is in turn a function of our structural parameter κ from equation (5). That is, we want to test whether the coefficient β_1 is sensitive to inter–transaction time, and if so, in which direction. The hot–potato hypothesis predicts a lower β_1 when inter–transaction times are short; the event–uncertainty hypothesis predicts a higher β_1 when inter–transaction times are short. These predictions derive from the relative importance of liquidity trading (σ^2_{Xj}) in the signal extraction problem.

Our final comment on the model concerns the assumption of a time–invariant desired inventory. First, note that with a slight re–interpretation the model can accommodate variability in desired inventories, that is, an I^*_i that varies through time. Consider the following model: $I^*_{it}=I_i+\delta(\mu_{it}-S_t)$, which is consistent with the linear demands arising from negative exponential utility, where the public information S_t represents the market price away from dealer *i*. Further, Q_{jt} is the only information available to dealer *i* that is not reflected in S_t. Under the assumptions of our model, $(\mu_{it}-S_t)$ is proportional to Q_{jt}. Accordingly, we write $(\mu_{it}-S_t)=\pi Q_{jt}$. Hence, we can express the desired inventory as: $I^*_{it}=I_i+\delta\pi Q_{jt}$. In estimation, I_i is absorbed in the constant. The estimate of β_1 now represents

$\left[\frac{1-\phi_k}{\phi_k \theta}\right] Q_{jt} + \left[\frac{\alpha}{\phi_k} - \alpha\right] \delta\pi$, whose significance still evinces an information effect, though we have to be more careful in interpreting its magnitude.

3. Data

Our dataset has significant advantages over FX data used in the past, in particular Reuters indications data [see for example Goodhart (1989), and Bollerslev and Domowitz (1993)]. The main shortcomings of the Reuters indications are three: first, these are only indications, not firm quotes at which dealers can transact; second, there is no measure of order flow or transaction prices; and third, the spreads in the indications dataset are 2 to 3 times the size of firm quotes in the inter-dealer market.

Our dataset consists of two linked components, covering the five trading days of the week August 3–7, 1992, from the informal start of trading at 8:30 EST to roughly 1:30 EST. The first component includes the time-stamped quotes, prices, and quantities for all the direct inter-dealer transactions of a single DM/$ dealer at a major New York bank. The second component comprises the same dealer's position cards, which includes all indirect (brokered) trades.

3.1. Dealer Data: Direct Quotes and Trades

The first component of the dataset includes the dealer's quotes, prices, and quantities for all direct transactions. The availability of this component is due to a recent change in technology in this market: the Reuters Dealing 2000–1 system. This system — very different from the system that produces the Reuters indications — allows dealers to communicate quotes and trades *bilaterally* via computer rather than verbally over the telephone.[1] Among other things, this allows dealers to

[1] Dealing 2000–1 is also very different than Dealing 2000–2. The former is wholly bilateral, while the latter is akin to an electronic broker, where multiple dealers participate.

request up to four quotes simultaneously, whereas phone requests are necessarily sequential. Another advantage is that the computerized documentation reduces the paperwork required of the dealers. Though use of this technology differs by dealer and is currently diffusing more widely, our dealer uses Dealing 2000–1 for nearly all of his direct interbank trades: less than 0.4% of all transactions were done over the phone over our sample week (as indicated on the position cards).

Each record of the data covering the dealer's direct trading includes the first 5 of the following 7 variables; the last two are included only if a trade takes place:

(1) The time the communication is initiated (to the minute, with no lag).

(2) Which of the two dealers is requesting the quote.

(3) The quote quantity.

(4) The bid quote.

(5) The offer quote.

(6) The quantity traded, (which provides Q_{jt}).

(7) The transaction price (which provides P_{it}).

This component of the dataset includes 952 transactions amounting to $4.1 billion.

Figure 2 provides an example of a dealer communication as recorded by the Dealing 2000–1 printout [see Reuters (1990) for more details]. The first word indicates that the call came "From" another dealer. Then comes the institution code and name of the counterparty, followed by the time (Greenwich Mean, computer assigned), the date (day first), and the number assigned to the communication. On line 3, "SP DMK 10" identifies this as a request for a spot DM/$ quote for up to $10 million. Line 4 provides the quoted bid and offer price: typically, dealers only quote the last two digits of each price, the rest being superfluous in such a fast–moving market. These two quotes correspond to a bid of 1.58<u>88</u> DM/$ and an offer of 1.58<u>91</u> DM/$. In confirming the transaction, the communication record provides the first three digits. Here, the calling dealer buys

$10 million at the D–mark offer price of 1.5891. The record confirms the exact price and quantity. In our dataset, transactions never take place within the spread; the transaction price always equals either the bid or the offer.

Figure 2

Example of a Reuters Dealing 2000–1 Communication

```
From  CODE     FULL NAME HERE    * 1250GMT  030892 */1080
Our Terminal : CODE    Our user    : DMK
    SP DMK 10
#  8891
    BUY

#   10 MIO AGREED
#   VAL 6AUG92
#   MY DMK TO FULL NAME HERE
#   TO CONFIRM AT 1.5891 I SELL 10 MIO USD
#
      TO CONFIRM AT 1.5891 I SELL 10 MIO USD
      VAL 6AUG92
      MY USD TO FULL NAME HERE  AC 0–00–00000
      THKS N BIFN
#
#   #END LOCAL#
#
^  ## WRAP UP BY DMK DMK 1250GMT 3AUG92

^  #END#

(  265 CHARS)
```

* "From" establishes this as an incoming call; the caller's four–digit code and institution name follow; "GMT" denotes Greenwich Mean Time; the date follows, with the day listed first; "SP DMK 10" identifies this as request for a spot, DM/$ quote for up to $10 million; "8891" denotes a bid of 88 and an offer of 91 (only the last two digits are quoted); the confirmation provides the complete transaction price, and verifies the transaction quantity.

3.2. Dealer Data: Position Cards

The second component of the dataset is composed of the dealer's position cards over the same five days covered by the direct–transaction data, August 3–7, 1992.

In order to track their positions, spot dealers record all transactions on hand-written position cards as they go along. An average day consists of approximately 20 cards, each with about 15 transaction entries.

There are two key benefits to this component of the dataset. First, it provides a very clean measure of the dealer's inventory I_t at any time since it includes both direct trades *and* any brokered trades. Second, it provides a means of error–checking the first component of the dataset.

Each card includes the following information for every trade:

(1) The signed quantity traded (which determines I_t),

(2) The transaction price, and

(3) The counterparty, including whether brokered.

Note that the bid/offer quotes at the time of the transaction are not included so this component of the dataset alone is not sufficient for estimating our model. Note also that each entry is not time–stamped; at the outset of every card, and often within the card too, the dealer records the time to the minute. Hence, the exact timing of some of the brokered transactions is not pinned down since these trades do not get confirmed via a Dealing 2000–1 record. Nevertheless, this is not a drawback for our purposes: the observations for our empirical model are the direct transactions initiated at our dealer's quoted prices; since the timing of these is pinned down by the Dealing 2000–1 records, and since these transactions appear sequentially in both components, the intervening changes in inventory due to brokered trades can be determined exactly.

3.3. Descriptive Statistics

Table 1 presents the data in the form of daily averages to convey a sense of the typical day's activity. This is masking some daily variation in the sample: the heaviest day (8/7/92) is a little less than twice as active as the lightest day (8/5/92). Note that this dealer averages well over $1 billion of inter–dealer trading daily (brokered trades are necessarily inter–dealer). With respect to quoting, because our dealer is among the larger in this market, he has $10 million "relationships" with many other dealers; that is, quote requests from other high–volume dealers that do not specify a quantity are understood to be good for up to $10 million. Note the tightness of the median spread. For comparison, the median spread in the Reuters indications dataset is DM 0.001, more than three times as large. A bid/offer spread of 3 pips is less than 0.02% of the spot price.

Table 1 Here

A natural concern is whether our dealer is representative of the larger dealers in the spot market. While we cannot answer this definitively, we offer a few relevant facts. First, he has been trading in this market for many years and is well–known among the other major dealers. Second, in terms of trading volume he is without a doubt one of the key players, trading well over $1 billion per day and maintaining $10 million quote relationships with a number of other dealers. Though this would probably not put him in the top five in terms of volume, he is not far back, possibly in the 5th to 15th range somewhere. In the end, our view is that he is representative, at least with respect to the issues addressed here. There is no doubt, however, that different trading styles exist.

3.4. Relevant Institutional Background

Here, we highlight two institutional factors relevant to our analysis: (i) trading limits imposed on dealers and (ii) trading on the IMM futures market. As for trading limits, there is an important distinction between intraday limits and overnight limits. At our dealer's bank, which is typical of major banks, there are no explicit intraday limits on senior dealers, though dealers are expected to communicate particularly large trades to their immediate supervisor (about $50 million and above for many banks in the current DM/$ market). In contrast, most banks impose overnight limits on their dealers. Currently, a common overnight limit on a single dealer's open position is about $75 million, considerably larger than the largest open position in our sample. Most dealers, however, close their day with a zero net position; carrying an open position means monitoring it through the evening, an unattractive prospect after a full day of trading. Our dealer ended his day with a zero net position each of the five days in our sample. Finally, though broader risk–management programs are in place at the bank for which our dealer trades, it is rare in FX that a dealer's position is hedged because it aggregates unfavorably with others; and when this does occur, it is typically without the participation of the individual dealer.

As for trading on the IMM futures market while dealing spot, this differs by dealer. We stress, though, that unlike equity markets, the spot FX market is many times larger than the futures market: in 1992 the average daily volume in New York in spot DM/$ was roughly $50 billion [New York FED (1992), adjusted for double counting]; in the same year the average daily volume on all IMM DM/$ contracts was less than $5 billion. As for our dealer, his position cards show that he traded less than $1 million daily in futures over the sample period, which is negligible relative to his daily spot volume. Like other spot dealers, he does listen to an intercom that communicates futures prices. However, this intercom is less important to a spot dealer than the intercoms connected to inter–dealer brokers in the spot market.

4. Estimation Results

We begin with our results from direct estimation of the model in equation (11), which are presented in Table 2. Though these estimates do not include any role for inter–transaction time, they provide a benchmark for the later results regarding the hot–potato and event–uncertainty hypotheses. Note that these estimates are essentially a replication of a result presented in Lyons (1993a). Accordingly, we refer readers to that earlier work for more detailed interpretation.

Table 2 Here

Given these benchmark results, henceforth we present only those coefficients which bear on the information content of order flow — namely variations of β_1 from equation 11. All non-reported coefficients remain significant at at least the 5% level, with the predicted signs and relative magnitudes. Presenting the results this way allows us to focus on the informational subtleties outlined in section 2.

4.1. The Core Model of Trading Intensity

Table 3 presents our estimates of the information content of order flow, distinguishing between short and long inter-transaction times. This is achieved via the introduction of dummy variables s_t and l_t (see the equation heading the table). The dummy s_t equals 1 if inter-transaction time is short, 0 otherwise; the dummy l_t equals 0 if inter-transaction time is short, 1 otherwise. Short inter-transaction times are defined two ways: less than 1 minute from the previous transaction and less than 2 minutes. The time stamps on our data are very precise, since they are assigned by the computer; however, they do not provide precision beyond the minute. Hence, less than 1 minute includes trades with the same time stamp; less than 2 minutes includes trades with time stamps differing by 1 minute or less. These categories bracket the mean inter-transaction time of 1.8 minutes. The second category corresponds to a break at the median inter-transaction time.

Table 3 Here

The results provide strong support for the hot-potato hypothesis over the event-uncertainty hypothesis. The coefficient β_1 — which measures the information effect of incoming trades with short inter-transaction times — is insignificant at conventional levels. In contrast, the coefficient β_1' — which measures the information effect of incoming trades with long inter-transaction times — is significant. Moreover, a test of the restriction that $\beta_1=\beta_1'$ is rejected at the 1% level in both cases. In summary, *trades occurring when transaction intensity is high are significantly less informative than trades occurring when transaction intensity is low.* This is the main result of the paper.

4.2. The Pattern of the Market

There is an additional testable implication of the hot–potato hypothesis: it follows directly from the story of bouncing inventories outlined in section 1 that these discretionary liquidity trades will tend to be in the same direction (i.e., have the same sign). The obverse is that clumped trading is more likely to be hot–potato (liquidity) trading if trades follow in the same direction. The implication for prices is that, even if Martingales, they are not necessarily Markov.

The test presented in Table 4 addresses this question: Is clumped order–flow less informative when transactions follow the same direction? Again, we introduce dummy variables, in this case s_t, o_t, and l_t (see the equation heading the table). The dummy s_t equals 1 if (i) inter–transaction time is short and (ii) the previous incoming trade has the same direction, 0 otherwise; the dummy o_t equals 1 if (i) inter–transaction time is short and (ii) the previous incoming trade has the opposite direction, 0 otherwise; the dummy l_t equals 0 if inter–transaction time is short, 1 otherwise. A short inter–transaction time is defined as less than the median of 2 minutes.

Table 4 Here

Once again, the results support the hot–potato hypothesis. The coefficient β_1 — short inter–transaction times and same direction — is insignificant. In contrast, the coefficient β_1' — short inter–transaction times and opposite direction — is significant. A test of the restriction that $\beta_1=\beta_1'$ is rejected at the 1% level. To summarize, *clumped trades occurring in the same direction are significantly less* *informative than clumped trades occurring in the opposite direction.*

4.3. Another Measure of Market Pace: Quote Intensity

The results of Table 4 highlight another important observation: though the hot–potato and event–uncertainty hypotheses make opposite predictions regarding the relation between information and trading intensity, they are not necessarily competing hypotheses. That is, both effects could be operative: hot–potato trading simply dominates when trading is most intense in this market.

To examine whether there is independent support for event–uncertainty, we exploit an "instrument" that is arguably more closely related to event–uncertainty than inventory–control. To understand this instrument, recognize that in Easley and O'Hara (1992) transaction intensity per se is the only dimension of trading intensity available for signalling the underlying state. The problem for our purposes is that transaction intensity is also the linchpin of the hot–potato model. Our dataset, on the other hand, includes a second dimension of trading intensity: quoting intensity. The roughly 4:1 ratio of not–dealt quotes to dealt quotes in Table 1 indicates that transactions alone may not be telling the full story. More important for discriminating event–uncertainty from hot–potato is the fact that quote requests per se typically signal heightened uncertainty and information gathering, whereas hot–potato transactions minimize on quote requests in order to unload inventory rapidly. In short, quoting intensity provides another vehicle for Easley and O'Hara.

Table 5 presents estimates of the information content of order flow, distinguishing between high and low quoting intensity as a measure of market pace. Once again we introduce dummy variables, in this case h_t and l_t (see the equation heading the table). The dummy h_t equals 1 if the total number of intervening quotes per minute is high, 0 otherwise; the dummy l_t equals 0 if the total number of intervening quotes per minute is high, 1 otherwise. The different definitions of a high number of intervening quotes appear in column one. These quotes are from the Dealing 2000–1 portion of the dataset, described in subsection 3.1.

Table 5 Here

These results provide support for the event–uncertainty hypothesis. The coefficient β_1 reflecting high quoting intensity is significant, whereas the coefficient β_1' reflecting low quoting intensity is insignificant. A test of the restriction that $\beta_1=\beta_1'$ is rejected at the 5% level in all three cases. To summarize, *trades occurring when trading intensity is high — where trading intensity is proxied by quoting intensity — are significantly more informative than trades occurring when trading intensity is low.*

5. Conclusions

Our results suggest that in FX: trading begets trading. The trading begotten is relatively uninformative, arising from repeated passage of idiosyncratic inventory imbalances among dealers. Clearly, this could not arise under a specialist microstructure. A broad implication is that a microstructural understanding of this market requires much richer multiple–dealer theory than now exists.

Our principal empirical findings are the following:

(1) Trades occurring when transaction intensity is high are significantly less informative than trades occurring when transaction intensity is low.

(2) Clumped trades occurring in the same direction are significantly less informative than clumped trades occurring in the opposite direction.

(3) Trades occurring when trading intensity is high — where trading intensity is proxied by quoting intensity — are significantly more informative than trades occurring when trading intensity is low.

We interpret results (1) and (2) as supportive of hot–potato trading among dealers in FX. We interpret result (3) as supportive of the Easley and O'Hara event–uncertainty hypothesis, though the vehicle differs from the transaction–focus of their paper. Taken together, the results highlight the potential complementarity between these seemingly polar views.

There is an important hardship in focusing on a dealership market like FX that warrants recognition. Empirical work on the specialist structure has the luxury of describing the behavior of a lone dealer. It is much more difficult to argue that by documenting the behavior of a single dealer in the FX market we have similarly captured the FX market. The data required to generate a more complete picture are out of the question given current availability. Nevertheless, the dealer we have tracked is without a doubt one of the key players in this market, trading well over $1 billion per day and maintaining $10 million quote relationships with a number of other dealers. Is he representative of dealers in the core of the wholesale spot market? We would argue yes, at least with respect to the issues addressed here. But, there is no doubt that different dealers have different trading styles.

References

Admati, A., and P. Pfleiderer (1988): "A Theory of Intraday Patterns: Volume and Price Variability," Review of Financial Studies, 1: 3–40.

Amihud, Y., and H. Mendelson (1980): "Dealership Market: Market Making with Inventory," Journal of Financial Economics, 8: 31–53.

Bessembinder, H. (1994): "Bid–Ask Spreads in the Interbank Foreign Exchange Markets," Journal of Financial Economics, 35: 317–348.

Bollerslev, T., and I. Domowitz (1993): "Trading Patterns and Prices in the Interbank Foreign Exchange Market," Journal of Finance, 48: 1421–1444.

Easley, D., and M. O'Hara (1987): "Price, Quantity, and Information in Securities Markets," Journal of Financial Economics, 19: 60–90.

Easley, D., and M. O'Hara (1992): "Time and the Process of Security Price Adjustment," Journal of Finance, 47: 577–605.

Flood, M. (1992): "Market Structure and Inefficiency in the Foreign Exchange Market," Journal of International Money and Finance, 13: 131–158.

Foster, D., and S. Vishwanathan (1990): "Variations in Volumes, Variances, and Trading Costs," Duke University Working Paper #88–108.

Garman, M. (1976): "Market Microstructure," Journal of Financial Economics, 3: 257–275.

Glosten, L., and L. Harris (1988): "Estimating the Components of the Bid–Ask Spread," Journal of Financial Economics, 21: 123–142.

Glosten, L., and P. Milgrom (1985): "Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents," Journal of Financial Economics, 14: 71–100.

Goodhart, C. (1988): "The Foreign Exchange Market: A Random Walk With a Dragging Anchor," Economica, 55: 437–460.

Goodhart, C. (1989): "'News' and the Foreign Exchange Market," Manchester Statistical Society Paper, October 17.

Goodhart, C., and L. Figliuoli (1991): "Every Minute Counts in Financial Markets," Journal of International Money and Finance, 10: 23–52.

Hasbrouck, J. (1988): "Trades, Quotes, Inventories, and Information," Journal of Financial Economics, 22: 229–252.

Hasbrouck, J. (1991): "The Summary Informativeness of Stock Trades: An Econometric Analysis," Review of Financial Studies, 4: 571–595.

Hausman, J., A. Lo, and C. MacKinlay (1992): "An Ordered Probit Analysis of Transaction Stock Prices," Journal of Financial Economics, 31: 319–379.

Ho, T., and H. Stoll (1983): "The Dynamics of Dealer Markets Under Competition," Journal of Finance, 38: 1053–1074.

Kyle, A. (1989): "Informed Speculation with Imperfect Competition," Review of Economic Studies, 56: 317–356.

Laux, P. (1991): "Dealer Market Structure and the Bid–Ask Spread: Theory and Empirical Evidence," mimeo, University of Texas at Austin.

Lee, C., and M. Ready (1991): "Inferring Trade Direction from Intradaily Data," Journal of Finance, 46: 733–746.

Lyons, R. (1994): "Information Externalities in the Microstructure of the Foreign Exchange Market," Berkeley Business School manuscript, revised version of NBER Working Paper #3889, October 1991.

Lyons, R. (1993a): "Tests of Microstructural Hypotheses in the Foreign Exchange Market," NBER Working Paper #4471, forthcoming in the Journal of Financial Economics.

Lyons, R. (1993b): "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange," NBER Working Paper #4467, September.

Madhavan, A., and S. Smidt (1991): "A Bayesian Model of Intraday Specialist Pricing," Journal of Financial Economics, 30: 99–134.

Manaster, S., and S. Mann (1993): "Life in the Pits: Competitive Marketmaking and Inventory Control," mimeo, University of Utah, May.

O'Hara, M. (1994): Market Microstructure Theory, Blackwell Publishers, forthcoming, October.

O'Hara, M., and G. Oldfield (1986): "The Microeconomics of Market Making," Journal of Financial and Quantitative Analysis, 21: 361–376.

Reuters Ltd. (1990): "The Reuter Dealing 2000–1 Service: User Guide, Version 3," London: Reuters Ltd.

Roell, A. (1990): "Dual–Capacity Trading and the Quality of the Market," Journal of Financial Intermediation, 1: 105–124.

Stoll, H. (1989): "Inferring the Components of the Bid–Ask Spread: Theory and Empirical Tests," Journal of Finance, 44: 115–134.

Appendix

Derivation of the Statistic Z_{jt} in Equation (4)

Beginning with equation (3):

$$Q_{jt}= \theta(\mu_{jt}-P_{it}) + X_{jt} \tag{3}$$

$$\Rightarrow \quad Q_{jt}/\theta + P_{it}= \mu_{jt} + X_{jt}/\theta$$

$$\Rightarrow \quad Q_{jt}/\theta + P_{it}= \lambda S_t+(1-\lambda)C_{jt} + X_{jt}/\theta \qquad \text{where } \lambda\equiv \sigma^2_\omega/(\sigma^2_\eta+\sigma^2_\omega)$$

$$\Rightarrow \quad Q_{jt}/\theta + P_{it}- \lambda S_t = (1-\lambda)(V_t+\bar{\omega}_{jt}) + \bar{X}_{jt}/\theta \qquad \text{since } C_{jt}= V_t+\bar{\omega}_{jt}$$

$$\Rightarrow \quad Z_{jt}\equiv \frac{Q_{jt}/\theta + P_{it}- \lambda S_t}{1-\lambda} = V_t + \bar{\omega}_{jt} + [1/\theta(1-\lambda)]\bar{X}_{jt}. \tag{4}$$

Derivation of the Price Representation in Equation (8)

Beginning with equation (6):

$$P_{it} = \mu_{it} - \alpha(I_{it}-I^*_i) + \gamma D_t \tag{6}$$

we can write:

$$\mu_{it} = \kappa_k S_t + (1-\kappa_k)Z_{jt} \qquad \text{where } \kappa_k\equiv\sigma^2_{Zk}/(\sigma^2_{Zk}+\sigma^2_\eta),\ k=s,l$$

$$= \kappa_k S_t + \left[\frac{1-\kappa_k}{1-\lambda}\right]\left[Q_{jt}/\theta + P_{it}- \lambda S_t\right]$$

$$= \kappa_k S_t - \left[\frac{\lambda(1-\kappa_k)}{1-\lambda}\right] S_t + \left[\frac{1-\kappa_k}{1-\lambda}\right]\left[Q_{jt}/\theta + P_{it}\right]$$

$$= \left[\kappa_k - \frac{\lambda(1-\kappa_k)}{1-\lambda}\right] S_t + \left[\frac{1-\kappa_k}{1-\lambda}\right]\left[Q_{jt}/\theta + P_{it}\right]$$

$$\equiv \phi_k S_t + (1-\phi_k)\left[Q_{jt}/\theta + P_{it}\right], \quad k=s,l \quad \text{since} \quad \left[\kappa_k - \frac{\lambda(1-\kappa_k)}{1-\lambda}\right] + \left[\frac{1-\kappa_k}{1-\lambda}\right] = 1$$

Note also that $0<\phi_k<1$ since $0<\kappa_k<1$, $0<\lambda<1$, and $\kappa_k>\lambda$ for both k=s and k=l. Each of these properties follows from the definitions of κ_k and λ and the fact that $\sigma^2_{Zj} = \sigma^2_\omega + [\theta(1-\lambda)]^{-2}\sigma^2_X$.

Substituting this expression for μ_{it} into equation (6) yields:

$$P_{it} = \phi_k S_t + (1-\phi_k)\left[Q_{jt}/\theta + P_{it}\right] - \alpha(I_{it}-I^*_i) + \gamma D_t$$

$$\Rightarrow \quad P_{it} = S_t + \left[\frac{1-\phi_k}{\phi_k\theta}\right] Q_{jt} - \left[\frac{\alpha}{\phi_k}\right](I_{it}-I^*_i) + \left[\frac{\gamma}{\phi_k}\right] D_t \quad (8)$$

Table 1

Overview Statistics
August 3–7, 1992

	Direct	Brokered
(1) Average # transactions daily	190	77
(a) incoming	170	
(b) outgoing	20	
(2) Average value transactions daily	$0.8 B	$0.4 B
(a) incoming	$0.65 B	
(b) outgoing	$0.15 B	
(3) Median transaction size	$3 M	$4 M
(a) incoming	$3 M	
(b) outgoing	$5 M	
(4) Average # quotes daily	924	
(a) made	502	
(b) received	422	
(5) Median quoted spread: Dealt	DM 0.0003	
(a) made	DM 0.0003	
(b) received	DM 0.0003	
(6) Median quoted spread: Not Dealt	DM 0.0003	
(a) made	DM 0.0003	
(b) received	DM 0.0005	

* Data for the dealer's direct (inter–dealer) quotes and transactions are from the Reuters Dealing 2000–1 communications. Incoming refers to transactions initiated by another dealer; outgoing refers to transactions initiated by our dealer. Made refers to quotes made by our dealer; received refers to quotes received by our dealer. The trades in these two columns reflect more than 95% of this dealer's trading; the trades that make up the remainder are executed either (i) over the phone, (ii) with a non–dealer customer, or (iii) in the futures market (IMM). Data for the dealer's brokered transactions are from the dealer's position sheets; it is not possible to identify the aggressor from these data. The dealer's trading day begins at 8:30 AM Eastern Standard Time, and ends around 1:30 PM on average.

Table 2

Benchmark Results

$$(11)\quad \Delta P_{it} = \beta_0 + \beta_1 Q_{jt} + \beta_2 I_{it} + \beta_3 I_{it-1} + \beta_4 D_t + \beta_5 D_{t-1} + \epsilon_{it}$$

	β_0	β_1	β_2	β_3	β_4	β_5	R^2
Estimated	–1.37 (–1.07)	1.34 (2.80)	–0.92 (–3.03)	0.72 (2.46)	10.85 (5.69)	–9.14 (–6.04)	0.22
Predicted		>0	<0	>0	>0	<0	

* T–statistics in parentheses. ΔP_{it} is the change in the transaction price (DM/\$) from t–1 to t. Q_{jt} is the dollar quantity transacted directly at dealer i's quoted prices, positive for buyer–initiated trades (i.e. effected at the offer) and negative for seller–initiated trades (at the bid). I_t is i's position at the end of period t. D_t is an indicator variable with value 1 if the trade is buyer–initiated, and value –1 if seller–initiated. The units of Q_{jt}, I_{it}, and I_{it-1} are such that a coefficient of unity implies a price impact of DM0.0001 for every \$10 million. The units of the indicator variable D_{t-1} are such that a coefficient of 10 implies DM0.0002/\$ between bid and offer at quantity zero. Estimated using OLS, with heteroskedasticity– and autocorrelation–consistent (first–order) standard errors. Sample: August 3–7, 1992, 842 observations.

Table 3

Is order–flow less informative when inter–transaction time is short?

$$\Delta P_{it} = \beta_0 + \beta_1 s_t Q_{jt} + \beta_1' l_t Q_{jt} + \beta_2 I_{it} + \beta_3 I_{it-1} + \beta_4 D_t + \beta_5 D_{t-1} + \epsilon_{it}$$

	β_1 (short)	β_1' (long)	Fraction short	$\beta_1=\beta_1'$ P–value
Inter–transaction time short if:				
Less than 1 minute	–0.01 (–0.01)	2.20 (3.84)	262/842	0.000
Less than 2 minutes	0.76 (1.63)	2.60 (3.40)	506/842	0.009

* T–statistics in parentheses. The coefficient β_1 measures the information effect of trades for which the time from the previous transaction is short (s_t=1 and l_t=0 in the equation in the heading), where short is defined in the first column. The coefficient β_1' measures the information effect of those trades for which the time from the previous transaction is long (s_t=0, l_t=1), where long is defined as not short. The Fraction short column presents the fraction of observations satisfying the corresponding definition of short inter–transaction times. In each case the remaining observations fall into the long category. The P–value column presents the significance level at which the null $\beta_1=\beta_1'$ can just be rejected. ΔP_{it} is the change in the transaction price (DM/$) from t–1 to t. Q_{jt} is the dollar quantity transacted directly at dealer i's quoted prices, positive for buyer–initiated trades (i.e. effected at the offer) and negative for seller–initiated trades (at the bid). The units of Q_{jt} are such that β_1=1 implies a price impact of DM0.0001 for every $10 million. I_t is i's position at the end of period t. D_t is an indicator variable with value 1 if the trade is buyer–initiated, and value –1 if seller–initiated. Estimated using OLS, with heteroskedasticity– and autocorrelation–consistent (first–order) standard errors. Sample: August 3–7, 1992, 842 observations.

Table 4

Is clumped order–flow less informative when transactions follow the same direction?

$$\Delta P_{it} = \beta_0 + \beta_1 s_t Q_{jt} + \beta_1' o_t Q_{jt} + \beta_1'' l_t Q_{jt} + \beta_2 I_{it} + \beta_3 I_{it-1} + \beta_4 D_t + \beta_5 D_{t-1} + \epsilon_{it}$$

β_1 (short & same)	β_1' (short & opposite)	β_1'' (long)	Fraction short & same	Fraction short & opposite	$\beta_1=\beta_1'$ P–value
–0.06 (–0.11)	1.90 (3.01)	2.64 (3.46)	276/842	230/842	0.009

* T–statistics in parentheses. The coefficient β_1 measures the information effect of trades that (i) have short inter–transaction times, defined as less than the median of 2 minutes, and (ii) have the same direction of the previous trade (s_t=1, o_t=0, and l_t=0 in the equation in the heading). The coefficient β_1' measures the information effect of trades that (i) have short inter–transaction times, defined as less than the median of 2 minutes, and (ii) have the opposite direction of the previous trade (s_t=0, o_t=1, l_t=0). The coefficient β_1'' measures the information effect of trades that have long inter–transaction times, defined as greater than or equal to the median of 2 minutes (s_t=0, o_t=0, l_t=1). The Fraction short & same column presents the fraction of observations satisfying the corresponding definition of short & same (similarly for the Fraction short & opposite column). The remaining 336/842 observations fall into the long category. The P–value column presents the significance level at which the null $\beta_1=\beta_1'$ can just be rejected. ΔP_{it} is the change in the transaction price (DM/\$) from t–1 to t. Q_{jt} is the dollar quantity transacted directly at dealer i's quoted prices, positive for buyer–initiated trades (i.e. effected at the offer) and negative for seller–initiated trades (at the bid). The units of Q_{jt} are such that β_1=1 implies a price impact of DM0.0001 for every \$10 million. I_t is i's position at the end of period t. D_t is an indicator variable with value 1 if the trade is buyer–initiated, and value –1 if seller–initiated. Estimated using OLS, with heteroskedasticity– and autocorrelation–consistent (first–order) standard errors. Sample: August 3–7, 1992, 842 observations.

Table 5

Is order–flow more informative when quoting intensity is high?

$$\Delta P_{it} = \beta_0 + \beta_1 h_t Q_{jt} + \beta_1' l_t Q_{jt} + \beta_2 I_{it} + \beta_3 I_{it-1} + \beta_4 D_t + \beta_5 D_{t-1} + \epsilon_{it}$$

	β_1 (high)	β_1' (low)	Fraction high	$\beta_1=\beta_1'$ P–value
Quoting intensity high if:				
≥3 intervening quotes per minute	2.16 (3.42)	0.87 (1.70)	$\frac{301}{842}$	0.046
≥4 intervening quotes per minute	2.41 (3.56)	0.84 (1.66)	$\frac{215}{842}$	0.026
≥5 intervening quotes per minute	2.72 (3.47)	0.89 (1.79)	$\frac{144}{842}$	0.025

* T–statistics in parentheses. The coefficient β_1 measures the information effect of those trades occurring when quoting intensity is high, (h_t=1, l_t=0), where high intensity is defined in the first column by the total number of quotes — both made and received — since the previous incoming transaction. The coefficient β_1' measures the information effect of those trades occurring when quoting intensity is low (h_t=0, l_t=1), where low intensity is defined as not high. The Fraction high column presents the fraction of observations satisfying the corresponding definition of high–intensity quoting. The P–value column presents the significance level at which the null $\beta_1=\beta_1'$ can just be rejected. ΔP_{it} is the change in the transaction price (DM/\$) from t–1 to t. Q_{jt} is the dollar quantity transacted directly at dealer i's quoted prices, where both are positive for buyer–initiated trades (i.e. effected at the offer) and negative for seller–initiated trades (at the bid). The units of Q_{jt} are such that β_1=1 implies a price impact of DM0.0001 for every \$10 million. I_t is i's position at the end of period t. D_t is an indicator variable with value 1 if the trade is buyer–initiated, and value –1 if seller–initiated. Estimated using OLS, with heteroskedasticity– and autocorrelation–consistent (first–order) standard errors. Sample: August 3–7, 1992, 842 observations.

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" --by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

210. Gerard Gennotte and Terry A. Marsh. "Variations in Economic Uncertainty and Risk Premiums on Capital Assets." May 1991.

211. Gerard Gennotte and Hayne Leland. "Low Margins, Derivative Securities, and Volatility." May 1991.

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 244

Optimal Cash Management for Investment Funds

by

Hayne Leland and Gregory Connor

March 1995

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH
Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Optimal Cash Management for Investment Funds

Hayne Leland and Gregory Connor

Haas School of Business
University of California, Berkeley
and
BARRA International

March 1995

Finance Working Paper #244

ABSTRACT

We consider the question of how much cash should be held by an investment fund for transactions purposes. Cash is needed to meet redemptions and rights offerings; it is generated by dividends and contributions. It is assumed the cumulative cash flow follows a random walk, perhaps with a drift. If transactions costs were zero, it would be optimal to keep zero cash balances, since cash reduces expected return and adds to tracking error. But keeping cash balances at zero would be very expensive in the presence of transactions costs, since random walks have infinite variation.

The optimal cash policy requires a "no trade" interval $[0, L^*]$. If cash balances are within this interval, no transfers between cash and portfolio securities takes place. If cash falls beneath zero, securities should be sold to return the cash balance to zero. If cash exceeds L^*, cash should be invested in the portfolio to reduce the cash balance to L^*.

We derive closed form solutions for L^*, and show how this responds to changes in transactions costs and other parameters of cash flows and portfolio returns. Finally, a closed form estimate of expected turnover associated with optimal strategies is derived.

OPTIMAL CASH MANAGEMENT FOR INVESTMENT FUNDS

1. Introduction and Summary

Many institutional portfolio managers face the following problem. They have a benchmark or target portfolio which they wish to mimic or exceed. Typically this portfolio has little or no cash. But due to random cash flows in and out of the portfolio, managers must have some cash available. If trading costs were zero, the optimal "inventory" of cash would be zero: cash would immediately be put into the target portfolio when received; cash would instantly be raised by security sales when needed.

But if trading is costly, and cumulative cash flows follow a random walk (perhaps with drift), maintaining a continuously zero inventory of cash would be very expensive. The frequent small adjustments would, over time, require a very large amount of trading. On the other hand, maintaining a large cash inventory would create tracking error with respect to the benchmark and a lower expected return from investment.

We use a dynamic programming approach to determine the optimal cash inventory, when the investor has a long time horizon. We focus on the case of trading costs which are proportional to the volume of securities sold. First, we show that the nature of the optimal policy will involve no purchase or sale of securities as long as the cash inventory remains within a fixed interval. If the cash on hand ever falls beneath zero, securities will then be sold to bring the inventory back to zero: negative cash positions are not permitted. If cash inventory builds up, no compensating adjustments are made until the cash position exceeds an amount L. At that point, securities are bought, but only in amounts which reduce the cash position back to L. We call the interval $[0,L]$ of cash holdings the "no trade" zone.

The paper derives a closed-form solution for the optimal $L = L^*$, and therefore the "no trade" interval. We show how the size of this interval depends on

- the costs of trading portfolio securities
- the risk and return of the target portfolio
- the drift and variance of random cash flows in and out of the fund
- the relative importance of tracking accuracy vs. trading costs.

Finally, we show how to predict the expected turnover (cash-to-securities and securities-to-cash transactions) which results from following the optimal cash management policy.

2. Formulation of the Problem

Let the target portfolio have expected rate of return μ_p and variance σ_p per time unit. We assume there is no cash in this target portfolio. Let k represent the proportional trading costs (e.g. $k = .01$, if one-way trading costs are 1%). For simplicity we assume that all securities in the target portfolio are bought or sold proportionately, so that the composition of the securities portfolio (and its risk characteristics) remains constant.

Let w_t be the proportion of the fund (per unit) held in cash at time t. Cash will randomly be generated or required as a result of dividends, rights offerings, net withdrawals, etc. Let δ_t denote this random cash flow per unit, which is normally and independently distributed through time with constant mean μ_δ and variance σ_δ per period.[1] The increase or decrease in cash position resulting from sales or purchases of the underlying portfolio at time t will be denoted D_t. It follows that the total change per unit in cash position between

[1] In the limit of continuous time, we assume that cumulative cash flows C follow a random walk, with increments $\delta = dC = \mu_\delta dt + \sigma_\delta dZ$, where Z is a standard Brownian motion. Thus $E(\delta) = \mu dt$ and $E(\delta^2) = \sigma^2 dt$.

periods is given by $w_{t+1} - w_t = \delta_t + D_t$. Trading costs in period t will be $k|D_t|$.

For related problems, it is well known that the optimal transactions strategy is to specify an interval over which no rebalancing between cash and the underlying portfolio takes place.[2] This interval is bounded below by 0, indicating that the fund cannot maintain a negative cash position. If, because of random cash outflows, the cash position becomes negative, sufficient sales of securities must be undertaken to return the cash position to zero.[3] L represents the upper bound of this no-action interval. If because of random cash inflows the cash position exceeds L, cash will be invested in the target portfolio to reduce the cash position to L. (It will be shown later that L does not depend upon t).

Given w_t and the above policy, the sequence of future cash positions per unit $\{w_{t+1}, ...\}$ will be uniquely determined by the random cash flows $\{\delta_{t+1},...\}$. If $w_\tau \in [0,L]$, then $w_{\tau+1} = w_\tau + \delta_\tau$. If $w_\tau + \delta_\tau < 0$, then $w_{\tau+1} = 0$ (and $D_\tau = -(w_\tau + \delta_\tau)$).
If $w_\tau + \delta_\tau > L$, then $w_{\tau+1} = L$ (and $D_\tau = (L - w_\tau - \delta_\tau)$).

Let $V(w_t, 0, L)$ be the expected discounted cost over an infinite horizon, when the period t begins with cash fraction w_t, and the no-action interval is $[0,L]$. (For simplicity, we shall henceforth suppress "0" as an argument of V). V depends upon three cost factors: the expected future transactions costs; the expected return loss due to cash holding; and the total future variance of the tracking error (relative to the benchmark portfolio) resulting from the cash holding w weighted by a monetizing parameter λ. The more important the accuracy of tracking the target, the larger will be λ. The expected return loss over the period $(t, t+1]$ will equal πw_{t+1}, where π is the risk premium (the difference between the

---

[2] See, for example, Magill and Constantinides [1979], Constantinides [1986], Hodges and Neuberger [1989], and Dumas and Luciano [1991].

[3] In the presence of fixed as well as proportional transactions costs, it will be optimal to return the cash position to a positive amount, rather than to zero. We discuss this case briefly in the Section 6.

expected return μ_p on the stock portfolio and the interest rate r per period). The variance of tracking error will be $\sigma_p w_{t+1}^2$.[4]

Finally, we allow for the possibility that the volatility reduction associated with the cash position over the period, $[(1 - (1-w_{t+1})^2]\sigma_p$, may provide an offsetting benefit, monetized by another parameter λ_1. Our base situation is when $\lambda > 0$ and $\lambda_1 = 0$. However, we shall also consider the situation when $\lambda = 0$ and $\lambda_1 > 0$. This latter formulation, while perhaps less realistic from the perspective of a money manager, is consistent with a "CAPM" formulation of investors who trade off risk and return.

We assume V exists and is twice differentiable. $V(w_t, L)$ can be expressed as

$$(1) \qquad V(w_t,L) = E\{\sum_{\tau=t}^{\infty} R^{\tau-t}[k|D_\tau| + \pi w_{\tau+1} + \lambda w_{\tau+1}^2 \sigma_p - \lambda_1[1-(1-w_{\tau+1})^2]\sigma_p\}$$

$R = 1/(1+r)$ is the discount rate, where r is the rate of interest per unit time. L affects the expected amount of trading and average size of w_t, and therefore affects V.

We may substitute for D_τ using the relationship $D_\tau = w_{\tau+1} - w_\tau - \delta_\tau$. From recursive summation,

$$(2) \qquad V(w_t, L) = E\{k|w_{t+1} - w_t - \delta_t| + \pi w_{t+1} + \lambda w_{t+1}^2 \sigma_p - \lambda_1[1-(1-w_{t+1})^2 \sigma_p] + R[V(w_{t+1}, L)]\}$$

We may expand $V(w_{t+1}, L)$ in a Taylor Series expansion to get

---

[4] We assume the following time sequencing. w_t is the cash position entering period t. Immediately thereafter, δ_t is realized and D_t is uniquely determined as previously described. Then $w_{t+1} = w_t + \delta_t + D_t$. The loss of expected return and tracking error are caused by w_{t+1} over the period $(t,t+1]$. w_{t+1} then is the cash position entering period $t+1$.

$$(3) \quad V(w_t,L) = E\{k|w_{t+1} - w_t - \delta_t| + \pi w_{t+1} + \lambda w_{t+1}{}^2\sigma_p - \lambda_1 [1-(1-w_{t+1})^2]\sigma_p + R[V(w_t,L) + V_w(w_t,L)(w_{t+1}-w_t) + .5V_{ww}(w_t,L)(w_{t+1}-w_t)^2]\}$$

where $V_w(w_t,L) = \partial V(w_t,L)/\partial w_t$, etc. This can be rewritten as

$$(4) \quad V(w_t,L)(1-R) = E\{k|w_{t+1} - w_t - \delta_t| + w_{t+1}\pi + \lambda w_{t+1}{}^2\sigma_p - \lambda_1 [1-(1-w_{t+1})^2]\sigma_p + RV_w(w_t,L)(w_{t+1}-w_t) + .5RV_{ww}(w_t,L)(w_{t+1}-w_t)^2\}.$$

We now consider when the time between successive periods, dt, is short, and $w_t + \delta_t$ is in the "no trade" interval [0, L]. It follows that $w_{t+1} - w_t - \delta_t = 0$, $E[w_{t+1} - w_t] = E[\delta_t] = \mu_\delta dt$, and $E[(w_{t+1} - w_t)^2] = E[\delta_t^2] = \sigma_\delta dt$. Note also that $R = e^{-rdt} = 1 - rdt + o(dt)$, where r is the continuous rate of interest, and $1 - R = 1 - e^{-rdt} = rdt - o(dt)$. Finally, $\sigma_p = \sigma_p dt$ and $\pi_t = \pi dt$. Substituting these relationships into (4) and ignoring terms of $o(dt)$ gives

$$rV(w_t,L)dt = E\{w_{t+1}\pi dt + \lambda w_{t+1}{}^2\sigma_p dt - \lambda_1 [1-(1-w_{t+1})^2]\sigma_p dt + V_w(w_t,L)\mu_\delta dt + .5V_{ww}(w_t,L)\sigma_\delta dt\}$$

Since $E[w_{t+1}] = w_t + \mu dt$, and $E[w_{t+1}{}^2] = w_t^2 + O(dt)$, and these terms are multiplied by dt, we have (ignoring terms of $o(dt)$)

$$rV(w_t,L)dt = w_t\pi dt + \lambda w_t^2\sigma_p dt - \lambda_1 [1-(1-w_t)^2]\sigma_p dt + V_w(w_t,L)\mu_\delta dt + .5V_{ww}(w_t,L)\sigma_\delta dt.$$

Dividing by dt gives the fundamental differential equation for $V(w_t,L)$ when $w_t \in$ [0, L]:

$$(5) \quad rV(w_t,L) = w_t\pi + \lambda w_t^2\sigma_p - \lambda_1 [1-(1-w_t)^2]\sigma_p + V_w(w_t,L)\mu_\delta + .5V_{ww}(w_t,L)\sigma_\delta$$

We henceforth omit the subscript "t" from w_t, since the solution is time independent. This in turn confirms that L will be independent of t, as previously asserted.

3. Solution to the Problem

For given L, the solution to the ordinary differential equation (5) is

$$V(w,L)=C_1e^{(x-z)w}+C_2e^{(x+z)w}+a_1+a_2w+a_3w^2 \tag{6}$$

where

$$z = \frac{(\mu_\delta^2+2r\sigma_\delta)^{\frac{1}{2}}}{\sigma_\delta}; \qquad x = -\frac{\mu_\delta}{\sigma_\delta}$$

and

$$a_1 = \frac{(\lambda+\lambda_1)\sigma_p}{r^2}\left(\frac{2\mu_\delta^2}{r}+\sigma_\delta\right); \quad a_2 = \frac{2(\lambda+\lambda_1)\sigma_p\mu_\delta}{r^2} + \frac{(\pi-2\lambda_1\sigma_p)}{r}; \quad a_3 = \frac{(\lambda+\lambda_1)\sigma_p}{r}$$

For $w > L$, $V(w,L) = V(L,L) + k(w - L)$; for $w < 0$, $V(w,L) = V(0,L) - kw$.

Boundary conditions, assuring "smooth pasting" at 0 and L, are

$$V_w(0,L) = (x-z)C_1+(x+z)C_2+a_2 = -k \tag{7i}$$

$$V_w(L,L) = (x-z)C_1e^{(x-z)L}+(x+z)C_2e^{(x+z)L}+a_2+2a_3L = k \tag{7ii}$$

where $V_w(w,L) = \partial V(w,L)/\partial w$, $V_L(w,L) = \partial V(w,L)/\partial L$, etc. Given L, it is possible to solve for C_1 and C_2 from the above two equations. We denote these $C_1(L)$ and $C_2(L)$.

Differentiating (7i) and (7ii) with respect to L yields

(8i) $$V_{wL}(0,L) = (x-z)C_{1L}+(x+z)C_{2L} = 0$$

and

(8ii) $$V_{wL}(L,L) = (x-z)^2C_1e^{(x-z)L}+(x+z)^2C_2e^{(x+z)L}+2a_3+(x-z)C_{1L}e^{(x-z)L}+(x+z)C_{2L}e^{(x+z)L} = 0$$

where $C_{1L} = \partial C_1(L)/\partial L$, etc.

To find the optimal L^*, we minimize (6) with respect to L, at arbitrary $w \in [0,L]$. Since only C_1 and C_2 are functions of L, the first order condition is

(9) $$V_L(w,L) = C_{1L}e^{(x-z)w}+C_{2L}e^{(x+z)w} = 0$$

When $w = 0$, equation (9) reduces to

(10) $$C_{1L} + C_{2L} = 0$$

Now (8i) and (10) can hold simultaneously if and only if

(11) $$C_{1L} = C_{2L} = 0$$

at $L = L^*$. But (11) implies equation (9) holds for all $w \in [0,L]$ at $L = L^*$. That is, the same L^* minimizes expected costs for all $w \in [0,L^*]$, not simply at $w = 0$. Condition (11) also allows simplification of (8ii) at $L = L^*$:

(12) $$V_{wL}(L,L) = (x-z)^2C_1e^{(x-z)L} + (x+z)^2C_2e^{(x+z)L} + 2a_3 = 0.$$

Equations (7i), (7ii), and (12) can be solved for the three unknowns C_1, C_2, and $L = L^*$.

4) An Example, with Comparative Statics

Consider an example where

r = .10 (10% annual riskfree rate)

σ_δ = .01 (10% standard deviation in annual cash flows)

μ_δ = 0 (no drift in annual cash flows)

σ_p = .04 (20% standard deviation in returns of the stock portfolio)

k = .01 (1% one-way transactions costs)

π = .10 (the expected return on the stock portfolio exceeds the return on cash by 10%)

λ = 10; λ_1 = 0.

Solving the appropriate equations gives

$$L^* = .04065$$

$$V(w, L^*) = 0.4 + w + 4w^2 - .0635\, e^{-4.472w} - .2894\, e^{4.472w}$$

Figure 1 graphs $V(w, L^*)$ as a function of w. For $w < 0$, $V(w, L^*) = V(0, L^*) - kw$. For $w > L^*$, $V(w, L^*) = V(L^*, L^*) + k(w - L^*)$.

In our example, the correct strategy is not to adjust the cash position as long as it remains between 0% and 4.065% of portfolio value. If it falls beneath zero, enough securities should be sold to return it to zero. If it exceeds 4.065%, enough securities should be sold to reduce the cash proportion to 4.065%. The following will increase the maximum cash fraction L^*:

> Higher transactions costs k.
> Higher volatility of cash flow σ_δ
> Lower expected cash flow μ_δ

> Lower volatility of the stock portfolio σ_p
> Lower risk premium π
> Lower tracking error weight λ

Consider the following separate parametric shifts and their effect on L^* and $V(0,L^*)$, in comparison with the base case where $L^* = .04065$ and $V(0,L^*) = .04713$:

Changed Parameter Value	Optimal L^*	$V(0, L^*)$
$k = .05$	.0840	.1110
$\sigma_\delta = .04$	.0757	.0983
$\mu_\delta = -.05$	.0434	.0458
$\sigma_p = .0225$	.0423	.0461
$\pi = .06$	.0485	.0383
$\lambda = 5$	.0565	.0460

An "emerging markets" portfolio might have $k = .05$, $\sigma_p = .10$, $\sigma_\delta = .04$, $\mu_\delta = -.05$, and $\pi = .20$. (We assume other parameters remain as in the base case). For this set of parameters, $L^* = .1127$ and $V(0, L^*) = .3257$. If L were mistakenly chosen at the same level as the base case (.04065), then $V(0, .04065) = .5381$, implying an expected cost 67% above the optimal level.

We briefly consider the case where $\lambda = 0$ but $\lambda_1 > 0$. This is consistent with a mean-variance optimizer who recognizes that cash lowers expected return, but also risk. The investor has no additional penalty for tracking error. By assumption, the investor chooses $w = 0$ in the absence of transactions costs. That is,

$$0 = \operatorname{argmax}_w \{\pi w - \lambda_1[1-(1-w)^2]\sigma_p\}$$

For this to be true

$$\lambda_1 = \pi/(2\sigma_p).$$

Given the other parameters of the base case, we find

$$L^* = .147.$$

The maximal cash position in this situation is considerably greater than before. This is because there is not a separate penalty for "tracking error." Although $w > 0$ is suboptimal (when transactions costs are zero), increasing w from zero initially has only a small cost, since the fall in expected return is balanced by a fall in overall portfolio risk. Since transactions costs are the same here, the optimum occurs at a greater L. Comparative statics, however, are similar in direction to the earlier case with $\lambda > 0$ and $\lambda_1 = 0$.

5. Expected Turnover of the Optimal Strategy

We have shown above how the interval $[0, L^*]$ is determined. But what are the expected trading costs and turnover of the optimal strategy? There is a "trick" that permits us to answer the question. Note that the differential equation (5) includes two terms associated with the cost of $w > 0$: the expected return loss πw, plus the tracking error w^2 weighted by λ.
The solution $V(w, L)$ gives the *total* cost of the strategy: the expected trading costs, plus the costs associated with $w > 0$.

Now consider a special case of $V(w, L)$, which we shall call $T(w, L)$, when the terms associated with the cost of $w > 0$ are set equal to zero. *This will represent the expected costs of trading alone.* The solution to (6) with the restrictions that $\pi = 0$, $\lambda = 0$ and $\lambda_1 = 0$ will generate $T(w, L)$. The boundary conditions are also the same as (7i) and (7ii), with $\pi = \lambda = \lambda_1 = 0$. From (6), denoting the constants as K_1 and K_2 (to distinguish them from C_1 and C_2):

$$(13) \qquad T(w, L) = K_1 e^{(x-z)w} + K_2 e^{(x+z)w},$$

with boundary conditions

(14')
$$(x\text{-}z)K_1 + (x+z)K_2 = -k$$

(14")
$$(x\text{-}z)K_1e^{(x-z)L} + (x+z)K_2e^{(x+z)L} = k.$$

Given the L^* determined by the solution to the original problem, we can use these boundary conditions to solve for the constants K_1 and K_2, and therefore for the function $T(w, L^*)$.

For the example considered in Section 4 above, with $L = .04065$, we find

$$T(w, .04065) = .01345\ e^{-4.472w} + .01122\ e^{4.472w}.$$

$T(w, L)$ represents total expected discounted transactions costs over the infinite time horizon. When $w = 0$, $T(0, .04065) = .02467$. Annual transactions costs will simply be $rT(w,L)$, and annual expected one-way turnover is $rT(w, L)/k$. This is graphed in **Figure 2**, for the base case. Depending on the current level of w, annual turnover is expected to be about 24.6%. It can be seen that turnover costs $T(w, L)$ are slightly more than half the total costs $V(w,L)$ in this case.

In contrast, the "emerging market" scenario described at the end of Section 4 incurs a cost T (when $w = 0$) of about 0.179, or .0179 annually. With transactions costs of 5%, this converts to an approximate annual turnover of about 35.9%. Note that although the variance of cash flows is much larger in comparison with the base case ($\sigma_\delta = .04$ vs. .01), the turnover is less than proportionately greater, because the "no trade" interval is almost three times as great ($L^* = .1127$ vs. .04065).

Figure 3 shows how turnover $T(0, L)$ depends upon L for base case parameters.[5] A good approximation for $T(0, L)$ is

(15) $$T(0, L) = \sigma_\delta m(.5 + \frac{1}{e^{mL} - 1})$$

where

$$m = \frac{(4\mu_\delta^2 + 2r\sigma_\delta)^{1/2}}{\sigma_\delta}$$

This is an exact expression when $\mu_\delta = 0$, in which case $m = (2r/\sigma_\delta)^{1/2}$. The expected turnover for $w \in (0,L]$ will generally be different from the expected turnover when $w = 0$; however, the difference will be quite small when $L < 0.15$.

6. Extensions

The model is easily extended to a minimum cash boundary other than zero. For example, say the optimal portfolio has a 3% cash position. Then the cash position associated with the cash inventory fall to -3%, and still not require borrowing outside the fund. Such a change could readily be incorporated in the analysis via a change in boundary conditions, as could a requirement that cash balances (say) never be less than 2%, rather than 0.

The existence of a fixed transactions cost, in addition to a percentage transactions cost, would alter the nature of the optimal policy. Now two nested intervals must be described: $[0, L^*]$, in which no trading takes place, and $[L_1, L_2]$, the cash positions to which one trades when trading is triggered, with $0 < L_1 < L_2 < L^*$. The derivation of the optimal intervals remains for the future.

---

[5] This is one-way (cash to securities or vice-versa) transactions. Often turnover is measured in "roundtrip" terms (securities to securities), in which case our turnover numbers would be divided by two.

7. Conclusions

We have derived precise answers for the optimal cash holding of an investment fund which has cumulative cash inflows and outflows described by a random walk (perhaps with a trend). By assumption, the fund desires cash only for transactions purposes. Cash may be generated by dividends and contributions, and may be demanded for rights offerings and redemptions. This cash would be instantly reinvested or created by sales of securities, were it not for transactions costs which would make such continuous reinvestment exceedingly costly.

The nature of the optimal cash policy is to allow cash to accumulate (or fall) randomly, with no selling or buying of securities, as long as the cash fraction is greater than or equal to zero, and less than or equal to an upper limit L^*. If the cash fraction falls beneath zero, securities must be sold in sufficient amount to return the cash position to zero. If the cash fraction exceeds L^*, the cash must be invested in securities such as to return the cash position to L^*.

A description of the optimal cash strategy therefore devolves into the determination of L*. This paper has developed a set of equations whose solution determines L^*, and relates it to transactions costs, the mean and variance of random cash flow process, the riskiness of investment portfolio, and the investor's tradeoff between trading costs and tracking error. A simple extension of the analysis allows prediction of the turnover associated with the optimal strategy.

REFERENCES

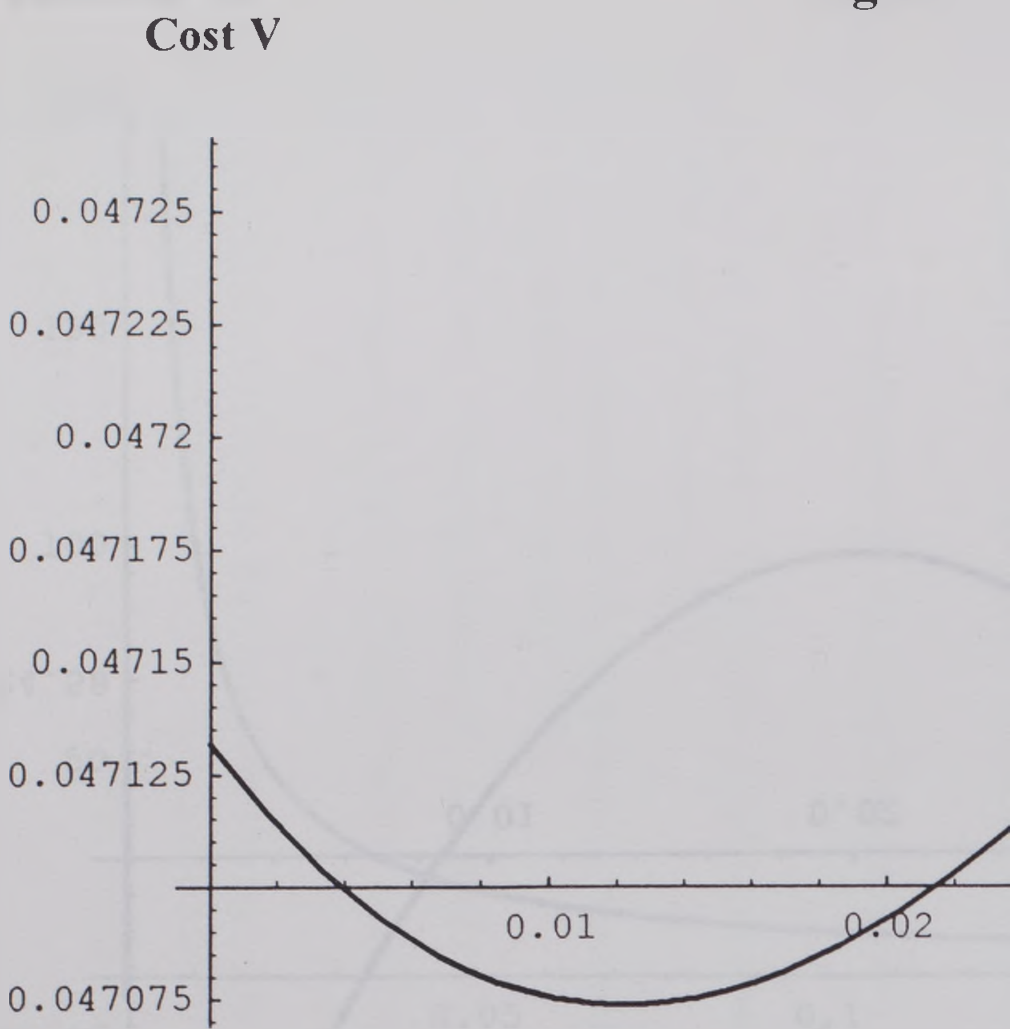
Constantinides, G. [1986], "Capital Market Equilibrium with Transactions Costs," *Journal of Political Economy* **94**.

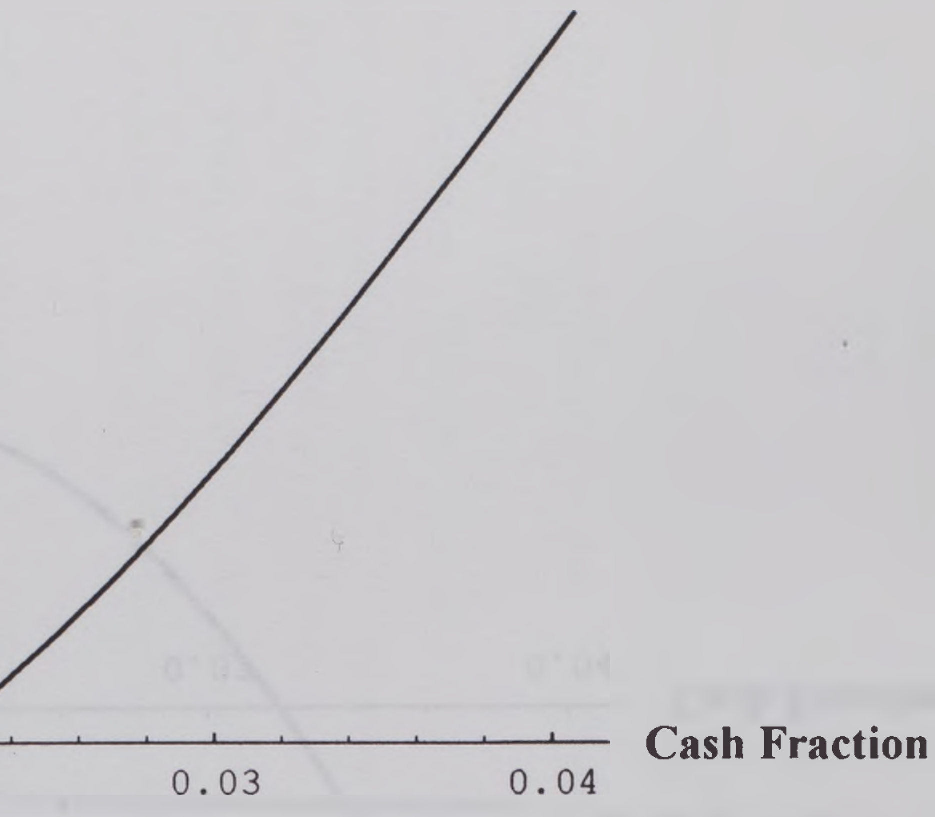
Dumas, B., and Luciano, E. [1991], "An Exact Solution to a Dynamic Portfolio Choice Problem Under Transactions Costs," *Journal of Finance* **46** (June), 577-96.

Hodges, S., and Neuberger, A. [1989], "Optimal Replication of Contingent Claims Under Transactions Costs", *RFMT* **8**, 222-239.

Magill, M., and Constantinides, G. [1979], "Portfolio Selection with Transactions Costs," *Journal of Political Economy* **87** (August), 673-700.

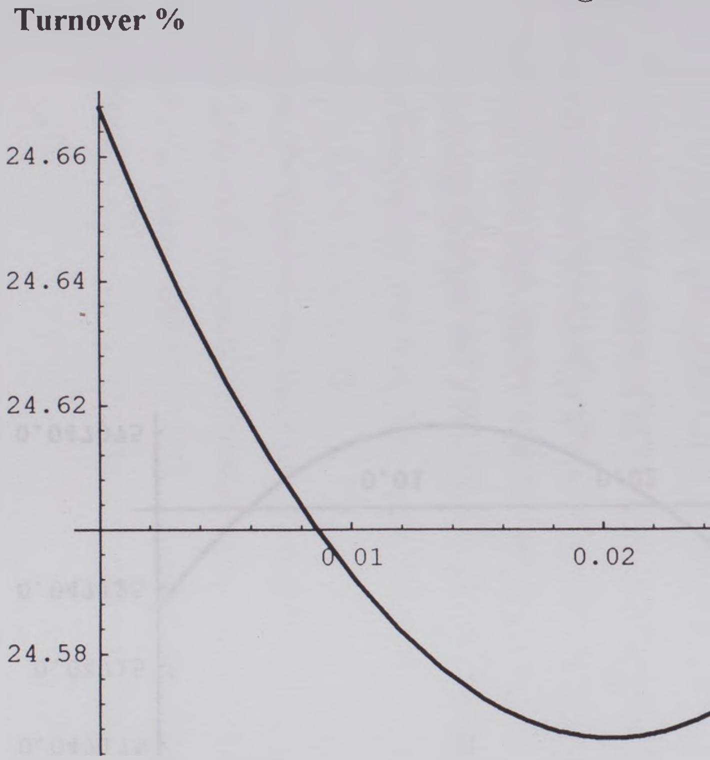
Figure 1


Cost V
0.04725
0.047225
0.0472
0.047175
0.04715
0.047125
0.047075
0.01
0.02



0.03
0.04
Cash Fraction

Figure 2



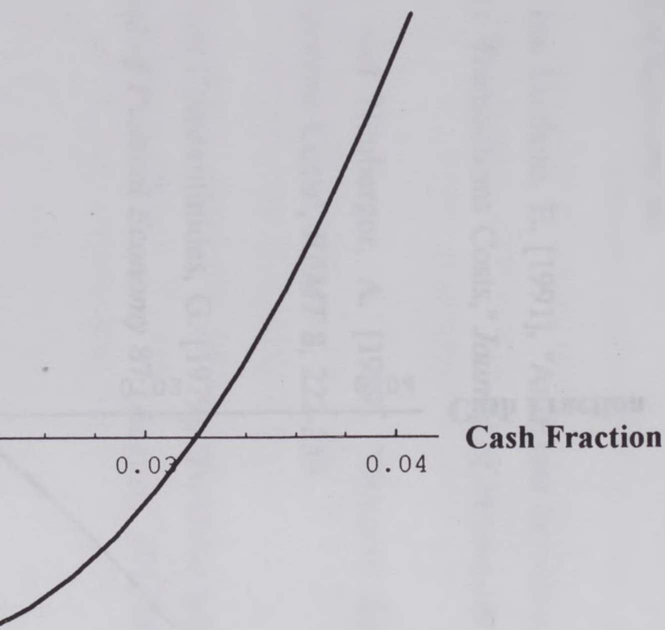
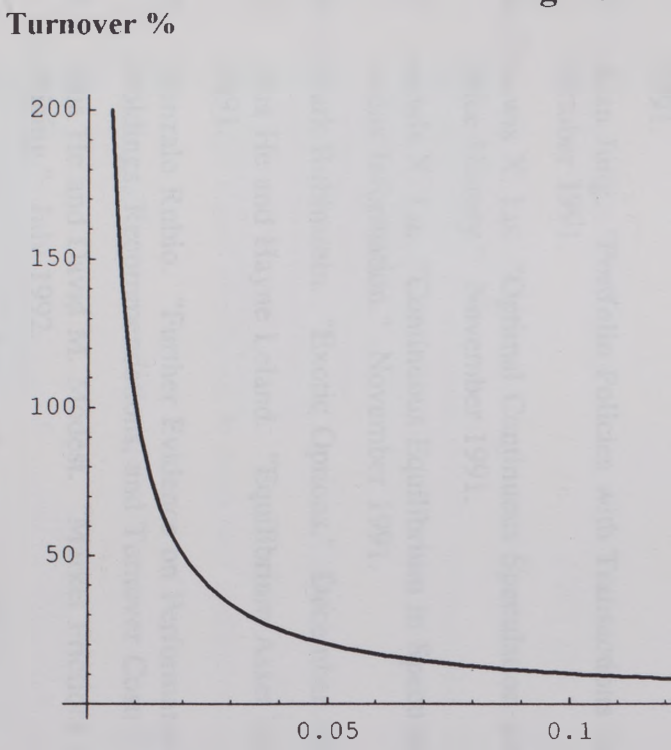

0.03
0.04
Cash Fraction

Figure 3



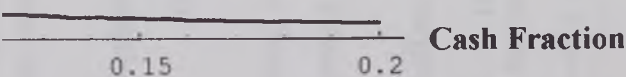
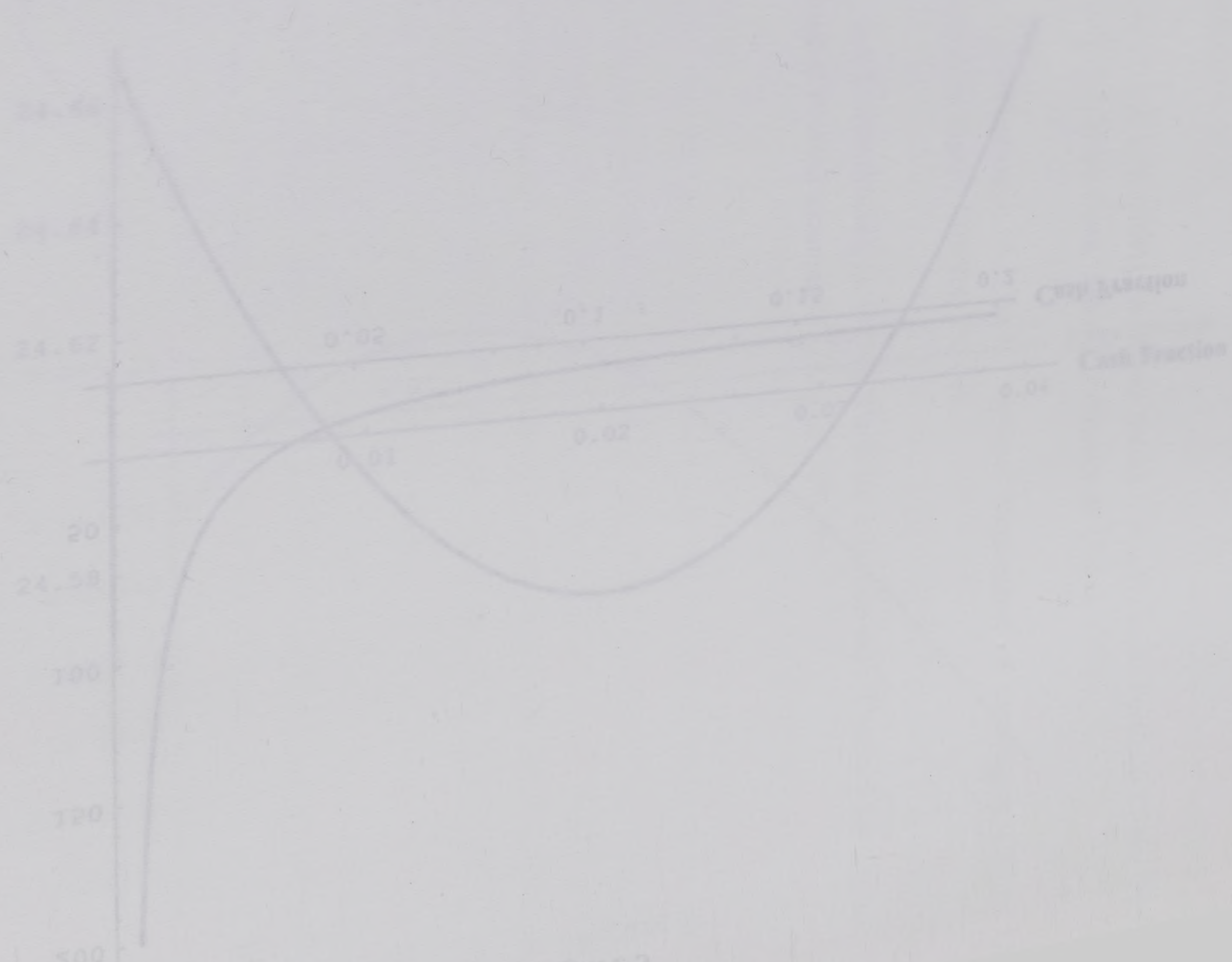

0.15
0.2
Cash Fraction

Figure 2



RESEARCH PROGRAM IN FINANCE WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" -- by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

210. Gerard Gennotte and Terry A. Marsh. "Variations in Economic Uncertainty and Risk Premiums on Capital Assets." May 1991.

211. Gerard Gennotte and Hayne Leland. "Low Margins, Derivative Securities, and Volatility." May 1991.

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

244. Hayne Leland and Gregory Connor. "Optimal Cash Management for Investment Funds." March 1995.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance
WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 245

On Revelation of Private Information in Stock Market Economies

by

Marcus Berliant

and

Sankar De

April 1995

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE
WALTER A. HAAS SCHOOL OF BUSINESS,
UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH
Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

On Revelation of Private Information in Stock Market Economies

Marcus Berliant
and
Sankar De

April 1995

Finance Working Paper #245

The authors would like to thank Dee Dechert, Jim Friedman, David Gordon, Frank Page, and Paul Romer for helpful comments. Any errors remaining in the paper are, however, the authors' responsibility.

Abstract

The notion that an agent in a given market can infer from the market price the (non-price) information received by other agents, as embodied in the existing studies of revealing rational expectations equilibrium, requires that the agent know the correct functional relationship between the non-price information of all agents and the resulting equilibrium price. This condition is usually restrictive and unsuitable as a description of reality. In this paper we show that this condition is also *unnecessary* in a stock market economy where producers or firms use their private information in their own optimization programs, which include stock purchases. Interestingly, this result does not extend to the case of consumers with private information.

Marcus Berliant
Department of Economics
Washington University, Campus Box 1208
St. Louis, MO 63130-4899
and
Department of Economics
University of Rochester
Rochester, NY 14627-0156

Sankar De
Indian Institute of
Management
Calcutta, India
and
Haas School of Business
University of California
Berkeley, CA 94720-1900

1. Introduction

The notion that an economic agent in a given market can use the equilibrium market price to make inferences about the (non-price) information received by other agents regarding the exogenous states of the environment, as embodied in models of rational expectations equilibrium in markets with differential information, requires that the agent know the correct functional relationship between the non-price information received by all market participants and the resulting equilibrium price (see, for example, Grossman (1978, 1981), Kihlstrom and Mirman (1975), and Radner (1979)). Given that the information-price function is known, the question whether an individual agent can glean the initial information of all agents from market prices, that is whether the information-price function is invertible on the set of admissible prices, assumes significance. This question is the focal point in the literature on fully revealing rational expectations equilibria. Yet the condition that lends significance to it, namely that the true relationship between the information signals of all agents taken together and the market prices is known to all, has been relatively little examined. However, this condition is usually very restrictive. Two conditions that are each sufficient (but not necessary) for knowledge of this relationship are the following. First, each agent could learn this relation over time; such a justification requires a model of learning such that convergence to the appropriate relation is assured. Second, each agent could have a good deal of information about the private characteristics and plans of other agents, along with a great computational ability. In any case, each agent must know the information-price relation, and this relation could be very complex.

Interestingly, this condition is unnecessary in one important situation. We show that, in a stock market economy characterized by producers with private information which they use in their own decision-making process, equilibrium stock prices reveal all market-relevant information under very general conditions even though the agents may not know the correct information-price function. We present formal models in Sections

II and III of this paper to demonstrate this result. The process through which this comes about can broadly be sketched as follows. The producers or the firms, acting on the basis of their private information, decide to own fractions of other firms and, by retaining shares, of themselves as well. Given certain general conditions, the profits from these activities based on their private information are in equilibrium a linear function of the stock prices and other publicly observable market data. As a result, the private calculations that go into their decision–making are of no consequence to the uninformed agents. This is the main result of our study. Note that this result does not extend to the case where consumers have private information about individual endowments or wealth since, unlike firms, they cannot own each other or issue claims against each other in a standard market setting. This creates an interesting asymmetry between consumers with private information and producers with private information in a stock market economy.

In Section II of this paper, we present a formal model to demonstrate our main result. In Section III we extend our model to a general equilibrium framework and show that the result holds generically in this setting. Section IV presents some implications of our study and our conclusions.

II. A Partial Equilibrium Model of Revelation When the Information–Price Relationship is Unknown

In this section we outline an optimization model for producers with private information in a stock market economy, and discuss a necessary condition for equilibrium that results in equilibrium prices, along with other observable market data, revealing all relevant information about the firms even though their private information as well as the function mapping that information into prices remain unknown.

Suppose there are n firms (n integer and finite) indexed by i and j. For simplicity, suppose that firms issue only one kind of security, namely stock. Further, suppose each firm issues only one unit of stock (through normalization). Firms can own not only their stock, but the stock issued by other firms as well. We define y^i_j to be the fraction of firm j's stock owned by firm i, a choice variable for the manager of firm i. Similarly, the variable y^i_i represents the fraction of its own stock retained by firm i in its vaults. It is the quantity the firm does not sell to anybody including its own employees. Of course, it is possible that $y^i_i = 0$. Let $Y^i \equiv (y^i_1,...,y^i_n)$. Further, let

$$Y \equiv \begin{bmatrix} y^1_1 & \cdots & y^1_n \\ \vdots & & \vdots \\ y^n_1 & \cdots & y^n_n \end{bmatrix}$$

The model has two dates in the sense that the managers of all firms in the market make their decisions about the activities of their firms at date 1. In other words, each manager solves his optimization problem at date 1. At date 1, there is uncertainty concerning the returns to the various activities. The uncertainty arises due to the stochastic nature of each firm's production set. This uncertainty is resolved only at date 2 when the firms realize the returns on their investments.

Formally, we proceed as follows. There are k physical commodities and n firms. The prices for the economy are denoted by (P,Q), where $Q = [Q_1,...,Q_k]'$ are prices for the physical commodities and $P = [P_1,...,P_n]'$ are stock prices. Prices (P,Q) lie in the simplex $\Delta \equiv \{p = (P,Q) \in \mathbb{R}^{n+k} \mid P_i \geq 0$ for all i, $Q_j \geq 0$ for all j, and $\sum_{i=1}^{n} P_i + \sum_{j=1}^{k} Q_j = 1\}$. At date 1, the decision-making problem of the manager of firm i involves two vectors of choice variables. First, he chooses y^i_j for j = 1, 2, ..., n, where y^i_j, as mentioned above, denotes the fraction of firm j's stock owned by firm i. Second, he chooses Z^i, where Z^i is defined to be the vector of inputs and outputs of the firm's production process (inputs negative, outputs positive). Let e^i_j be the initial endowment of firm j's stock held by firm i, let $e_j \equiv \sum_{i=1}^{n} e_j^i$, and let $e \equiv [e_1,...,e_n]'$. Notice that $e_j \leq 1$ for each j. In the partial equilibrium model discussed in this section, we neglect discussion of the consumer sector, which could also have initial endowments of stock as well as stock ownership in equilibrium. We define

$$E \equiv \begin{bmatrix} e_1^1 & \cdots & e_n^1 \\ \vdots & & \vdots \\ e_1^n & \cdots & e_n^n \end{bmatrix}$$

The information available to the manager of the firm i at date 1 when he makes his investment decisions is now described. First, the market value of one unit of firm j's stock P_j, for j = 1, 2, ..., n, is known to the manager of firm i. It is not necessary for the first part of our study that the stock market be perfectly competitive; it is simply assumed that the manager of firm i knows the price schedule that he faces. Our structure is general enough to include imperfectly competitive markets where some prices could be choice variables for the manager of firm i. Next, it is assumed that the matrix of stock purchases by all other firms,

$$\begin{bmatrix} y_1^1 & y_2^1 & \cdots & y_n^1 \\ \vdots & & & \vdots \\ y_1^{i-1} & y_2^{i-1} & \cdots & y_n^{i-1} \\ y_1^{i+1} & y_2^{i+1} & \cdots & y_n^{i+1} \\ \vdots & & & \vdots \\ y_1^n & y_2^n & \cdots & y_n^n \end{bmatrix}$$

as well as the matrix of initial stockholdings represented by E are known to the manager of firm i when he solves his maximization problem. Finally, the manager of firm i observes an exogenous random vector ϵ^i that represents his personal information about the alternative states of nature. Let $\epsilon \equiv (\epsilon^1,...,\epsilon^n)$. Further, let $T^i(\epsilon^i) \subseteq \mathbb{R}^k$ be the technology available to the manager of firm i at date 1. Since ϵ^i is known only to the manager of firm i, $T^i(\epsilon^i)$ is private information as well. Such private information about a firm's technology available only to the firm's manager could arise, for example, from the outcome of research and development expenditures. The form of private information assumed in this model is reasonable and general.

The objective function of the manager of firm i is now introduced. Of course, it is unclear what firms maximize under incomplete markets (e.g. expected profits), and we abstract from this issue by simply assuming that there is an objective function. We now define the function V^i $(Y^i; Z^i; (P,Q))$ to be manager i's objective function. The value of the function depends on the stock purchases of firm i, its production activities begun at date 1, and market prices. Note that the manager of firm i does not necessarily know Z^j, $j \neq i$. Note that the function V^i is, by its construction, very general in nature.[1] It does not necessarily represent the expected value of the firm's uncertain return distribution at date 2 conditional on the given arguments of the function, though we could, of course, use it to mean just that.

Finally, we assume that the markets for securities are perfect; in other words, there are no taxes, transaction costs, etc.

The consumer sector is peripheral to the results in this section of the paper; it will be detailed in the next section, where it is needed to complete the model for general equilibrium analysis.

---

[1]It is possible to let V^i depend not only on Y^i but on the entire matrix of stock purchases Y, so that $V^i = V^i(Y; Z^i;(P,Q))$. This would allow V^i to take account of the possible covariances between the return vectors of firms to a greater degree than the formulation in the main body of the text permits. However, this further generalization does not alter any of our results and is, therefore, unnecessary for our exposition.

Given $(P,Q) \in \Delta$, Y^i, and Z^i, we write firm i's profits or returns per unit of stock at date 2 as π_i. Let $\pi \equiv [\pi_1,\ldots,\pi_n]'$. With the notations and the set of assumptions as described above, the manager of firm i faces the following problem at date 1 (an explanation follows immediately):

(1) (a) $\max\limits_{\substack{Y^i \\ Z^i \\ \text{(and possibly } (P,Q))}} \quad V^i(Y^i;\ Z^i;\ (P,Q))$

subject to

(b) $Z^i \in T^i(\epsilon^i)$

(c) $\pi_j \equiv \sum_{h=1}^{n} [y_h^j \cdot (\pi_h - P_h) + e_h^j \cdot P_h]$ for $j = 1,\ldots,n$.

The significance of the objective function itself is obvious. Apart from making production decisions, the manager of firm i must choose which stocks to buy for his company in order to maximize his objective. Note that by purchasing stock of other firms, the manager changes the return vector of his own firm. Further, in an imperfectly competitive market, (P,Q) can be a decision variable for the manager.

Expression 1(b) represents the technological constraints on the firm. It simply says that input–output combinations must be in a set that depends on the realization of the random variable, where the latter is private information. As a consequence, both Z^i and the realized production set are also private information.

Expression 1(c) is a set of consistency conditions on profits or returns at date 2. Given the assumption that Y^j $(j \neq i)$, the matrix E, and the prices (P,Q) are known to the manager of firm i at date 1, condition 1(c) has a natural interpretation. Recall that both P_j and π_j are expressed in dollars per unit of stock. Expression 1(c) means that at date 2 the total profit of firm j is the sum of: (i) the profits earned from speculative as well as production activities, $\sum_{h=1}^{n} y_h^j \cdot (\pi_h - P_h)$, which represents stock purchases[2] multiplied by per–share net profit (that is, final worth less cost) of the concerned firms; and (ii) $\sum_{h=1}^{n} e_h^j \cdot P_h$, the value of the stock that firm j is endowed with. This is a simple accounting identity that holds in equilibrium no matter what state of the world is realized. It is an *ex post* condition. Note that implicit in it is the production process of firm j, in that both Z^j and ϵ^j help determine equilibrium values of P_j and π_j, j = 1,...,n.

A very important point that we wish to emphasize here is that the firm manager believes that ***in equilibrium*** the consistency conditions 1(c) will be satisfied. Out of equilibrium, the manager has no reason to believe that they will be satisfied. Thus, they are a property of an anticipated equilibrium, not of demand. This is similar in spirit to the use of the information–price relationship by agents in a rational expectations model, where the relationship is anticipated to hold in equilibrium, but the relationship is used by agents to obtain information employed in formulating demand. A consequence of this condition, reflected in the results derived below, is that ***in equilibrium*** a firm's manager cannot believe that the profit or returns to his firm is higher or lower than a linear function of the equilibrium share prices.

If the constraint set is compact and V^i is continuous, then a solution to problem (1) exists. In fact, under the usual conditions, an explicit solution to problem (1) can

---

[2]The securities of firm j owned by firm i is actually a series

$$y_j^i + \sum_{\substack{k=1 \\ k \neq i \\ k \neq j}}^{n} y_k^i \, y_j^k + \sum_{\substack{k=1, \ell=1 \\ k \neq 1, \ell \neq i \\ k \neq j, \ell \neq j}}^{n} \; \sum^{n} y_k^i \, y_\ell^k \, y_j^\ell + \ldots$$

This accounts for both direct and indirect ownership. It is implicit in the interdependent set of linear equations 1(c).

be found using Lagrangean methods. However, that is not the purpose of this paper.

Define I to be the nxn identity matrix.

Theorem 1: Let Y be an equilibrium matrix of firm shareholdings, and suppose that I–Y is invertible. Then it must be true that equilibrium prices and profits satisfy:

$$\pi = [I-Y]^{-1}\cdot[E-Y]\cdot P. \qquad (2)$$

Remark. This is a partial equilibrium result, since we use only condition 1(c) (which holds in equilibrium) but not market clearing conditions, and the consumer sector is absent. *Any agent who knows* Y, E, *and* P *can figure out* π *without knowledge of the relationship between private information,* ϵ, *and prices*, P.

Proof: In vector form, 1(c) can be written as $\pi = Y\cdot[\pi-P] + E\cdot P$. Hence $[I-Y]\cdot\pi = [E-Y]\cdot P$, and since [I–Y] is assumed to be invertible, $\pi = [I-Y]^{-1}\cdot[E-Y]\cdot P$.

Q.E.D.

Corollary 1: Suppose that each firm is endowed with all of its own stock, E = I, and that Y is an equilibrium matrix of firm shareholdings with I–Y invertible. Then a necessary condition for an equilibrium is that $\pi = P$.

Proof: A trivial application of Theorem 1.

Corollary 2: Suppose that Y is an equilibrium matrix of firm shareholdings with Y non–negative. Suppose further that for each firm, in equilibrium there exists a consumer who owns a fraction of that firm: $\sum_i y_j^i < 1$ for j = 1, 2, ..., n. Then a necessary condition for an equilibrium is that $\pi = [I-Y]^{-1}\cdot[E-Y]\cdot P$.

Proof: Suppose that I–Y (where I is the nxn identity matrix) is not invertible. Then 1 is a characteristic root of Y.

Let $s = \max_j \sum_{i=1}^{n} y_j^i$. By assumption, $s < 1$. Recall that each element of Y is non–negative. Let r be the maximal characteristic root of Y, and let x be a column vector of n ones. Then, if the matrix Y′ represents the transpose of Y, $Y' x \leq s \cdot x$. By Debreu and Herstein (1953, Lemma*), since every element of x is strictly positive, $r \leq s < 1$. Hence 1 cannot be a characteristic root of Y; the hypothesis is false and I–Y is invertible.

The remainder of the proof follows from Theorem 1.

Q.E.D.

The results above indicate that if each manager faces problem (1), then stock prices in conjunction with stock purchases as well as initial holdings, which are assumed to be known to all, are sufficient to infer date 2 returns or profits anticipated by the managers acting on the basis of their private information. In the case where a firm is endowed with all of its own stock, the profits of a firm as seen by its manager are the same as their market prices.

Intuitively, the results indicate the following. If the maximization programs of the managers of the firms are executed in a consistent manner (i.e., one that satisfies 1(c)), managerial private information is of no consequence to the uninformed agents, since the equilibrium values of all relevant variables must be consistent with publicly observable data.

Of course, it is not always the case that the main assumption of Theorem 1, that I–Y is invertible, is satisfied. Hence, we show next that this assumption holds in a generic sense.

In a natural extension of the model discussed above, there would be an infinite number (continuum) of producers and hence an infinite number (continuum) of securities in a perfectly competitive stock market. The natural generalization of the matrix Y is a linear operator Y* from the space of stock prices (or present values) into itself that is known to all managers. The analog of the invertibility property of the matrix I–Y is the invertibility property of I–Y*, where I is the identity operator; see Rudin (1973, p. 98).

III. A General Equilibrium Model of Revelation When the Information–Price Relationship is Unknown

In this section, we shall complete the partial model of the previous section by adding markets for goods and a consumer sector. Properties of this general equilibrium extension of the partial equilibrium model are then examined as follows. A "completely revealing" concept of competitive equilibrium is defined, where prices reveal all market–relevant private information in the economy. Theorem 2 shows that a hypothesis of Corollary 2 holds generically. Theorem 3 demonstrates the existence of an equilibrium in the model with no uncertainty. The main result of the paper, Theorem 4, demonstrates the generic existence of a fully revealing equilibrium, and follows directly from Theorems 2 and 3. Finally, we examine the welfare properties of a fully revealing equilibrium allocation in Theorem 5.

Since the focus of this paper is clearly not on the consumer sector, we shall abbreviate the model. Naturally, in this section we assume competitive behavior on the part of all agents.

The consumption space for the model is $\mathbb{R}_+^{n+k}$, where the first n components represent the securities. Prices reside in the n + k − 1 dimensional simplex Δ. There are m consumers, i=1,...,m, where consumer i has a continuous, quasi–concave utility function $u^i : \mathbb{R}_+^{n+k} \to \mathbb{R}$. The *ex post* utility of a consumer can depend on the realizations of the random variables ϵ^i, but since only an *ex ante* equilibrium (for the consumers) is obtained, utility is not directly a function of the realizations. For example, consumer i could be an expected utility maximizer. One interpretation of this framework is that u^i is a reduced form utility that only depends on security ownership through a budget constraint. Consumer i has endowment $\omega^i \in \mathbb{R}_+^{n+k}$. Hence, consumer i is endowed with ω_j^i units of the stock of firm j (j = 1,...,n). Let $\omega_j \equiv \sum_{i=1}^m \omega_j^i$, and note that $\omega_j + e_j = 1$ for all j. We shall parameterize economies by $\omega = (\omega^1,...,\omega^m) \in \Omega \equiv \{\omega \in (\mathbb{R}_+^{n+k})^m \mid \omega_j = 1 - e_j \text{ for } j = 1,...,n\}$, where $0 \leq e_j \leq 1$ for $j = 1,...,n$. We say that a property is *generic* if it holds everywhere on Ω except for a measurable subset of Lebesgue measure zero.

Let the demand correspondence for consumer i, $f^i: \Delta x \mathbb{R}_+^{n+k} \to \mathbb{R}_+^{n+k}$ be defined by $f^i(p,\omega^i) \equiv \{x \in \Omega \mid p\cdot x \leq p\cdot \omega^i \text{ and } u^i(x) \geq u^i(z) \text{ for all } z \in \mathbb{R}_+^{n+k} \text{ with } p\cdot z \leq p\cdot\omega^i\}$. Define the aggregate excess demand correspondence by $f(p,\omega) \equiv \sum_{i=1}^m [f^i(p,\omega^i)-\omega^i]$.

On the production side, we assume that shareholdings must be non-negative, and that prices are parametric. Define the supply correspondence for firm i by $\beta^i(p,\epsilon^i) \equiv \{(Y^i,Z^i) \in \mathbb{R}_+^n x T^i(\epsilon^i) \mid (Y^i,Z^i) \text{ solves (1)}\}$. Let $\hat{e}$ be the vector with e in the first n places and 0 in the next k places. The aggregate excess supply correspondence is defined by $\beta(p,\epsilon) \equiv \sum_{i=1}^n \beta^i(p,\epsilon^i) - \hat{e}$.

An *equilibrium price* given ω and ϵ is a $p \in \Delta$ such that $0 \in \beta(p,\epsilon) + f(p,\omega)$.

<u>Theorem 2:</u> Suppose that $f^i(\cdot)$ and $\beta^i(\cdot)$ are single-valued, while $f^i(\cdot)$ and $\beta^i(\cdot,\epsilon^i)$ are continuously differentiable (the latter for each given ϵ^i). For each realization of the random vector ϵ, the property that $\pi = [I-Y]^{-1}\cdot[E-Y]\cdot P$ when evaluated at an equilibrium price holds generically in consumer endowments.

Remarks: 1) By assuming that the excess demand and supply correspondences are functions and C^1, we abstract from the problems associated with deriving these properties from primitives, as this is not our focus and would distract us from the main objective of this work. (Of course, this is in the tradition of Debreu (1970).) For indications of how this derivation could be accomplished, we refer to Mas–Colell (1985). We should note, however, that the structure of production employed here differs substantially from classical production theory.

2) Notice also that it might be possible to parameterize economies by producer endowments of stock. There are two drawbacks to this alternative approach. First, it would probably require a more complicated technical argument (as well as further assumptions) to show that this parameterization is regular, since producers do not have budget constraints. Second, genericity in this sort of parameterization might exclude a case of interest: the case when each producer is endowed with all of its own stock.

3) In the example where the utility of a consumer depends only on physical commodity consumption while asset holdings only affect the budget, it might be thought that asset demand might not be single–valued, particularly in the case where there is no uncertainty, so stocks are used only to store value. This problem can be solved by simply assuming "artificial" preferences over assets, for instance using an additively separable form for a global utility combining the sum of the true utility over physical commodities with the artificial utility over assets. The artificial utility would have to satisfy the usual smooth economy assumptions (for instance, CES) so that the global utility would have the desired properties, implying that demand is single–valued and C^1.

4) Of course, when ϵ takes on only finitely many values, the conclusion of the Theorem holds generically in consumer endowments for *all* ϵ. This notion is closer to the standard concept of completely revealing equilibrium found in the literature,

where economies are parameterized by utility functions.

Proof:[3] Fix ϵ. That Ω is a regular parameterization is proved in Mas–Colell (1985, p. 227); the proof for the "Edgeworth box" economy there covers the case considered here. Mas–Colell (1985, Proposition 8.3.1, p. 320) implies that except for a set $B \subseteq \Omega$ of Lebesgue measure zero, $\partial f/\partial p + \partial\beta/\partial p$ has 0 as a regular value. By Mas–Colell (1975, Proposition 5.8.13, p. 229), the collection of $(p,\omega) \in \Delta x \Omega$ such that $\beta(p,\epsilon) + f(p,\omega) = 0$ for the first n+k–1 equations is a C^1 manifold of the same dimension as Ω. Fix $\omega \in \Omega$, $\omega \notin B$, and let p be an equilibrium price for ω. Suppose that $\sum_{i=1}^{n} y_j^i = 1$ holds in equilibrium for some j. Then $f_j(p,\omega) + \omega_j = 0$, where the subscript j refers to stock j (j = 1,...,n) in the vector belonging to $\mathbb{R}^{n+k-1}$. Suppose further that $\partial y_j/\partial\omega \neq 0$. Hence, along the directions defined by the manifold, which is given locally by $\partial p(\omega)/\partial\omega \equiv [\partial f/\partial p + \partial\beta/\partial p]^{-1}[\partial f/\partial\omega]$, there is a direction of parameter ω movement, a consequent direction of price movement, and a resulting direction of y_j movement such

---

[3]It should be noted that there are some technical complications in this structure relative to classical economies, mainly because it is not assumed that Walras' law holds in aggregate due to the unusual maximization problem faced by producers. In this theorem, the standard technique for renormalizing prices (say, the price of the last commodity is 1), and eliminating both a price and a market clearing condition to account for the redundancy embedded in Walras' law might not work. However, we can eliminate a price and an equation regardless, as we are only looking for a necessary condition for an equilibrium. Thus, we set $p_{n+k} = 1$ (where prices are no longer in the simplex), and eliminate commodity n+k from our calculations. Thus, we do not deal with an equilibrium manifold, but rather a manifold for which the first n+k–1 markets clear.

that nearby manifold holdings of the stock of firm j by all firms exceeds 1. This is a contradiction, as the manifold is defined by market clearance in all stocks. So $\partial y_j/\partial\omega = 0$ for all j, and hence $\partial\beta_j/\partial\omega = 0$ for all j. Thus, $\partial\beta/\partial p = 0$ for directions of price movement along the manifold. Moreover, since the consumers own no stock in firm j at this equilibrium and market clearance is maintained in all stocks on the manifold, in the same manner as for the firms, it must be the case that $\partial f_j/\partial p = 0$ along the directions defined by the manifold as well. Hence $\partial f/\partial p + \partial\beta/\partial p$ is singular at the equilibrium, a contradiction. So for every $\omega \in \omega$, $\omega \notin B$, $\sum_{i=1}^{n} y_j < 1$ for all $j = 1,\ldots,n$. The remainder of the Theorem follows from Corollary 2.

Q.E.D.

In order to prove that an equilibrium exists, it is necessary to go back to the primitives rather than relying on assumptions about supply and demand. This is due to the fact that Walras' law might fail in aggregate (since producers have no budget constraint) for prices that are not equilibrium prices. We now assume that u^i is locally non – satiated, i.e. $\forall\ x \in \mathbb{R}_{+}^{n+k}$ and $\forall\ \delta > 0$ there exists $x' \in \mathbb{R}_{+}^{n+k}$, $\|x - x'\| < \delta$ with $u^i(x') > u^i(x)$. The endowment of consumer i is given by $\omega^i \in \mathbb{R}_{++}^{n+k}$ (a suitable irreducibility assumption could be used in place of the interiority assumption if, for example, one wishes to endow each firm with all of its own stock). A production sector is a collection of n firms, $i = 1,\ldots,n$. Each firm i has continuous function V^i as specified in equation (1). Let V_i be quasi–concave in $(Y^i;Z^i)$ for fixed values of its other arguments. Assume that T^i is closed and convex with $0 \in T^i(\epsilon^i)$ for each i and ϵ^i. In this section, we assume perfect competition, so each firm takes prices as given. The manager of firm i solves (1) taking prices and the stock purchases of other firms as given. Let π_i be the profit function for firm i, $\pi_i(Y^i;Z^i;p) \equiv p\cdot(y_1^i-e_1^i,\ldots,y_n^i-e_n^i;Z^i)$. The crucial assumption concerning V^i is that $\pi_i(Y^i;Z^i;P) < 0$ implies $V^i(Y^i;Z^i;\epsilon^i;p) <$

$V^i(e^i_1,\ldots,e^i_n;0;p)$. That is, if profits are negative according to market prices, then a firm can do better by inaction.

Following Shafer and Sonnenschein (1975), a *competitive equilibrium* is $(p;x^1,\ldots,x^m;Y;Z^1,\ldots,Z^1) \in \Delta x(\mathbb{R}^{n+K}_+)^m x[0,1]^{n^2} x\mathbb{R}^{kn}$ such that $\sum_{i=1}^m [x^i - \omega^i] \leq \sum_{j=1}^n (Z_j; \sum_{i=1}^n [e^i_j - y^i_j])$, $p\cdot x^i = p\cdot\omega^i$ for $i = 1,\ldots,m$, $p\cdot x \leq p\cdot\omega^i$ implies $u^i(x) \leq u^i(x^i)$ for $i = 1,\ldots,m$, and for each $j = 1,\ldots,n$ for any private and public information observed by agent j, (Y^j,Z^j) solves (1).

<u>Theorem 3:</u> There exists a general equilibrium for a stock market model.

<u>Proof:</u> We apply Shafer and Sonnenschein (1975), where the stock purchases of firms act similar to externalities in the model. Using classical techniques, such as those in Arrow and Debreu (1954), we can bound the consumption and production sets without loss of generality. As in these papers, we add another agent, the market player, who has prices Δ as a choice set and the objective of maximizing the value of excess demand. It is easy to verify that, given our assumptions in this section, there exists an equilibrium in the sense of Shafer–Sonnenschein. To show that this equilibrium is also a competitive equilibrium, it suffices to check that $\sum_{i=1}^m [x^i - \omega^i] \leq \sum_{j=1}^n (\sum_{i=1}^n [e^i_j - y^i_j];Z^j)$. To see this, note that by local non–satiation, for each consumer i $p\cdot x^i = p\cdot\omega^i$. For each producer i, the last assumption on V^i implies that $\pi^i \geq 0$ at equilibrium, so that the value of excess demand at equilibrium is non–positive. Since the market agent maximizes the value of excess demand, it must be the case that excess demand for each commodity and stock is non–positive.

Q.E.D.

This proof might be of independent interest since it is not specific to the model with a stock market or asymmetric information, but allows general objective functions for producers. The proof can also be generalized in many standard directions, such as the use of incomplete or intransitive preferences.

An equilibrium is called *completely revealing* if when each agent solves its optimization problem, it faces no uncertainty.

Theorem 4: Suppose that the only uncertainty relevant to any agent's optimization problem is in the profits of the firms; in other words, there is no uncertainty in the model other than ϵ^i, which enters only in firm i's technology. Under the assumptions of Theorems 2 and 3, for each realization of the random vector ϵ, generically there exists a completely revealing equilibrium.

Proof: Apply Theorem 3 to obtain a set $B \subseteq \Omega$ of measure zero such that for any economy not in B, all agents can derive π from known variables. Then apply Theorem 2 to this economy in the case where there is no uncertainty (i.e. π is known to all). This equilibrium is completely revealing.

Q.E.D.

Given ϵ, a *feasible allocation* for the economy ω is a vector $(x^1,\ldots,x^m;Z^1,\ldots,Z^n;Y)$ $\in (\mathbb{R}_+^{n+k})^m x T^1(\epsilon^1) x \cdots x T^n(\epsilon^n) x [0,1]^{n^2}$ such that $\sum_{i=1}^m x^i + \sum_{i=1}^n (Z^i,Y^i) \leq \omega$. Given ϵ, a *Pareto optimum* is a feasible allocation $(x^1,\ldots,x^m;Z^1,\ldots,Z^n;Y)$ such that there is no other feasible allocation $(\bar{x}^1,\ldots,\bar{x}^m;Z^1,\ldots,Z^n;Y)$ with $u^i(\bar{x}^i) \geq u^i(x^i)$ for all i, with strict inequality holding for some i.

Theorem 5: For each consumer i, let u^i be locally non–satiated. For each producer j, suppose that V^j is a monotonic function of π^j *when the producer faces no uncertainty.* Then any completely revealing equilibrium allocation is Pareto optimal.

The proof of this Theorem is standard.

Several remarks are in order. First, on the consumer side, this model involves no uncertainty other than in realized profits of the firms, so there is no random component in the direct utility, only in the budget constraint. This is like most of the finance literature, where there is uncertainty in (future) prices and returns only. Unlike the incomplete markets literature, multiple budget constraints are not present. An interesting example of V^i fitting into our framework is the expected profit function. Finally, consumers do not learn the ϵ^i's in this model, but only learn the profits of firms in equilibrium.

IV. Implications and Conclusions

We have shown above that, in a stock market economy with producers with private information, equilibrium prices, in conjunction with certain other publicly observable variables such as stock purchases and stock endowments, reveal profits or returns anticipated by the producers acting on the basis of their private information. This result is shown to hold under general conditions and does not require that the agents know the true relationship between the initial information of all agents in the economy and the resulting market prices.

In this context, there is an interesting asymmetry between consumers with private information and producers with private information. Our result that a well–functioning stock market reveals information through equilibrium stock prices does not extend to the case where consumers have private information about individual endowments or wealth, simply because the consumers, unlike firms, cannot own each other or issue claims against each other in a standard market setting.

An implication of allowing firms to buy stock in other firms is to make their return vectors interdependent. That is, the manager of a firm can change its return vector (in a linear fashion) by buying stock in other firms. This provides a way to determine the securities offered in equilibrium endogenously. Of course, it must eventually be combined with a model explaining the entry and exit of firms.

Note that our results hold whether markets are complete or not, as long as stock purchases as well as initial stockholdings of firms are observable. In the context of incomplete markets or temporary equilibrium theory (see Green (1973)), the solution to the firm's maximization problem in the presence of asymmetric information will once again reveal some inside information that the firm might have. If consumers and other firms observe this information and condition on it, then their expectations may be more similar to one another than if they did not. This, in turn, might aid in establishing the existence of an equilibrium which requires, as Hart (1974) has shown, that the expectations of the agents must be "sufficiently similar." For further explanation, see Page (1987, Propositions 4.3 and 5.3).

References

ARROW, K., AND DEBREU, G. (1954), "Existence of an Equilibrium for a Competitive Economy," Econometrica, 22, 265–290.

DEBREU, G. (1970), "Economies with a Finite Set of Equilibria," Econometrica, 38, 387–392.

DEBREU, G., and HERSTEIN, I.N. (1953), "Nonnegative Square Matrices," Econometrica, 21, 597–607.

GREEN, J. (1973), "Temporary General Equilibrium in a Sequential Trading Model with Spot and Future Transactions," Econometrica, 41, 1103–24.

GROSSMAN, S. (1978), "Further Results on the Informational Efficiency of Competitive Stock Markets," Journal of Economic Theory, 18, 81–101.

GROSSMAN, S. (1981), "An Introduction to the Theory of Rational Expectations Under Asymmetric Information," Review of Economic Studies, 48, 541–559.

HART, O. (1974), "On the Existence of Equilibrium in a Securities Model," Journal of Economic Theory, 9, 293–311.

KIHLSTROM, R. and MIRMAN, L. (1975), "Information and Market Equilibrium," Bell Journal of Economics and Management Science, 6, 357–376.

MAS–COLELL, A. (1985), The Theory of General Economic Equilibrium: A Differentiable Approach (Cambridge: Cambridge University Press).

PAGE, F. (1987), "On Equilibrium in Hart's Securities Exchange Model," Journal of Economic Theory, 41, 392–404.

RADNER, R. (1979), "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices," Econometrica, 47, 655–678.

RUDIN, W. (1973) Functional Analysis (New York: McGraw–Hill).

SHAFER, W., and SONNENSCHEIN, H. (1975), "Equilibrium in Abstract Economies Without Ordered Preferences," Journal of Mathematical Economics, 2, 345–348.

References

ARROW, K. AND DEBREU, G. (1954), "Existence of an Equilibrium for a Competitive Economy," Econometrica, 22, 265–290.

DEBREU, G. (1970), "Economies with a Finite Set of Equilibria," Econometrica, 38, 387–392.

DEBREU, G. and HERSTEIN, I.N. (1953), "Nonnegative Square Matrices," Econometrica, 21, 597–607.

GREEN, J. (1973), "Temporary General Equilibrium in a Sequential Trading Model with Spot and Future Transactions," Econometrica, 41, 1103–23.

GROSSMAN, S. (1976), "Further Results on the Informational Efficiency of Competitive Stock Markets," Journal of Economic Theory, 18, 81–101.

GROSSMAN, S. (1981), "An Introduction to the Theory of Rational Expectations Under Asymmetric Information," Review of Economic Studies, 48, 541–559.

HART, O. (1974), "On the Existence of Equilibrium in a Securities Model," Journal of Economic Theory, 9, 293–311.

KIHLSTROM, R. and MIRMAN, L. (1975), "Information and Market Equilibrium," Bell Journal of Economics and Management Science, 6, 357–376.

MAS-COLELL, A. (1985), The Theory of General Economic Equilibrium: A Differentiable Approach (Cambridge: Cambridge University Press)

PAGE, F. (1987), "On Equilibrium in Hart's Securities Exchange Model," Journal of Economic Theory, 41, 392–404.

RADNER, R. (1979), "Rational Expectations Equilibrium: Generic Existence and the Information Revealed by Prices," Econometrica, 47, 655–678.

RUDIN, W. (1973) Functional Analysis (New York: McGraw–Hill)

SHAFER, W., and SONNENSCHEIN, H. (1975), "Equilibrium in Abstract Economies Without Ordered Preferences," Journal of Mathematical Economics, 2, 345–348.

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" -- by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua He. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

244. Hayne Leland and Gregory Connor. "Optimal Cash Management for Investment Funds." March 1995.

245. Marcus Berliant and Sankar De. "On Revelation of Private Information in Stock Market Economies." April 1995.

246. Sankar De, Mark Fedenia, and Alexander J. Triantis. "Effects of Competition on Bidder Returns." April 1995.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance
WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 246

Effects of Competition on Bidder Returns

by

Sankar De

Mark Fedenia

and

Alexander J. Triantis

April 1995

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Effects of Competition on Bidder Returns

Sankar De, Mark Fedenia
and
Alexander J. Triantis

April 1995

Finance Working Paper #246

Sankar De is at the University of California at Berkeley. Mark Fedenia and Alexander Triantis are at the University of Wisconsin-Madison. We thank Dean Johnson for valuable research assistance and an anonymous referee, Michael Fishman, Larry Merville, William Putsis, Joan Schmit and Moon Song for comments on an earlier version. We are grateful to Anand Desai for making his tender offer database available to us.

Abstract

This study offers several new perspectives on the effects of competition in takeover contests on bidder returns. Using a more extensive database than existing studies and employing several different measures of success in a takeover, we find that success in competitive acquisitions decreases shareholder wealth relative to failure and also relative to success in observed single-bidder takeovers. Further, we consider and test a number of hypotheses regarding bidder returns, including hypotheses suggested by the preemptive bidding theory. In general, our results indicate lack of support for the predictions of preemptive bidding theory and for the hypotheses linking the method of payment and the observed level of competition. We also test hypotheses relating to returns across the multiple events in a multiple-bid contest that competition among bidders generates. The results of these tests underscore the importance of timing *as well as* success of a bid to the bidder's subsequent performance.

1 Introduction

In assessing the effects of competition on bidder returns, several recent studies have documented and compared returns to bidders in single-bidder and multiple-bidder contests.[1] We re-examine some of those tests using a more extensive database and a methodology that accounts for limited dependent variables. We measure the impact on bidder returns of: (1) bidding by multiple bidders, (2) revision of bids by the same bidder in a takeover contest, (3) bid outcome, and (4) the medium of payment used in the bid. We also test predictions of bidder returns from the preemptive bidding theory of Fishman (1988) and theory linking the medium of payment with the level of competition in a takeover contest [Fishman (1989) and Berkovitch and Narayanan (1990)]. Finally, we use bidder returns across the multiple events associated with contested bids in order to test new hypotheses implied by synergy theory and by Roll's (1986) hubris theory.

Since the success of a bid is not defined unambiguously, we employ several alternative criteria to classify the outcome of a tender offer. Three of the criteria are discrete measures of success: (1) acquisition of any target shares; (2) acquisition of target shares exceeding or equal to the number sought in the tender offer; and (3) acquisition of a controlling number of target shares. In addition, we use two continuous measures: the number of target shares purchased as a proportion of the number of shares sought or tendered for, and the number of target shares held at the conclusion of the tender process as a proportion of the total number of target shares outstanding. In this study, for the most part we focus on results which hold

[1]See, among others, Boebel and Harris (1988), Bradley, Desai and Kim (1988), and James and Wier (1987).

for most, if not all, measures of success.[2] Another noteworthy feature of our study is the extensive sample of tender offers that we use to conduct our tests. The sample includes 958 bids made by non-financial U.S. corporations in 660 tender offer contests during the period 1962-88. This sample is considerably larger than those in existing empirical studies, primarily due to the longer observation period.

We find that successful bidders in multiple-bidder contests earn significantly negative returns and fare significantly worse than successful bidders in single-bidder contests. In contrast to Bradley et al. (1983, 1988), we find that, while the successful bidders in multiple-bid contests appear to experience negative returns in the announcement period, unsuccessful bidders in such contests earn positive returns, regardless of the measure of success employed. In single-bid contests, however, both successful and unsuccessful bidders earn similiar statistically insignificant positive returns. Taken together, our results underscore that success in competitive acquisitions is costly to shareholders of acquiring firms. Interestingly, we also find that revision of bids in an observed single-bidder contest also diminishes bidder returns. Such revision of bids may stem from potential competition. If so, our results imply that both potential competition and actual competition have a negative impact on bidder returns.

We confirm earlier research indicating that bidder returns following the passage of the Williams Act through 1984 appear to be lower than corresponding bidder returns before the Act. We find, however, that bidder returns post-1984 are more favorable. A possible explanation for this finding is that the negative impact of the delay and disclosure provisions

[2]By contrast, the existing studies typically use a single success classification. Bradley et al. (1988) and Jarrell and Poulsen (1989), for example, use the first discrete criterion, whereas Eckbo, Giammarino and Heinkel (1990) employ the third criterion. Boebel and Harris (1988) use a more extreme criterion, namely delisting of the target company.

of the Williams Act on bidder returns were mitigated in the late eighties by lax enforcement of antitrust and other regulatory restrictions on acquisition activities.

Contrary to one prediction of preemptive bidding thory, we find that returns to the bidders in single-bidder cases are not higher than the returns to the first bidders in multiple-bidder contests. Our results also do not support the preemptive bidding theory concerning cost of information acquisition before and after the passage of the Williams Act. Relating the method of payment to the level of competition, we find that pure cash and pure stock offers generate more competition than mixed (cash and stock) offers. Further, we find no evidence that cash offers are associated with competition any differently than are stock offers. Our results, therefore, support neither Fishman (1989) nor Berkovitch and Narayanan (1990).

Our analysis of bidder returns across announcement dates in multiple-bidder contests hints at a combination of synergy and first-mover advantage in the takeover bids. The abnormal returns to a bidding firm in the period surrounding the announcement of its bid are positively related to its abnormal returns when a competing bidder subsequently puts in a bid for the target. The correlation is significant if the first of these two bidders is also the first bidder in the takeover contest. When we control for the success of the first of two consecutive bids, success of the first bid has a significantly positive impact on first bidder returns at the second bidder's announcement date. So, there are gains to being first *as well as* successful in a multiple-bidder contest.

The rest of the paper is organized in the following manner. In Section 2, we discuss the hypotheses that we test in this study. We discuss our data and methodology in Section 3. We present our results in Section 4 and our conclusions in Section 5.

2 Testable hypotheses

2.1 Predictions of preemptive bidding theory

In Fishman's (1988) theory of preemptive bidding, at the beginning of the bidding game for a particular target each potential bidder has a separate, but common knowledge, distribution of probable gains (synergy) from acquiring the target. However, any bidder can acquire information about the precise value of the target at a cost. Upon observing the first bidder's initial offer, a second bidder updates his prior beliefs about the first bidder's private valuation and he decides whether to acquire information and submit a competing bid. The first bidder, in determining an initial offer, takes into account the actions of any subsequent bidder. Strategic interaction between the bidders ensures that the higher the private valuation of the first bidder, the lower is a second bidder's expected outcome from entering the takeover contest. In equilibrium, a first bidder may acquire the target by making a preemptive high-premium bid to signal a high private valuation and deter a second bidder from competing. Alternatively, the first bidder makes a low-premium bid, indicating a low private valuation. A second bidder enters, an English auction for the target ensues, and a takeover occurs at the second-highest valuation. Preemptive bidding theory therefore implies that the first bidder's return will be higher in a single-bid contest than in a multiple-bid contest. For a proper test of this prediction, it is necessary to compare returns to first bidders in multiple-bidder contests, regardless whether they are successful or not, i.e., returns to bidders that clearly fail to preempt competition, with returns to the bidding firms which preempt other bidders (as evidenced by their successful first bid).

In reality, the bidding process may not resemble a frictionless English auction. Costs are associated with submission and revision of bids (e.g., direct costs such as fees to investment bankers, counsels, and consultants, and indirect costs such as loss of executive time). If counterbidding costs are significant, lower information acquisition costs will result in higher threshold offers for targets. The resulting impact on bidder returns in single-bid contests is, however, unclear. While an increase in the threshold offer would necessitate a higher preemptive bid, the downward impact on bidder returns may be offset by bidders with higher private valuations than before engaging in preemptive bidding. The theory, however, predicts that lower information acquisition costs would result in higher returns, in general, for the first bidder in a multiple-bidder contest, a higher threshold indicating higher average valuation across all bidders below the threshold.[3] To test this prediction, we examine bidder returns around the Williams Act. This Act mandates pre-offer filing of key information and a minimum period during which tender offers must remain open. Assuming the Act has reduced the cost of acquiring information for the competing bidders, the preemptive bidding theory predicts that the first bidder return in the period following the Williams Act should be higher than that in the period before the Act.

2.2 Predictions of medium of payment theory

Fishman (1989) and Berkovitch and Narayanan (1990) develop theories in which the medium of payment has a direct bearing on the level of competition among the potential bidders. In Fishman (1989), a bidder with a high private estimate of the value of an acquisition

[3] In the event that competition among the bidders drives the surplus entirely to the targets, the net impact on bidder returns may be insignificant. However, this is an extreme outcome.

makes a cash offer, as opposed to a stock or a mixed offer, in order to exclude the target shareholders from sharing the gains. Since such a bid is also likely to deter other bidders from competing, a testable implication of this model is that competing bidders are less likely to be observed following an initial cash offer than following an initial non-cash offer. In Berkovitch and Narayanan (1990), the target receives a higher *dollar* amount from a high-synergy acquirer but, given that the low-synergy acquirer faces greater competition, a higher *proportion* of the total synergy from a low-synergy acquirer. Under asymmetric information, the high-synergy acquirer uses a higher proportion of cash and the low-synergy acquirer uses a higher proportion of equity in their respective offers. Since neither has the desire to mimic the other, this arrangement works out as a separating equilibrium. As potential competition increases, the use of cash, both in dollar amount and as a proportion of the total offer, increases. Furthermore, with actual competition only the *lowest-synergy* bidder makes a non-cash offer. Hence, we examine whether cash offers are more likely to be seen than non-cash offers in observed multiple-bidder takeover contests.

2.3 Predictions of overbidding versus synergy theories

The overbidding and synergy theories of takeovers have implications for bidder returns at different announcement dates in a multiple-bidder contest. Conventional takeover theory suggests that takeovers unlock positive synergies between the target and bidding firms that partly accrue to the bidder. Alternatively, Roll (1986) argues that the managers of bidding firms, infected by hubris, overestimate the value of their targets and overpay. Schleifer and Vishny (1988) believe that managers overpay because their selection of acquisition targets is guided by objectives other than value maximization. Overbidding theory predicts that the return to a bidder is negative when the bidder makes it bid. However, when another bidder makes a competing bid, the return to the first bidder at that date should be positive since the competing bid lowers the probability of the first bidder's success. In contrast, if synergy motivates the takeover bid, then we expect the bidder to experience a positive stock market reaction to the first bid, followed by a negative reaction when another bidder enters the contest. The negative reaction is a consequence of the lower probability of realizing the synergy.

3 Data and methodology

3.1 Sample selection and data description

Our primary database consists of tender offers made by firms registered with the *NYSE* or the *AMEX*, as well as *OTC* firms, during the period from July 1962 through December 1988. Various sources were used to make the information in the database as complete as possible. For the period prior to January 1981, our primary source was the database used in Bradley et al. (1983).[4] Following this period, the list of tender offers was obtained from Douglas Austin & Associates' Tender Offer Statistics (for the years 1981–86), and from Merrill Lynch's annual Mergerstat Review (for 1987 and 1988).[5] Relevant information about these tender offers was pieced together from the Wall Street Journal Index, from the

[4]This is based, in part, on databases compiled by Bradley (1980) and Dodd and Ruback (1977), and contains tender offers during the 1958-80 period in which either the target or the bidder was listed on the NYSE or the AMEX.

[5]The Austin database compiles tender offers using SEC 14D-1 filings. Mergerstat lists tender offers for which both purchase price and earnings of the seller were available.

MERGER library of LEXIS, and from the Wall Street Journal Abstracts in NEXIS.

From the set of all tender offers made during the study period, bids were excluded if there was incomplete or unreliable information about the number of bids or bidders in a tender offer contest, the exact date of the tender offer, the method of payment proposed for the tender offer, whether the bidding firm was able to purchase any target shares, or about event period returns (e.g., if not available from either *CRSP NYSE/AMEX* or *NASDAQ* databases). Bids were also excluded if the bidding firm had its stock traded for less than 50 days during the estimation interval [-200,-51] or if the target firm was privately held. Bids made for a given target were assumed to be part of the same tender offer contest if they followed a previous bid by no more than 120 trading days. Though any such cutoff would be arbitrary to some degree, we ascertained that a 120-day cutoff is consistent with the duration of tender offer contests in our database.[6]

The above selection criteria resulted in the final sample consisting of a total of 958 bids made by 739 bidding firms in 660 tender offer contests for target firms.[7] The descriptive statistics of the tender offers in our final sample are presented in Table 1. Panel A shows 367 contests in our final sample involving a single offer from one bidder, 93 involving more than one bid from a single bidder (for a total of 184 bids), and 200 contests with competition among bidders (for a total of 407 bids). Our sample of 958 bids from 660 contests is considerably larger than those in existing studies.[8] Primarily three factors account for our larger sample

[6]Almost two-thirds of all consecutive bids in multiple-bid contests in our sample are within 20 trading days of each other and 90% are within 50 trading days of each other. In all but five of the 293 multiple bid contests the total duration of the contest is shorter than 120 days, and three-quarters are shorter than 30 trading days.

[7]Without the application of the selection criteria with respect to the availability of return data, our sample consists of 2,106 bids made by 1,624 bidding firms in a total of 1,316 contests.

[8]For example, Bradley et al. (1988) employ a sample of 236 tender offer contests consisting of 163 single-bidder and 73 multiple-bidder contests. The sample in Boebel and Harris (1988) includes 139 tender offers.

size: 1) our study period is longer (as discussed in the following paragraph); 2) our sample includes both successful and unsuccessful bids; and 3) our sample includes tender offers made by *OTC* firms.

Insert Table 1 Here

Panel B indicates the frequency with which tender offers occurred in the four separate time periods covered by our study. The frequencies are presented in terms of the number of contests as well as the number of bids initiated in each period. The first three periods (7/62–6/68, 7/68–12/80, and 1/81–12/84) are the same as in Bradley et al. (1988), with the first period ending with the passage of the Williams Act. The last period (1/85–12/88) provides an additional 303 contests (45.9% of the total) and 440 bids (45.9% of the total). Thus, the inclusion of this four–year period almost doubles our sample size.

In Panel C, the bids are classified according to the method of financing proposed: cash, stock, or mixed (both cash and stock). The panel indicates that the overwhelming majority of bids in our sample, 84.1%, involved cash financing exclusively. Panel D of table 1 presents descriptive statistics on acquisition of target shares by the bidders in our sample. The bidders held an average of 11.12% (but a median of only 0.45%) of the targets' shares prior to the announcement of their bids. They sought an average of 63.71% of the shares through their bids, but were successful in acquiring only 33.18% on average.

As mentioned earlier, our study examines both unsuccessful and successful bids. There are several definitions of success that may be appropriate. Many existing studies [e.g., Bradley et al. (1988)] classify a bid as successful if the bidder is able to purchase any of the target's outstanding shares. We use the binary variable *SUCCESS* to correspond to

Among the existing studies, the sample of 770 tender offers in Jarrell and Poulsen (1989) is the largest.

this definition. An alternative definition, represented by the variable *OUTCOME*, is that the bid is successful if the bidder acquires at least the number of shares that it seeks in its tender offer. Further, the binary variable *CONTROL* reflects whether or not the bidder gains control of the target firm by holding 50% or more of the outstanding shares through the exercise of the tender offer. In addition, we use two continuous measures of success: number of shares purchased as a proportion of the number of shares sought (*PUR/SOUGHT*), and the number of shares owned as a proportion of the total number of shares outstanding at the conclusion of the tender process (*OWN/OUTSTD*). Table 2 presents data regarding the five success classification variables for the sample of tender offers in our study.

Insert Table 2 Here

Panel A presents a correlation matrix for the five success classification variables (where the binary variables equal one if the bid is successful and zero otherwise).[9] Each entry in the matrix shows the Pearson correlation coefficient between two corresponding success measures. Below each of these correlation coefficients is the number of bids for which we are able to obtain sufficient information to calculate both success measures.[10] The correlation matrix amply documents the fact that the success measures are far from perfect substitutes for each other, suggesting that results based on success may be sensitive to how success is defined. Whenever possible, we use all five definitions in our analysis and note any discrepancies which arise.

[9]Note that more than one bid for a given target may be classified as successful by some of the success classification criteria.

[10]Recall that our final sample includes only bids for which we had sufficient information to determine whether they were successful in acquiring any target shares. As a result, we are able to classify all 958 bids in the sample as successful or otherwise by the classification measure *SUCCESS*. However, we have sufficient information to classify only 826 bids by the classification measure *OUTCOME*, 547 bids by the measure *CONTROL*, and 540 bids by all three binary classification measures.

Panel B of Table 2 classifies the bids in our sample as successful or unsuccessful using the three binary success classification criteria. Note that, of the three criteria, *OUTCOME* accounts for the highest proportion of unsuccessful bids (72.4%). In view of the fact that it is the most restrictive criterion – a bid is not successful unless the number of shares acquired is at least equal to the number sought – this finding should come as no surprise. Panel C presents some descriptive statistics regarding the two continuous classification criteria.

3.2 Methodology

The abnormal return for firm i at time t was calculated as $AR_t^i = R_{it} - \alpha_i - \beta_i R_{mt}$ where R_{it} is the return for firm i, R_{mt} is the market return (using the $CRSP$ equally-weighted index), and α_i and β_i are firm specific market model parameters estimated by an ordinary least squares regression. The estimation interval is $[-200, -51]$ trading days relative to the event date ($t = 0$), where the event is the announcement of the tender offer as reported in the Wall Street Journal Index. Each firm included in the sample had its stock traded for at least 50 trading days during the estimation period. In order to minimize the impact of possible confounding events, the cumulative abnormal return for firm i (CAR_i) is computed over the two-day event period $[-1, 0]$: $CAR^i_{[-1,0]} = AR^i_{-1} + AR^i_0$. We calculate average cumulative abnormal returns for different subsamples of our data, and test whether these are significantly different from zero, and from each other. In regression equations, we control for truncated data and discrete data for the dependent variable, as discussed below.

When observations are systematically excluded from a sample, truncation bias may affect parameter estimates.[11] Eckbo, Maksimovic and Williams (1990) discuss why merger data

[11] See Davidson and MacKinnon (1993) for a detailed discussion of problems that arise from truncation

samples exhibit this characteristic. Abnormal returns associated with announcements of mergers, or tender offers, may occur if managers possess private information not yet revealed in market prices. Since shareholders infer that managers believe the voluntary event to have a positive net present value, this truncates the residual term that measures the value of inside information. As in Eckbo et al. (1990), we assume that a tender offer is made only if the return to bidder i as perceived by its managers, $x_i\gamma + \eta_i$, is positive, where x_i is a vector of publicly observable variables, γ is a weighting vector of constants (the parameters that are estimated) and η_i is the bidder's private information (which is distributed normally with mean zero and variance ω^2). The truncation bias can be eliminated by including the following non-linear term in the regression: $E(\eta_i \mid \eta_i \geq -x_i\gamma) = \omega\frac{n(z_i)}{N(z_i)}$, where $z_i = \frac{x_i\gamma}{\omega}$, and n and N are the standard normal density and distribution functions, respectively [see Eckbo et al. (1990, Section 1)]. Thus, the regression equation becomes

$$CAR^i = x_i\gamma + \omega\frac{n(z_i)}{N(z_i)} + \epsilon_i$$

We employ non-linear least squares to estimate the parameters in γ, using the estimated parameters from the OLS regression (i.e., dropping the non-linear term) as initial values.[12]

To test the hypotheses in Section 2.2 concerning the linkage between method of payment and the number of bidders, we use event count data (the number of bidders in a contest). Regression methods developed for continuous dependent variables lead to considerable inefficiencies when used in regressing discrete, and positively-skewed, dependent variables against

and techniques that can be used to obtain consistent estimators.

[12]In tests where the dependent variable is the average abnormal return to a bidder at the announcement of a bid by *another* bidder, there is no similar truncation problem. This follows from the fact that, in such cases, the bid is not a voluntary event initiated by the first bidder's management, and, as Eckbo et al. (1990) assume, each bidder's private information is independently distributed across the bidding firms.

explanatory variables (in our case, the method of payment).[13] Therefore, we use a regression procedure that explicitly accounts for the discrete distribution of the dependent variables.

The population that we examine is the set of firms that receive tender offers.[14] Let y_i, $i = 1, \ldots, n$, denote the number of subsequent bidders following the first bidder in each contest. We assume that this number is Poisson distributed.[15] The Poisson distribution assigns a positive probability to any non-negative event count, including a zero event count, i.e., cases where there is only a single bidder. The density function for y_i is: $f_P(y_i|\lambda_i) = \frac{e^{-\lambda_i}\lambda_i^{y_i}}{y_i!}$ for $\lambda_i > 0$ and $y_i = 0, 1, 2, \ldots$, where λ_i is the expected value of the Poisson variable (i.e., the average number of bidders not including the first bidder).

Under the Poisson regression model, λ_i is assumed to be an exponential-linear function of a vector of explanatory variables, x_i, such that: $E(y_i) = \lambda_i = e^{x_i\beta} = e^{\beta_0+\beta_1 x_{1i}+\beta_2 x_{2i}}$. For our purposes, β $(= [\beta_0, \beta_1, \beta_2]')$ is a vector of effect parameters corresponding to mixed (intercept), and cash and stock financing explanatory dummy variables (x_1 and x_2). To estimate these coefficients, we construct the log-likelihood function (dropping the term that does not depend on β), $\ln L(\beta|y) = \sum_{i=1}^{N} y_i x_i \beta - e^{x_i\beta}$, and maximize over β. We also report asymptotic standard errors from our regression model. The coefficients β_1 and β_2 indicate whether cash or stock bids lead to a larger or smaller number of bidders than mixed bids.

[13]See Amburgey and Carroll (1984) for a discussion of these inefficiencies and the use of the Poisson process in event count regressions.

[14]Alternatively, we could take as the underlying population all firms that either did receive, or could *potentially* receive, tender offers from bidding firms. We could then look at the unconditional expected number of bidders taking into account a large subset of the firms in the population receiving no tender offers. If we assume that the number of bidders is Poisson distributed, we can use a Truncated-at-Zero Poisson distribution to estimate the underlying unconditional mean of the Poisson distribution using data observed from the conditional (truncated) distribution. This approach was attempted, but the results implied that only about one-third of the total population of firms did not receive any bids. Since this seems too low, we hesitate to follow this estimation technique for the regression.

[15]While other discrete distributions, such as the negative binomial distribution, are used in event count models, we find that the Poisson distribution provides the best fit to our data.

4 Results

4.1 Bidder returns in single- and multiple-bidder contests

We first re-examine tests in the existing literature on bidder returns, using our larger database, employing different measures of success, and controlling for a larger number of relevant variables than have been included in these studies. In Table 3, we report results concerning the joint effects of the level of competition and the success of a bid on the associated bidder returns. A four-way analysis of bidder returns is reported: single (SB) versus multiple (MB) bidders and success (Y) versus failure (N) (using all three discrete measures of success). Three results in the table are noteworthy. First, we find that successful bidders in multiple-bidder contests experience *significant* negative abnormal returns. Other studies, such as Bradley et al. (1988) and Stulz et al. (1990), have found negative but insignificant returns to successful bidders in such cases. Second, successful bidders in multiple-bidder contests fare significantly worse than successful bidders in single-bidder contests. Though other studies report a similar result [see Bradley et al. (1988), Boebel and Harris (1988), and Stulz et al. (1990)], none of them have found the difference between the mean returns for the two groups as statistically significant. In our study, the two results are observed for each of the three success classifications and are highly significant in each case. We believe that our larger sample size explains the statistical significance of our results. Third, we find that, while successful bidders in multiple-bidder contests experience significant negative returns, unsuccessful bidders in such contests earn either insignificant or positive returns. Moreover, the difference in average returns between the two groups is statistically significant in all three cases. To the best of our knowledge, no existing study on the subject has made this observation. In fact, Bradley et al. (1983, 1988) arrive at precisely the opposite conclusion. They find that unsuccessful bidders in multiple-bidder contests experience a significant decline in their preoffer value.[16]

Insert Table 3 Here

Another interesting result in this connection concerns the abnormal returns to successful bidders in single-bidder contests. These returns are positive for all success classifications. However, they are significant (at the 10% level) only in the case where the success measure is *SUCCESS*.[17] Interestingly, as can be seen from the footnote to Table 3 which further breaks down this successful single-bidder case, the bidder returns are positive and significant (at the 5% level) if the single-bidder bids only once, but negative (although insignificant) if the bidder bids more than once. Thus, for bidders that acquire some shares of the target, the returns are positive only if the bidder does not need to revise its initial bid with a higher offer. This result indicates that revision of bids, possibly resulting from the threat of potential competition, causes dilution of bidder returns in single-bidder contests in much the same way as does actual competition in multiple-bidder contests.

In Table 4, we examine whether the payment method proposed for a bid and its period of origination, in addition to its success and the number of bidders in the associated tender offer contest, affect bidder returns. Consistent with the results in Table 3, we find that the multiple-bidder dummy variable (*COMPETITION*) has a significantly negative coefficient

[16]They conclude that "once a firm finds itself in a bidding war, it is better to win than to lose, even though in winning the firm's shareholders may suffer a capital loss" (Bradley et al. (1988), p. 29).

[17]Bradley et al. (1988) also find significant positive announcement period returns to successful bidders in single-bidder contests using this success measure.

for three of the five measures of success.[18] In addition, success in a tender offer contest, measured by any of the five success classification variables, leads to lower returns than the corresponding returns to unsuccessful bidders. The returns are significantly lower for four of the five measures, but not for the *SUCCESS* criterion. Conceivably, success by other measures such as *OUTCOME* and *CONTROL*, which are more demanding than *SUCCESS*, requires more overbidding by the bidder firms to persuade the target shareholders to tender their shares.

Insert Table 4 Here

From Table 4, the returns in the first two periods (7/68-12/80 and 1/81-12/84) following the Williams Act are uniformly lower than the returns in the pre-Williams Act period.[19] Further, they are highly significantly lower (at the 1% level) in three of the five regressions for the period 7/68-12/80. Interestingly, in the last post-Williams Act period (1/85-12/88), the returns are statistically indistinguishable from the pre-Williams Act period except in one case, and they are higher in that case. A possible explanation is that the negative impact of the delay and disclosure provisions of the Williams Act on bidder returns was mitigated in the late 1980's by the lax enforcement of antitrust regulations.

Finally, noting that the intercept terms in the regression equations in Table 4 capture the effect of cash financing, stock financing appears to generate uniformly lower returns than cash financing. These results are generally consistent with existing empirical studies that

[18]Though we might expect the coefficient to be significantly negative everywhere, recall from Table 3 that unsuccessful bids in multiple-bidder contests generate neutral or positive returns, presumably making the negative impact of competition less pronounced for all bids taken together. Only two existing studies have looked at the effects of both multiple bids and the means of payment on bidder returns. Both Franks et al. (1988) and Boebel and Harris (1988) find no statistically significant relationship between multiple bids and bidder returns after controlling for the method of payment.

[19]This result is consistent with Bradley et al. (1988) and Jarrell and Poulsen (1989).

report insignificant abnormal returns to the bidding firms in cash-financed acquisitions but significantly negative abnormal returns in stock-financed acquisitions.[20]

4.2 Tests of preemptive bidding theory

We conducted tests of the implications of preemptive bidding theory discussed in Section 2.1. We first examine whether there is any significant difference between the returns to the bidders in observed single-bidder cases and the returns to the first bidders in multiple-bidder contests at the time of the announcement. To make sure that the bids in the single-bidder cases indeed fit the label preemptive, we consider only successful single-bidder cases where the bidder bids *only once*.[21] Panels A, B, and C of Table 5 show that the announcement period abnormal returns to the single-bidders are positive, but insignificant except when the success measure is *SUCCESS*. The returns to the first bidders in multiple-bidder contests are, however, significantly positive and higher. Even though the returns for the two groups are not significantly different from one another, our results oppose the prediction of preemptive bidding theory.

Insert Table 5 Here

As discussed in Section 2.1 above, in order to test predictions of preemptive bidding theory concerning the cost of information acquisition we analyze the return to the first bidder in a

[20]See Asquith, Bruner and Mullins (1987), Travlos (1987), and Franks, Harris and Mayer (1988). In contrast, Cornett and De (1991) find significant positive announcement period bidder returns in mergers of commercial banks regardless of the method of payment chosen. Further, they find an insignificant difference between bidder returns associated with different payment methods.

[21]In our sample, there are 315, 128, and 147 such observations corresponding to the three success classifications, *SUCCESS*, *OUTCOME*, and *CONTROL*, respectively. (While our sample consists of a total of 394, 162, and 184 bids in successful single-bidder cases based, respectively, on the three discrete measures, of them, 79, 34, and 37, respectively, represent bids in cases where the bidders made more than one bid each.) On the other hand, our sample includes 107 multiple-bidder contests (see Table 1).

multiple-bidder contest before and after the Williams Act. The theory predicts that, given lower costs of information acquisition expected in the period following the Williams Act, observed first bidder returns in multiple-bidder contests should be higher during this period than in the pre-Act period. Our results for returns before and after the Williams Act in Panel D are again not consistent with the prediction of preemptive bidding theory. The abnormal returns to the first bidder in a multiple-bidder contest are statistically indistinguishable in the two periods. While this result does not support preemptive bidding theory, because of the paucity of observations in the pre-Williams period one should be cautious in inferring from this finding a firm rejection of preemptive bidding theory.

4.3 Tests of medium of payment theory

In Panel A of Table 6, we present results from our test designed to determine whether the observed level of competition and the medium of payment chosen are linked as suggested by Fishman (1989). We regress the number of bidders *following* the first bidder in each contest on the method of payment for the first bid in the contest. As discussed in Section 3.2, we employ a log-linear specification in a Poisson regression model. Thus, the antilog of the intercept coefficient gives the expected number of bidders following the initial bidder in a contest (i.e., the expected total number of bidders minus one) associated with mixed cash and stock bids. The antilog of the coefficient of each dummy variable measures the relative frequency of bidders (in excess of one) for either cash-financed or stock-financed bids compared to the incidence of mixed bids. For example, if the coefficient is negative, the antilog of this coefficient measures how many fewer (multiplicatively) bidders are on average associated with the form of financing represented by the explanatory variable.

Insert Table 6 Here

The results in Panel A indicate the following. If the initial bid in a tender offer contest is mixed, it leads to an expected number of 0.19 ($= e^{-1.674}$) bidders following the first bidder in the contest. In this case, the expected number of subsequent bidders is significantly below unity ($= e^0$) at the 1% level. The panel also indicates that an initial all-cash offer is expected to lead to significantly more additional bidders than a mixed offer, precisely 2.14 ($= e^{.761}$) times more. The results for initial all-stock offers are similar; they lead to 2.81 ($= e^{1.032}$) times as many additional bidders as mixed offers. Furthermore, we have checked that the difference between cash and stock offers is insignificant. This result refutes the prediction in Fishman (1989) that competing bidders are more likely to be observed following an initial non-cash offer than following an initial cash offer.

We also investigate, in Panel B, the prediction in Berkovitch and Narayanan (1990) that, in contests involving competition, cash offers are more likely to be observed than offers of other kinds. It does appear that, for the 407 bids in multiple-bidder contests, cash offers (348 or 85.5% of the total) overwhelm both mixed offers (22 or 5.4%) and stock offers (37 or 9.1%). However, as the panel indicates, the proportion of cash offers is virtually the same in the subsamples of tender offers which exhibit no competition at all. Cash offers accounted for 83.4% of the total number of bids in single bidder/single bid cases and 82.6% of the total number of bids in single bidder/multiple bids cases. Since the relative frequency of the cash offers is virtually uniform for competition and no-competition cases in our sample, we reject the prediction in Berkovitch and Narayanan (1990).

4.4 Tests of overbidding versus synergy theories

To test an implication of overbidding versus synergy theories, we examine whether there is any significant relationship between the returns to a given bidder measured at two dates: when that bidder announces its bid, and again when a competing bidder subsequently announces its bid. Table 7 reports the abnormal returns to the first in a pair of consecutive bidders in a multiple-bid contest at the first bidder's announcement date and again at the competing bidder's announcement date. In our sample, there are a total of 407 bids made in 200 multiple-bidder contests, yielding a total of 207 pairs of consecutive bids. Eliminating the pairs of consecutive bids made by the same bidders, we are left with 200 observations.

Insert Table 7 Here

In panel A, we see that the abnormal return to the first bidder in a pair is positive, both at the first announcement date (1.01%) and at the second announcement date (.57%). However, only the return at the first announcement date is significantly different from zero. In light of the hypotheses discussed in Section 2.3, the results point in favor of synergy. The synergy could result in a positive abnormal return, on average, at the first bid announcement. However, since the appearance of the second bidder reduces the probability of the first bidder being successful (or results in the first bidder being successful only at a higher offer price, reducing its share of the takeover gains), the synergy hypothesis predicts a neutral or even negative effect on the first bidder's return at the second announcement date.

For further confirmation of this interpretation, we investigated whether the appearance of a second bidder in a multiple-bidder contest actually lowers the probability of the success of the first bidder in our sample. We found it is indeed so for all three of the discrete success measures. In the overall sample including single-bidder and multiple-bidder contests, using the three binary success measures, the first bidders are successful in 71.0% ($SUCCESS$), 33.7% ($OUTCOME$), and 56.2% ($CONTROL$) of the contests. In multiple-bidder contests, the corresponding success rate for the first bidder is considerably lower in each case: 21.1% ($SUCCESS$), 11.4% ($OUTCOME$), and 15.3% ($CONTROL$).

We examine whether the abnormal returns at the two announcement dates are significantly related. In Panel B, the regression of the return to the first bidder at the second bid announcement date (RET_TWO) on the return at the first announcement date (RET_ONE) yields an insignificant coefficient estimate of 0.104. In Panel C, however, we find that by controlling for the timing of the pair of competitive bids, a significant relationship emerges. The variable $LATE \cdot RET\_ONE$ is a slope indicator variable of RET_ONE and $LATE$, where the latter variable equals one if the first bid in the pair is not the first bid of the tender offer process and zero otherwise. The regression estimates indicate that bid pairs that appear after the first bid of the process may have different characteristics than bid pairs that include the first bid. In particular, the relationship between the bidder returns at the two announcement dates is now significantly positive (at the 10% level) if the first bid in the pair is also the first bid in the tender offer process, and negative, though not significantly so, if the bid pair is later in the tender offer contest. The former significant result cannot be explained fully by either the overbidding or synergy hypothesis. However, it hints at the presence of a first mover advantage in a takeover contest.

The final three models in Panels D, E, and F document the importance of the success of the first bid in explaining the first bidder's return at the second bid date. We employ all three discrete success classifications. The coefficient for the success indicator variable is significant in all regressions and is between 2.6%-4.2%, indicating that if the first bid is successful the returns to the bidder are on average significantly higher at the second bid date than if the first bid were unsuccessful. A possible explanation is that the second bidder bids at a higher premium than the offer price of the first bidder, resulting in a potential gain to the first bidder if the first bidder has been successful in acquiring some target shares. In addition, the second bid may indicate that the first bidder has been successful in identifying a valuable target, thus increasing the market's perception of the value of the first bidder's acquisition program, or more generally of the ability of the bidding firm's management to identify potentially valuable investments.

Taken together, our findings indicate significant subsequent returns from being first *and* successful in a multiple-bidder contest, particularly if the returns from the first move are favorable.[22] It is an interesting result, particularly in view of our earlier finding reported in Table 3 that success in all multiple-bidder contests taken together generates, on average, significant negative returns. This disparity is partly due to the fact that in many multiple-bidder contests the successful bids are the ultimate bids in the process; they are not followed by other bids. Further, recall that the results in Table 3 reflect a bidder's abnormal return at the announcement of its *own* bid, and not at the announcement of a bid by a subsequent bidder.

[22]Bradley et al. (1988) also examine the joint effect of timing and success on bidder returns in multiple-bidder contests. They find that first-bidder acquirers experience high positive returns, while late-bidder acquirers experience a significant wealth loss. However, their test is quite different from the ones reported here. They consider the stock price performance of successful bidders over a variable event window extending from the first bid in a tender offer contest to the ultimately successful bid. Their approach does not permit the kind of investigation we attempt in this case, namely the impact of a bid by a subsequent bidder on the returns to the first bidder in a tender offer contest.

5 Conclusions

In the first part of our study, where we replicate tests in the literature using a larger database, several different measures of success and non-linear regressions, we find that success in an overtly competitive acquisition process, such as an observed multiple-bidder contest, is usually costly to the shareholders of the successful bidding firm. On the other hand, success in an observed single-bidder contest as well as failure in a multiple-bidder contest is usually not costly. Interestingly, however, revision of bids in an observed single-bidder contest diminishes bidder returns. If such revision of bids stems from potential competition, our results indicate a negative impact of actual *as well as* potential competition on bidder returns.

We test predictions of preemptive bidding theory and hypotheses linking the medium of payment and the observed level of competition, and fail to find support for them. The results of our tests of predictions regarding bidder returns across announcement dates in multiple-bidder contests appear to be intriguing. They indicate that the abnormal returns to a bidding firm in the period surrounding the announcement of its bid are positively related to its abnormal returns when a competing bidder subsequently puts in a bid for the same target, significantly so if the first of the two bidders is also the first bidder in the takeover contest. The results seem to hint at a combination of synergy and first mover advantage at work. Upon closer examination, we find that the success of the first bid has a positive and significant impact on the abnormal returns to the first bidder at the second announcement date. Taken together, these findings indicate significant rewards to being first as well as successful in a multiple-bidder contest.

References

Amburgey, T. L. and Carroll, G. C.: 1984, Time-series models for event counts, *Social Science Research* **13**(1), 38–54.

Asquith, P., Bruner, R. and Mullins, D.: 1987, Merger returns and the form of financing, *Working paper*, Harvard University.

Berkovitch, F. and Narayanan, M.: 1990, Competition and the medium of exchange in takeovers, *Review of Financial Studies* **3**(2), 153–174.

Boebel, R. and Harris, R.: 1988, Competition in the bidding market and shareholder wealth, *Working paper*, University of Virginia.

Bradley, M.: 1980, Interfirm tender offers and the market for corporate control, *Journal of Business* **54**(4), 345–376.

Bradley, M., Desai, A. and Kim, E. H.: 1983, The rationale behind interfirm tender offers: Information or synergy?, *Journal of Financial Economics* **11**(1), 183–206.

Bradley, M., Desai, A. and Kim, E. H.: 1988, Synergistic gains from corporate acquisitions and their division between the stockholders of target and acquiring firms, *Journal of Financial Economics* **21**(1), 3–40.

Cornett, M. and De, S.: 1991, Medium of payment in corporate acquisitions: Evidence from interstate bank mergers, *Journal of Money, Credit, and Banking* **23**(4), 767–776.

Davidson, R. and MacKinnon, J. G.: 1993, *Estimation and Inference in Econometrics*, Oxford University Press, Inc., New York, New York.

Dodd, P. and Ruback, R.: 1977, Tender offers and stockholder returns, *Journal of Financial Economics* **5**(3), 351–373.

Eckbo, B., Maksimovic, V. and Williams, J.: 1990, Consistent estimation of cross-sectional models in event studies, *Review of Financial Studies* **3**(3), 343–365.

Eckbo, E., Giammarino, R. and Heinkel, R.: 1990, Asymmetric information and the medium of exchange in takeovers: Theory and tests, *Review of Financial Studies* **3**(4), 651–675.

Fishman, M.: 1988, A theory of preemptive takeover bidding, *RAND Journal of Economics* **19**(1), 88–101.

Fishman, M.: 1989, Preemptive bidding and the role of the medium of exchange in acquisitions, *Journal of Finance* **44**(1), 41–58.

Franks, J., Harris, R. and Mayer, C.: 1988, Means of payment in takeovers: Results for the U.K. and the U.S. in A. Auerbach (ed) corporate takeovers: Causes and consequences, *Working paper*, National Bureau of Economic Research, Inc.

James, C. and Wier, P.: 1987, Returns to acquirers and competition in the acquisition market: The case of banking, *Journal of Political Economy* **95**(2), 355–370.

Jarrell, G. and Poulsen, A.: 1989, The returns to acquiring firms in tender offers: Evidence from three decades, *Financial Management* **18**(3), 12–19.

Roll, R.: 1986, The hubris hypothesis of corporate takeovers, *Journal of Business* **59**(2), 197–216.

Schleifer, A. and Vishny, R.: 1988, Value maximization and the acquisition process, *Journal of Economic Perspectives* **2**(1), 7–20.

Stulz, R., Walkling, R. and Song, M.: 1990, The distribution of target ownership and the division of gains in successful takeovers, *Journal of Finance* **45**(3), 817–833.

Travlos, N.: 1987, Corporate takeover bids, methods of payment, and bidding firms' stock returns, *Journal of Finance* **42**(4), 943–963.

Table 1: **Descriptive statistics of tender offers.** Panel A classifies tender offer contests by the number of bidders involved and the number of bids received by a target. Panel B indicates the number of bids received by targets in each of four periods, as well as the number of contests which began in each of the periods. Panel C shows the number of bids which proposed the following three financing methods: cash, stock, and mixed (i.e., both cash and stock). Panel D provides information concerning the percentage of the target's outstanding shares held by a bidder at the date of the announcement of a bid, the percentage of the target's shares that the bidder proposes to purchase through the tender offer, and the percentage of shares that the bidder succeeds in obtaining through its offer.

	Contests		*Bids*		*Bidders*	
A. Classification of offers						
	Frequency	%	Frequency	%	Frequency	%
Single Bidder, Single Bid	367	55.6	367	38.3	367	49.7
Single Bidder, Multiple Bids	93	14.1	184	19.2	93	12.6
Multiple Bidders	200	30.3	407	42.5	279	37.8
Total	660	100.0	958	100.0	739	100.0
B. Time Period						
	Frequency	%	Frequency	%		
7/62–6/68	47	7.1	53	5.5		
7/68–12/80	143	21.7	213	22.2		
1/81–12/84	167	25.3	252	26.3		
1/85–12/88	303	45.9	440	45.9		
Total	660	100.0	958	100.0		
C. Medium of payment						
			Frequency	%		
Cash bids			806	84.1		
Stock bids			67	7.0		
Mixed bids			85	8.9		
Total			958	100.0		

D. Bidder's position	Mean	Median	Standard Deviation
% of Target Shares Owned	11.12	0.45	18.63
% of Target Shares Sought	63.71	56.58	29.43
% of Target Shares Purchased	33.18	19.77	36.35

Table 2: **Success classification variables: correlation matrix, frequency of success, and other descriptive statistics.** Five success classification variables are used. Three of these are binary variables: *SUCCESS* equals one if the bidder acquires at least one share of the target and zero otherwise; *OUTCOME* equals one if the bidder acquires at least as many shares as were sought and zero otherwise; *CONTROL* equals one if the bidder holds at least 50% of the shares outstanding at the exercise of the tender offer and zero otherwise. The other two success measures are continuous variables: *PUR/SOUGHT* is the number of shares purchased as a fraction of the number sought in the tender offer bid; *OWN/OUTSTD* is the fraction of the target's total shares outstanding which the bidding firm owns at the exercise of the tender offer. In panel A, each entry in the matrix shows the Pearson correlation coefficient between two corresponding (row and column) success measures. Below each of these correlation coefficients is the number of bids for which information to calculate the two success measures is available. Panel B presents the frequency of success based on the three binary success classification variables. Panel C presents descriptive statistics for the two continuous classification variables.

A. Correlation Matrix

	SUCCESS	*OUTCOME*	*CONTROL*	*PUR/SOUGHT*	*OWN/OUTSTD*
SUCCESS	1.000				
	958				
OUTCOME	0.528	1.000			
	826	826			
CONTROL	0.776	0.276	1.000		
	547	540	547		
PUR/SOUGHT	0.712	0.653	0.556	1.000	
	826	826	540	826	
OWN/OUTSTD	0.848	0.306	0.930	0.627	1.000
	547	540	547	540	547

B. Frequency of Success

Successful	540(56.4%)	228(27.6%)	249(45.5%)		
Unsuccessful	418(43.6%)	598(72.4%)	298(54.5%)		
Total	958	826	547		

C. Descriptive Statistics

Mean				0.625	0.443
Median				0.709	0.400
Standard Deviation				0.750	0.400

Table 3: **Bidder abnormal returns classified by success and number of bidders.** Percentage cumulative abnormal bidder returns on days -1 and 0 are examined to see if there is any significant relationship between bidder returns and success in single-bidder and multiple-bidder contests. Success of a bid is measured by three variables: *SUCCESS* equals *Y* if the bidder acquires at least one share of the target and *N* otherwise; *OUTCOME* equals *Y* if the bidder acquires at least as many shares as were sought and *N* otherwise; *CONTROL* equals *Y* if the bidder holds at least 50% of the shares outstanding at the exercise of the tender offer and *N* otherwise. Single bidder contests are denoted by *SB*, while multiple bidder contests are denoted by *MB*. N indicates the number of bids in a particular category. The *p*-value indicates the probability that the null hypothesis, $CAR_{[-1,0]} = 0$, is correct. * indicates that the null hypothesis is rejected at the 10% significance level. Similarly, ** indicates rejection at the 5% level, and *** indicates rejection at the 1% level.

	N	$CAR_{[-1,0]}$	*Standard Error*	*p-value*	*i/j* → ↓	$\Pr\left(CAR^i_{[-1,0]} = CAR^j_{[-1,0]}\right)$ 1	2	3
A. Success measure is SUCCESS								
MB & N	261	0.325	0.311	0.296	1			
MB & Y	146	−1.278***	0.416	0.002	2	0.002***		
SB & N	157	−0.195	0.401	0.626	3	0.305	0.061*	
SB & Y †	394	0.459*	0.253	0.070	4	0.740	0.000***	0.168
Total	958							
B. Success measure is OUTCOME								
MB & N	281	−0.027	0.299	0.929	1			
MB & Y	66	−1.778***	0.617	0.004	2	0.011**		
SB & N	317	0.514*	0.281	0.068	3	0.188	0.001***	
SB & Y	162	0.032	0.394	0.936	4	0.906	0.014**	0.320
Total	826							
C. Success measure is CONTROL								
MB & N	150	0.772*	0.410	0.060	1			
MB & Y	65	−1.452**	0.623	0.020	2	0.003***		
SB & N	148	−0.132	0.413	0.750	3	0.121	0.078*	
SB & Y	184	0.329	0.370	0.375	4	0.422	0.014**	0.407
Total	547							
† BIDS_ONCE & Y	315	0.633**	0.283	0.026				
BIDS>ONCE & Y	79	−0.235	0.566	0.678				
Total	394							

Table 4: **Regression of bidder abnormal returns with success, number of bidders, time period, and medium of payment as independent variables.** Bidder return is the percentage cumulative abnormal return estimated on the *Wall Street Journal* announcement date and the previous day (i.e., $CAR_{[-1,0]}$). Each regression uses a different success measure. *SUCCESS* equals one if the bidder acquires at least one share of the target and zero otherwise; *OUTCOME* equals one if the bidder acquires at least as many shares as were sought and zero otherwise; *CONTROL* equals one if the bidder holds at least 50% of the shares outstanding at the exercise of the tender offer and zero otherwise. *PUR/SOUGHT* is the number of shares purchased as a fraction of the number sought in the tender offer bid; *OWN/OUTSTD* is the fraction of the target's total shares outstanding which the bidding firm owns at the exercise of the tender offer. *COMPETITION* equals one if there are multiple bidders in the contest and zero otherwise. The *PERIOD* variables are dummy variables. *MIXED* equals one if the method of payment chosen is a mixture of cash and stock and zero otherwise. Similarly, *STOCK* equals one if stock is exclusively used and zero otherwise. Parentheses contain asymptotic standard errors. * indicates that the null hypothesis is rejected at the 10% significance level. Similarly, ** indicates rejection at the 5% level, and *** indicates rejection at the 1% level.

	(1)	(2)	(3)	(4)	(5)
INTERCEPT	1.357***	1.195***	1.035*	0.975**	0.608***
	(0.150)	(0.320)	(0.606)	(0.416)	(0.082)
COMPETITION	−0.482	−0.970*	0.110	−0.783*	−0.178***
	(0.441)	(0.514)	(0.674)	(0.476)	(0.048)
SUCCESS	−0.239				
	(0.366)				
OUTCOME		−0.995*			
		(0.562)			
CONTROL			−0.621***		
			(0.035)		
PUR/SOUGHT				−0.371***	
				(0.145)	
OWN/OUTSTD					−0.742***
					(0.082)
PERIOD 7/68–12/80	−1.686***	−1.210	−1.091***	−1.059	−1.190***
	(0.122)	(0.877)	(0.419)	(0.903)	(0.117)
PERIOD 1/81–12/84	−1.803**	−1.408**	−1.190	−1.225*	−1.227***
	(0.832)	(0.585)	(0.983)	(0.717)	(0.136)
PERIOD 1/85–12/88	−0.584	0.012	0.405	0.089	0.952**
	(0.362)	(0.102)	(0.684)	(0.559)	(0.443)
MIXED	−0.359	1.275	1.572	1.201	2.163**
	(0.533)	(0.865)	(1.068)	(0.840)	(0.946)
STOCK	−0.956	−1.422**	−2.187**	−1.551*	−1.607*
	(0.858)	(0.644)	(0.892)	(0.850)	(0.872)
Number of Observations	958	826	547	826	547
R^2	0.018	0.027	0.045	0.026	0.049

Table 5: **Tests of preemptive bidding theory.** *First Bidder* is the average return to the first bidder at the first bid date in a multiple-bidder contest. *Single Bidder* is the average return to successful bidders in single-bidder, single-bid contests. Bidder return is the percentage cumulative abnormal return estimated on the *Wall Street Journal* announcement date and the previous day (i.e., $CAR_{[-1,0]}$). N indicates the number of observations in each category. The first three panels use different success classifications: $SUCCESS$ requires that the bidder acquire at least one share of the target; $OUTCOME$ requires that the bidder acquire at least as many shares as were sought; $CONTROL$ requires that the bidder hold at least 50% of target stock at the exercise of the tender offer. The p-value indicates the probability that the null hypothesis, $CAR_{[-1,0]} = 0$, is correct. * indicates that the null hypothesis is rejected at the 10% significance level. Similarly, ** indicates rejection at the 5% level, and *** indicates rejection at the 1% level.

	N	$CAR_{[-1,0]}$	Standard Error	*p*-value
A. First bidder (in multiple bidder contest) vs. successful single bidder (SUCCESS)				
First Bidder	107	1.341***	0.489	0.006
Single Bidder	315	0.633**	0.285	0.027
$\Pr\left(CAR^{First\ Bidder}_{[-1,0]} = CAR^{Single\ Bidder}_{[-1,0]}\right) = 0.212$				
B. First bidder (in multiple bidder contest) vs. successful single bidder (OUTCOME)				
First Bidder	107	1.341**	0.559	0.017
Single Bidder	128	0.143	0.511	0.780
$\Pr\left(CAR^{First\ Bidder}_{[-1,0]} = CAR^{Single\ Bidder}_{[-1,0]}\right) = 0.115$				
C. First bidder (in multiple bidder contest) vs. successful single bidder (CONTROL)				
First Bidder	107	1.341***	0.517	0.010
Single Bidder	147	0.498	0.447	0.266
$\Pr\left(CAR^{First\ Bidder}_{[-1,0]} = CAR^{Single\ Bidder}_{[-1,0]}\right) = 0.222$				
D. First bidder bidding before vs. after the Williams Act				
Before	4	5.598*	2.998	0.065
After	103	1.176*	0.591	0.049
$\Pr\left(CAR^{Before\ WA}_{[-1,0]} = CAR^{After\ WA}_{[-1,0]}\right) = 0.151$				

Table 6: **The effect of medium of payment on the number of bidders and bids in a contest.** Panel A reports Poisson regression results. The dependent variable *NUMBER OF BIDDERS* is the number of bidders following the initial bidder in each contest. The independent variables are medium of payment dummies for the first bid in each contest: $CASH = 1$ if bidder uses only cash; $STOCK = 1$ if bidder uses only stock. The Poisson regression uses a log-linear specification. The antilogs of the dummy variable coefficient estimates are multipliers which indicate the relative number of bidders expected (in excess of one), given the medium of payment, compared with mixed financing. *** indicates that the null hypothesis of a zero parameter value is rejected at the 1% level. Panel B presents the frequency of cash, stock and mixed offers for bids in single bidder and multiple bidder contests.

A. $NUMBER\ OF\ BIDDERS = e^{\beta_0+\beta_1 CASH+\beta_2 STOCK}$ (N=660)

Parameter	Estimate	Asymptotic Standard Error	p-value
β_0	-1.674***	0.289	0.000
β_1	0.761***	0.296	0.010
β_2	1.032***	0.365	0.005

B. Method of payment for bids in single and multiple bidder contests (N = 958)

	Cash		*Stock*		*Mixed*	
	Frequency	%	Frequency	%	Frequency	%
Single Bidder, Single Bid	306	83.4	16	4.4	45	12.2
Single Bidder, Multiple Bids	152	82.6	14	7.6	18	9.8
Multiple Bidders	348	85.5	37	9.1	22	5.4
All Bids	806	84.1	67	7.0	85	8.9

Table 7: **Tests of bidder abnormal returns for the first of two consecutive bidders in a multiple-bidder contest at the announcement dates of bids by the first and second bidders in the pair.** The percentage cumulative abnormal return is estimated on the *Wall Street Journal* announcement date and the previous day (i.e., $CAR_{[-1,0]}$). Panel A presents the abnormal return to the first bidder when it announces its bid (RET_ONE) and the abnormal return to this bidder when a competitor announces its bid in the contest (RET_TWO). The model in Panel B regresses RET_TWO on RET_ONE to verify whether there is a significant relationship between the two return series. Panel C adds the term $LATE \cdot RET\_ONE$ where $LATE$ equals one if the first bid in the pair is not the first bid in the tender offer contest, and zero otherwise. In Panels D, E and F we introduce three success measures into the regression. $SUCCESS$ equals one if the first bidder acquires at least one share of the target and zero otherwise; $OUTCOME$ equals one if the first bidder acquires at least as many shares as were sought, and zero otherwise; $CONTROL$ equals one if the first bidder holds at least 50% of the shares outstanding at the exercise of the tender offer and zero otherwise. The p-value indicates the probability that the null hypothesis, $CAR_{[-1,0]} = 0$ is correct. * indicates that the null hypothesis is rejected at the 10% significance level. Similarly, ** indicates rejection at the 5% significance level and *** indicates rejection at the 1% significance level.

Parameter	Estimate	Standard Error	p-value	N	R^2
A. The first bidder performance at two announcement dates					
Mean RET_ONE	1.010**	0.391	0.011	200	
Mean RET_TWO	0.570	0.432	0.189	200	
B. $RET\_TWO = \beta_0 + \beta_1 RET\_ONE$					
β_0	0.465	0.438	0.290	200	0.009
β_1	0.104	0.078	0.185		
C. $RET\_TWO = \beta_0 + \beta_1 RET\_ONE + \beta_2 LATE \cdot RET\_ONE$					
β_0	0.396	0.440	0.370	200	0.018
β_1	0.160*	0.089	0.073		
β_2	−0.246	0.187	0.190		
D. $RET\_TWO = \beta_0 + \beta_1 RET\_ONE + \beta_2 LATE \cdot RET\_ONE + \beta_3 SUCCESS$					
β_0	−0.235	0.462	0.612	200	0.079
β_1	0.143*	0.086	0.098		
β_2	−0.237	0.182	0.195		
β_3	4.178***	1.155	0.000		
E. $RET\_TWO = \beta_0 + \beta_1 RET\_ONE + \beta_2 LATE \cdot RET\_ONE + \beta_3 OUTCOME$					
β_0	−0.354	0.459	0.441	171	0.076
β_1	0.145*	0.084	0.088		
β_2	−0.166	0.181	0.359		
β_3	4.044***	1.286	0.002		
F. $RET\_TWO = \beta_0 + \beta_1 RET\_ONE + \beta_2 LATE \cdot RET\_ONE + \beta_3 CONTROL$					
β_0	0.089	0.401	0.223	105	0.055
β_1	0.051	0.068	0.752		
β_2	−0.247	0.158	0.121		
β_3	2.577*	1.326	0.055		

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" -- by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

212. Nils H. Hakansson. "Welfare Economics of Financial Markets." June 1991

213. Gerard Gennotte and Alan Jung. "Investment Strategies under Transaction Costs: The Finite Horizon Case." August 1991.

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

244. Hayne Leland and Gregory Connor. "Optimal Cash Management for Investment Funds." March 1995.

245. Marcus Berliant and Sankar De. "On Revelation of Private Information in Stock Market Economies." April 1995.

246. Sankar De, Mark Fedenia, and Alexander J. Triantis. "Effects of Competition on Bidder Returns." April 1995.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance

WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 247

Anatomy of an ARM: Index Dynamics and Adjustable Rate Mortgage Valuation

by

Richard Stanton

and

Nancy Wallace

April 1995

iber

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Anatomy of an ARM:
Index Dynamics and Adjustable Rate
Mortgage Valuation

Richard Stanton and Nancy Wallace

Haas School of Business
University of California, Berkeley

April 1995

Finance Working Paper #247

The authors gratefully acknowledge helpful comments and suggestions from Kerry Back, Dwight Jaffee, Steve Leroy, Richard Meese, and seminar participants at U.C. Berkeley and the 1994 meetings of the American Real Estate and Urban Economics Association. They also acknowledge financial assistance from the Berkeley Program in Finance, from the Dean Witter Foundation, from the U.C. Berkeley Committee on Research, and from the U.C. Berkeley Center for Real Estate and Urban Economics.

CCD
12-6-95

Abstract

This paper analyzes the dynamics of the commonly used indices for Adjustable Rate Mortgages, and systematically compares the effects of their time series properties on adjustable rate mortgage prepayment and value. Our ARM valuation methodology allows us simultaneously to capture the effects of the dynamics of the index, discrete coupon adjustment, and caps and floors. It allows us either to calculate an optimal prepayment strategy for mortgage holders, or to use an empirical prepayment function. We find that the dynamics of the ARM indices, including both their average levels and their speeds of adjustment to interest rate shocks, introduce significant variation in the value of the prepayment option across ARMs. Valuation methodologies that ignore the time series properties of the index with respect to current rates will therefore systematically misprice adjustable rate mortgages.

1 Introduction

Recent surveys of major thrifts and mortgage bankers (See *Inside Mortgage Finance*) indicate that there are twelve commonly used indices for adjustable rate mortgages in the U.S. This finding is a significant change from 1985 surveys by the Federal Home Loan Mortgage Corporation and the United States League of Savings Institutions. These found that the one year constant maturity Treasury index accounted for between two thirds and eighty percent of all ARM lending ([10], [1]). There is no available information on the index market share of the outstanding stock of ARMs or the periodic flow of ARM originations.

Despite the variety of available ARM indices, it is remarkable that most contingent claims ARM valuation strategies do not explicitly account for the time series characteristics of the underlying index. Instead, published ARM valuation models implicitly assume that the ARM coupon resets with the contemporaneous term structure (See for example Kau et al. [12], McConnell and Singh [15]). There are no previous studies that systematically compare the effects of the times series properties of different ARM indices on the valuation of ARMs in a contingent claims framework with endogenous prepayment.

The results of the Ott [20] ARM duration study, and numerous recent studies of the time series properties of EDCOFI,[2] suggest that the commonly used indices do not adjust instantaneously to changes in contemporaneous spot rates. The strength of this empirical evidence suggests that we should reconsider the instantaneous reset assumptions found in most ARM valuation models. These models should instead be based on empirically tested specifications of the time series properties of the indices relative to observed term structure data.

This paper compares the time series dynamics of the most commonly used ARM indices, includes these dynamics in an ARM valuation model, and determines their impact on prepayment option and mortgage value. We build upon techniques developed by Kau et al. [12], Kishimoto [13], and Stanton and Wallace [25]. We use finite difference techniques to solve the pricing equation, taking into account all the main contractual features of the ARM index. A major advantage of this strategy is that it allows us either to determine endogenously the optimal prepayment policy for mortgage holders, or to use an empirically derived prepayment function. A second advantage is that it enables us to price ARMs in which lags in the index interact with other contract elements, such as caps, to induce path dependence in the mortgages' cash flows.

The paper is in three sections. The empirical specification for ARM indices is discussed in section 2. Section 3 discusses the valuation methodology, and analyzes the effects of index dynamics, caps and margins on the value of ARM contracts. Section 4 concludes the paper.

2 Dynamics of ARM Indices

The ARM indices that dominate the market are:

1. The one year constant maturity Treasury yield,

2. The Federal Housing Finance Board (FHFB) national average contract interest rate,

[2] See, for example, Cornell [3], Crockett et al. [5], Hayre et al. [9], Nothaft [18], Nothaft and Wang [19], Passmore [21], Roll [24], Stanton and Wallace [25].

3. The Eleventh District Cost-of-Funds Index (EDCOFI),[3]

4. The five year Treasury note rate,

5. One year LIBOR.

The one year Treasury rate reflects the average yields of all existing Treasury securities with one year of maturity remaining. The yield is determined from the closing market bids on actively traded Treasury Securities in the over-the-counter market, as disclosed by the five leading U.S. government securities dealers. The index is computed as a weekly average, and the Federal Reserve Board publishes this yield in its weekly H-15 statistical release. The five year constant maturity index is computed in the same way for existing Treasury securities with five years of remaining maturity.

The FHFB contract interest rate is the weighted average of initial mortgage interest rates paid by home buyers for loans originated during the first five business days of every month. The weights are determined by the type, size and location of the lender. The index is constructed by the Federal Housing Finance Board and reported on a monthly basis. The Eleventh District Cost-of-funds Index is computed from the book values of liabilities for all insured savings and loan (S&L) institutions in the Eleventh District (institutions in California, Nevada, and Arizona). The index is the ratio of the month-end total interest expenses on savings accounts, advances, and purchased funds to the average book value of these liabilities from the beginning to the end of the month. The ratio is adjusted with an annualizing factor so that the interest expenses are comparable across months.

The historical values of the ARM indices from July 1981 through May 1993 are plotted against the 3-month Treasury rate in Figure 1.[4] The plot shows that EDCOFI and the FHFB average contract rate display considerably less volatility than the Treasury and LIBOR series. EDCOFI appears to lag the Treasury series by several months. This should be expected, given that it is based on book yields, which can only change when a liability matures. The FHFB average contract rate looks rather like EDCOFI, with a spread of approximately 200 basis points. The FHFB average contract also lags the Treasury rates.

Considering the construction of EDCOFI and the FHFB average contract rate, and our plots of these indices relative to market rates, a partial adjustment model[5] is a reasonable representation for the movements of EDCOFI, the FHFB average contract rate, and one year LIBOR. For a given index, I_t, the model can be written as

$$I_t = \alpha + \beta r_t + \gamma I_{t-1} + \epsilon_t, \tag{1}$$

where r_t is an instantaneous spot rate, and ϵ_t is an error term. The coefficient β indicates the effect of the spot rate on the index each period, and γ indicates the speed at which the index adjusts. The extremes in the adjustment dynamics would be $\beta = 0$, where the index does not move at all with market rates, and $\gamma = 0$, where the index moves perfectly with the spot rate (the usual implicit assumption).

[3]EDCOFI is the ratio of the month end total interest expenses on savings accounts, advances, and purchased funds to the average book value of these liabilities from the beginning to the end of the month, calculated for all insured savings and loan (S&L) institutions in the Eleventh District (California, Nevada, and Arizona).

[4]All of the data series, except EDCOFI, were obtained from CITIBASE. The EDCOFI series data were obtained from the Office of Thrift Supervision.

[5]See Ott [20], Cornell [3], Passmore [21], Roll [24] and Stanton and Wallace [25] for further discussion and justification of this specification.

Ignoring the error term, if the index starts at a value I_0, and the interest rate remains at a constant level r, the value of the index at any later time is given by[6]

$$I_t = \left(1 - \gamma^t\right) \frac{\alpha + \beta r}{1 - \gamma} + \gamma^t I_0. \tag{2}$$

This is a weighted average of the long run value of I_t and its initial value. The speed of convergence is governed by the value of γ. The half-life, the number of periods required to reach half way between the two values, is the solution to

$$\gamma^{t_{1/2}} = \frac{1}{2}, \tag{3}$$

which is

$$t_{\frac{1}{2}} = -\frac{\log(2)}{\log(\gamma)}. \tag{4}$$

Note that substituting $\gamma = 0$ into equation 2 yields the correct result for instantaneous adjustment,

$$I_t = \alpha + \beta r. \tag{5}$$

Because we are interested in the adjustment of observed ARM indices to the instantaneous spot rate, we estimate the partial adjustment models using the three month Treasury rate as a proxy for the instantaneous spot rate.[7] The estimation results are reported in Table 1. All the indices are estimated in levels.[8] We estimate the partial adjustment model for EDCOFI using dummies for January and February to account for seasonality. Because of problems with both serial correlation and heteroscedasticity, we estimate the partial adjustment model for EDCOFI using the Newey and West [17] instrumental variable procedure to obtain a heteroscedasticity and autocorrelation consistent covariance matrix. The partial adjustment model for the FHFB average contract rate was estimated using OLS, because neither heteroscedasticity nor serial correlation violations was observed. The one year LIBOR partial adjustment model was estimated using instruments for the first lag of one year LIBOR, and then using Yule-Walker estimation methods and an AR(2) specification for second stage estimation of the model. The R^2, Breusch-Pagan [2] tests for heteroscedasticity, and the Durbin tests for AR(1) errors are also reported.

---

[6] This solution can be verified by inserting it into equation 1, with ϵ_t set to zero.

[7] This choice was made because the three month Treasury rates offered the shortest term rate with reasonable large trading volume. The one month Treasury rates reflect very erratic trading volume over our period of analysis

[8] Augmented Dickey-Fuller [6] tests of the form

$$\Delta x_t = \mu + \gamma^* x_{t-1} + \sum_{j=1}^{p-1} \phi_j \Delta x_{t-j} + \epsilon_t$$

were performed on all series using twelve lagged differences to control for possible seasonality. We were unable to reject the null hypothesis that there are unit roots in market rates and in the indices. Phillips and Perron [23] nonparametric unit root testing procedures were also applied, with the same result. Tests for the cointegration of the indices and Treasury rates, using Johansen [11], showed that they are not cointegrated. However, because the series are relatively short, and it is well known that the low power of standard unit roots tests often leads to acceptance of the null hypothesis of a unit root in many economic time series (Kwiatkowski et al. [14], Faust [8]), we rely on our strong priors that our interest rate series are mean reverting rather than explosive, and undertake all our estimation in levels of interest rates.

EDCOFI responds a little more quickly to the three month Treasury rate than the FHFB average contract rate. One year LIBOR responds faster than either, keeping very close track of the three month Treasury rate. This is shown in Figure 2, which shows the effect of an instantaneous 1% shift in the riskless interest rate on each of the three indices. Each index starts at its long run level (the level it would reach if r stayed at 7.5% for ever), and the graph shows what happens when r jumps from 7.5% to 8.5%. Besides the obvious lags in EDCOFI and the FHFB rate, one other interesting feature of the graph is the difference between the levels of the three series. EDCOFI is approximately 0.6% higher than r in this region; LIBOR is approximately 1.2% higher, and the FHFB rate is almost 3% higher. This means that for a given margin, a loan based on the FHFB rate will have significantly higher payments (and hence value) than one based on either of the other two series. The pricing impact of the lags in the indices' adjustment processes can only be evaluated in an option pricing framework that accounts for the adjustment frequency, caps and prepayment characteristics of ARM contracts.

3 Valuation

This section develops an algorithm for valuing adjustable rate mortgages. The algorithm can handle all of the important features of the contract, including the model for movements in EDCOFI developed in section 2. We can either use an empirical prepayment function (as commonly used in Wall Street valuation models), or derive endogenously the optimal prepayment strategy for mortgage holders. This latter strategy has the advantage that it is robust to possible changes in the economic environment, such as changes in the interest rate process, which would have an unquantifiable effect on an empirical prepayment function. In addition, using the optimal prepayment strategy allows us to determine an upper bound for the value of the prepayment option possessed by mortgage holders. The algorithm is based on techniques developed by Kau et al. [12], Kishimoto [13], and Stanton and Wallace [25].

3.1 Main Features of an ARM Contract

Coupon rate, C_t. The coupon rate on an ARM changes at each "reset date". The coupon determines the monthly cash flows on the mortgage until the next reset date. The monthly cash flow equals that on a fixed rate mortgage with the same time to maturity, same remaining principal balance, and same coupon rate as the ARM.
Underlying Index, I_t. The adjustment rule for the coupon rate specifies a particular index to which the rate is tied.
Margin, m. At each coupon reset date, the new rate is set by adding a margin, m (e.g. 2%), to the prevailing level of the underlying index (subject to certain caps, discussed below).
"Teaser" rate, C_0. It is common for the initial coupon rate to be lower than the "fully indexed" rate given by adding the margin to the initial level of the index. The initial rate, C_0, is often referred to as a "teaser" rate.
Annual cap, Δ. ARM contracts usually specify a maximum adjustment in the coupon rate at each reset period (e.g. 2% per year).
Lifetime caps, $\overline{C}$ and $\underline{C}$. ARM contracts usually specify an overall maximum coupon rate over the life of the loan, $\overline{C}$ (e.g. the initial rate plus 6%), and a minimum coupon rate over the life of the

loan, $\underline{C}$.

Reset Frequency. The coupon rate on an ARM contract adjusts at prespecified intervals. This interval is usually every 6 months or one year. In this paper, we assume yearly adjustment. If month t is a coupon reset date, the new coupon rate is given by

$$C_t = \max\left[\underline{C},\, C_{t-1} - \Delta,\, \min\left[I_t + m,\, C_{t-1} + \Delta,\, \overline{C}\right]\right] \tag{6}$$

3.2 Interest Rates

To value the mortgage, we need to make assumptions about the process governing interest rate movements. We use the Cox, Ingersoll and Ross [4] one-factor model. In this model, the instantaneous risk-free interest rate, r_t, satisfies the stochastic differential equation

$$dr_t = \kappa(\mu - r_t)\, dt + \sigma\sqrt{r_t}\, dz_t. \tag{7}$$

This equation says that, on average, the interest rate r converges toward the value μ. The parameter κ governs the rate of this convergence. The volatility of interest rates is $\sigma\sqrt{r_t}$. One further parameter, λ, which summarizes risk preferences of the representative individual, is needed to price interest rate dependent assets.

The parameter values used here are those estimated by Pearson and Sun [22], using data from 1979–1986. These values are

$$\begin{aligned} \kappa &= 0.29368, \\ \mu &= 0.07935, \\ \sigma &= 0.11425, \\ \lambda &= -0.12165. \end{aligned}$$

The long run mean interest rate is 7.9%. Ignoring volatility, the time required for the interest rate to drift half way from its current level to the long run mean is $\ln(1/2)/(-\kappa) \approx 2.4$ years.

3.3 Other factors affecting ARM value

The value of an ARM depends not only on the current interest rate, r_t, but on the whole path of interest rates since its issue. This determines the current coupon rate, C_t, the current level of the underlying index, I_t (which in turn determines future movements in the coupon rate), and the current remaining principal balance, F_t. These three variables summarize all relevant information about the history of interest rates. By adding these as extra state variables, we return to a Markov setting where all prices can be written as a function only of the current values of a set of underlying state variables.

Write B_t for the value of a non-callable bond which makes payments equal to the promised payments on the ARM. The mortgage holder's position can be decomposed into a short position in B_t (the scheduled payments on the mortgage) plus a long position in a call option on B_t, with (time varying) exercise price F_t. Writing M_t for the market value of the mortgage, and O_t for the value of the prepayment option, we have

$$M_t = B_t - O_t \tag{8}$$

Since B_t does not depend on the mortgage holder's prepayment decision, minimizing his or her liability value is equivalent to maximizing the value of the prepayment option, O_t. Write

$$\begin{aligned} B_t &\equiv B(r_t, I_t, C_t, F_t, t), && (9)\\ O_t &\equiv O(r_t, I_t, C_t, F_t, t). && (10)\end{aligned}$$

All values are homogeneous of degree one in the current remaining principal amount, F_t. Thus, if each month we value a mortgage with $1 remaining principal, we can scale up or down as necessary for different principal amounts. Define normalized asset values (values per $1 of remaining principal) by

$$\begin{aligned} \hat{B}_t &= B_t/F_t, && (11)\\ &\equiv \hat{B}(r_t, I_t, C_t, t). \\ \hat{O}_t &= O_t/F_t, && (12)\\ &\equiv \hat{O}(r_t, I_t, C_t, t). \end{aligned}$$

3.4 Valuation with one State Variable

Given the interest rate model defined by equation 7, write $V(r,t)$ for the value of an asset whose value depends only on the current level of r_t and time, and which pays coupons or dividends at some rate $\delta(r_t, t)$. This value satisfies the partial differential equation[9]

$$\frac{1}{2}\sigma^2 r V_{rr} + [\kappa\mu - (\kappa + \lambda)r]\, V_r + V_t - rV + \delta = 0, \tag{13}$$

which can be solved for V, subject to appropriate boundary conditions.

Natural boundaries for the interest rate, r, are 0 and ∞. Rather than working directly with r, define the variable y by

$$y = \frac{1}{1+\gamma r}. \tag{14}$$

for some constant $\gamma > 0$,[10] The infinite range $[0, \infty)$ for r maps onto the finite range $[0, 1]$ for y. The inverse transformation is

$$r = \frac{1-y}{\gamma y}. \tag{15}$$

Equation 14 says that $y = 0$ corresponds to "$r = \infty$" and $y = 1$ to $r = 0$. Next, rewrite equation 13 using the substitutions

$$\begin{aligned} U(y,t) &\equiv V(r(y), t), \qquad \text{so} && (16)\\ V_r &= U_y \frac{dy}{dr}, && (17)\\ V_{rr} &= U_y \frac{d^2y}{dr^2} + U_{yy}\left(\frac{dy}{dr}\right)^2, && (18)\end{aligned}$$

[9] We need to assume some technical smoothness and integrability conditions (see, for example, Duffie [7]).

[10] The larger the value of γ, the more points on a given y grid correspond to values of r less than, say, 20%. Conversely, the smaller the value of γ, the more points on a given y grid correspond to values of r greater than, say, 4%. We are most interested in values of r in an intermediate range. Therefore, as a compromise between these two competing objectives, we choose $\gamma = 12.5$. The middle of the range, $y = 0.5$, then corresponds to $r = 8\%$.

to obtain

$$\frac{1}{2}\gamma^2 y^4 \sigma^2 r(y) U_{yy} + \left(-\gamma y^2 \left[\kappa\mu - (\kappa + \lambda) r(y)\right] + \gamma^2 y^3 \sigma^2 r(y)\right) U_y + U_t - r(y) U + \delta = 0. \quad (19)$$

We can solve equation 19 using a finite difference algorithm. Finite difference algorithms replace derivatives with differences, and approximate the solution to the original partial differential equation by solving the set of difference equations that arise. We use the Crank-Nicholson algorithm.[11]

Represent the function $U(y, t)$ by its values on the finite set of points,

$$y_j = j\,\Delta y, \quad (20)$$

$$t_k = k\,\Delta t, \quad (21)$$

for $j = 0, 1, \ldots, J$, and for $k = 0, 1, \ldots, K$. Δy and Δt are the grid spacings in the y and t dimensions respectively. $\Delta y = 1/J$, and Δt is chosen for convenience to be one month, making a total of 360 intervals in the time dimension. Write

$$U_{j,k} \equiv U(y_j, t_k), \quad (22)$$

for each (j, k) pair. The Crank-Nicholson algorithm rewrites equation 19 in the form

$$MU_k = D_k, \quad (23)$$

where M is a tridiagonal matrix, U_k is the vector $\{U_{0,k}, U_{1,k}, \ldots, U_{I,k}\}$, and D_k is a vector whose elements are functions of $U_{j,k+1}$. This system of equations relates the values of the asset for different values of y at time t_k to its possible values at time t_{k+1}. To perform the valuation, we start at the final time period, when all values are known, and solve equation 23 repeatedly, working backwards one period at a time.

3.5 Extension to multiple state variables

In general, when asset prices depend on more than one state variable plus time, solution of the resultant partial differential equation becomes numerically burdensome. In this case, the additional variables, I_t and C_t, are functions of the path of interest rates, and so they introduce no additional risk premia. This allows us to extend the Crank-Nicholson finite difference algorithm to handle the multiple state variable case. The extensions required are:

1. Allow values to depend on C_t and I_t as well as r_t and t, allowing for dependence between the processes governing movements in these variables.

2. Scale values to correspond to \$1 remaining principal.

3. Handle caps, floors and teaser rates.

[11] See, for example, McCracken and Dorn [16].

In addition to the finite sets of values for y and t defined above, define a finite set of values for I and C by

$$\begin{aligned} I_l &= l\Delta I, & (24)\\ C_m &= m\Delta C, & (25) \end{aligned}$$

for $l = 0, 1, \ldots, L$, and for $m = 0, 1, \ldots, M$. ΔI and ΔC are the grid spacings in the I and C dimensions respectively. We are now solving for values on the points of a 4-dimensional grid, whose elements are indexed by the values of (j, k, l, m). Write the value of an asset whose cash flows depend on these state variables as

$$U_{j,k,l,m} \equiv U(y_j, t_k, I_l, C_m), \tag{26}$$

for each (j, k, l, m). I and C are functions of the path of interest rates. Over the next instant, the movement in r completely determines the movements in both I and C. Assume that movements in EDCOFI are described by the equation

$$I_{t+1} = g(I_t, r_{t+1}), \tag{27}$$

so that EDCOFI this month is a deterministic function of EDCOFI last month, plus the short term riskless rate this month (the models estimated above are of this type). Define l^* by

$$\begin{aligned} I_{l^*_{1,j,l,m}} &\approx g(I_l, r_{j+1}), & (28)\\ I_{l^*_{0,j,l,m}} &\approx g(I_l, r_j), & (29)\\ I_{l^*_{-1,j,l,m}} &\approx g(I_l, r_{j-1}). & (30) \end{aligned}$$

In words, l^* gives the closest index to the value of I next period given the current values of r, I and C, and three possible values of r next period (up, the same, and down). Assuming that next month is a coupon reset date (since otherwise, the coupon rate next month will just be the same as the coupon rate this month), define m^* similarly, to give the index of C next period given the current values of r, I and C, and the value of r next period. m^* is determined by the interplay between the current coupon C_t, the index I_t, the margin m, and the caps $\overline{C}$, $\underline{C}$ and Δ. Note that the effects of caps, floors and teaser rates are all automatically captured in this definition of m^*.

We can now generate a set of finite difference equations for each pair (l, m). For example, the approximation for the time derivative now becomes

$$U_t(y_j, t_k, I_l, C_m) \approx \left(U_{j,k+1,l^*_{0,j,l,m},m} - U_{j,k,l,m}\right) / \Delta t, \tag{31}$$

if t_{k+1} is *not* a coupon reset period, and

$$U_t(y_j, t_k, I_l, C_m) \approx \left(U_{j,k+1,l^*_{0,j,l,m},m^*_{0,j,l,m}} - U_{j,k,l,m}\right) / \Delta t, \tag{32}$$

if t_{k+1} *is* a coupon reset period. This allows us to write down one set of systems of equations like equation 23 for each (l, m) pair. These equations are independent of each other, so we can solve them for each (l, m) pair in turn, looping over l and m to calculate values at every grid point at time t_k. A simplified version of this is shown graphically in Figure 3. Each horizontal plane corresponds to a grid of values in (r,t) space. There is a separate such grid for each value of I_t (as shown in the figure), and for each value of C_t.[12] As in the standard Crank-Nicholson algorithm, we value the asset by solving a set of difference equations, just like equation 23, for each (r,t) plane. The difference equation for the value of the asset at any point involves its values at six points, corresponding to the current time, t, and the following time period, $t+1$, and interest rates r_i, r_{i-1} and r_{i+1}. Note that the values at time t all sit on the current (r,t) plane, while the values for next period may be on other planes, from equations 31 and 32 (for example, in Figure 3, if the interest rate moves from r_i to r_{i+1} next period, the index moves from I_j to I_{j+1}. Similarly, if the interest rate moves from r_i to r_{i-1} next period, the index moves from I_j to I_{j-1}). We can solve the equations for each (r,t) plane separately, rather than having to consider them all simultaneously. This is because the interaction between different (r,t) planes only occurs in the values at date $t+1$. By the backward nature of the solution methodology, when we are calculating values at time t, we can regard all values at date $t+1$ as known, so this only affects the calculation of the right hand side of equation 23.

The final step in the process is to deal with the normalization of asset prices to correspond to a remaining principal balance of $1. This is possible because, at any time, we know exactly how much principal will be repaid over the next one month. Given a coupon rate C_t and a current remaining principal F_t, the usual amortization formula tells us the value of F_{t+1}, regardless of any possible movements in r_t, I_t or C_t. The values stored in the grid for next period correspond to $1 in remaining principal *next* period. These need only to be multiplied by F_{t+1}/F_t (a function only of C_t) to make them correspond to $1 of remaining principal today.

3.6 Valuation Results

The extended Crank-Nicholson algorithm described above was used to value 30 year adjustable rate mortgages. Starting in month 360, the algorithm works backward to solve equation 23 one month at a time, calculating the normalized bond value, $\widehat{B}_t$. For the option, the same process gives the value conditional on its remaining unexercised for the next month. This value must then be compared with the option's intrinsic value ($\max[0, \widehat{B}_t - 1]$), to determine whether prepayment is optimal. $\widehat{O}_t$ is set to the higher of these two values, and the mortgage value is calculated from the relationship

$$\widehat{M_t} = \widehat{B}_t - \widehat{O}_t.$$

Figures 4–15 show the results for different underlying indices and different contract terms. For ease of comparison, every mortgage shown has an annual coupon reset frequency, and a lifetime cap of 13.5%. Figure 4 shows the values of the bond (the promised coupon payments, with no prepayment option), the mortgage, and the prepayment option, for different values of the interest rate r, with the current value of the FHFB rate set to 8.5%. The coupon rate adjusts annually to equal the prevailing value of the FHFB rate (with no additional margin). The prepayment option has a value of at least 4% of the remaining principal balance on the loan. Its value decreases as interest rates increase. Figure 5 shows the values of the bond, the mortgage, and the prepayment option for different values of the FHFB rate, keeping the riskless interest rate equal to 7.5%. The higher the current value of the FHFB rate (and the current coupon rate on the mortgage), the higher the value of the underlying instrument and the value of the prepayment option. There is a discontinuity in the

[12] Imagine Figure 3 repeated in the direction perpendicular to the page.

slope of the graph at the point where the FHFB rate equals 13.5%. Below this value, as the FHFB rate increases, the graph shows the value of mortgages with progressively higher values of *both* the underlying index *and* a higher index. At 13.5% the cap becomes binding, and from then on, while the index increases, the coupon rate remains fixed at 13.5%. The graphs of both bond value and option value are almost flat beyond this point. This is because further increasing the FHFB rate does not increase the coupon rate, merely the expected time before the cap ceases to bind.

Figures 6–15 focus on the impact on the value of the prepayment option of the index used, and the size of the reset margin. Figures 6 and 7 look again at mortgages based on the FHFB rate, with reset margins ranging from -0.5% to 1% over the FHFB rate. For each reset margin, the value of the prepayment option decreases in r for a given value of the FHFB rate, and for a given r the value increases in the FHFB rate, almost flattening off after 13.5%. The value increases in the reset margin, but note that in both Figures 6 and 7, the values for different reset margins for high values of r, or high values of the FHFB rate, appear to converge or even cross slightly.[13] The reason for this is that once the cap is binding, due to the high mean level of the FHFB rate and its slow movement relative to shifts in the term structure, it is likely to stay binding for a long time. With a binding coupon cap, the size of the margin is irrelevant; all mortgages have a coupon rate of 13.5%, and the value of the underlying bond is almost independent of the reset margin. The value of the prepayment option is high. For a margin of 1%, and a value of 7.5% for r, the prepayment option is worth at least 10% of the remaining principal balance on the mortgage. This is a function of the slow movement in the index, and also the generally high level of this index relative to the other indices studied (see Figures 1 and 2).

Figures 8 and 9 show the value of the prepayment option contained in ARMs based on EDCOFI. Comparing these graphs with those for FHFB loans, the spread between option values for loans with different reset margins is larger for loans based on EDCOFI. This is for two reasons. First, EDCOFI reacts faster to movements in r than does the FHFB rate. Second, its mean level is lower. Together, these imply that even if the coupon cap is binding today, it is likely not to do so in the near future, with the result that different margins imply significantly different bond, and hence option, values. The overall value of the options is lower than for FHFB, for the same reset margin. This is because FHFB is in general about 2% higher than EDCOFI (see Figure 1). Prepayment option values for EDCOFI loans with a 2% margin, and FHFB loans with a 0% margin, are similar across much of the range of possible index values. For high levels, the FHFB option is more valuable, as the effect of slow movements in the underlying index becomes more important.

Figures 10 and 11 show the value of the prepayment option for LIBOR based loans. Figures 12 and 13 show its value for ARMs based on the one year T-Bill rate, and Figures 14 and 15 look at loans based on the five year T-Note rate. The graphs for LIBOR and the one year T-Bill rate look very similar. This is because LIBOR tracks the short term interest rate closely (see Table 1 and Figure 1). The values for LIBOR loans are higher, since LIBOR is on average higher than the one year T-Bill rate (see Figure 2). For a given reset margin, the prepayment options contained in loans based on the five year T-Note rate are rather more valuable. This is because the five year rate is in general substantially higher than the one year rate, so the underlying bond is more valuable, and prepayment is more likely to be optimal (see Figure 1).

---

[13] The lines crossing is an artifact of the discrete approximation to the true asset value.

4 Summary

This paper analyzes the valuation of adjustable rate mortgages based on the most commonly used indices,

1. The one year constant maturity Treasury yield,
2. The Federal Housing Finance Board (FHFB) national average contract interest rate,
3. The Eleventh District Cost-of-Funds Index (EDCOFI),
4. The five year Treasury note rate,
5. One year LIBOR.

We find that a simple partial adjustment model closely describes the behavior of EDCOFI, the FHFB average contract rate, and one year LIBOR, and we develop an ARM valuation methodology which allows us simultaneously to capture the effects of index dynamics, discrete coupon adjustment, and caps and floors. Our methodology allows us either to calculate an optimal prepayment strategy for mortgage holders, or to use an empirical prepayment function. We use the conduct a systematic comparison of the properties of ARMs based on the different indices, and show in particular that the prepayment options embedded in most of these ARMs usually have significant value, a fact which is often overlooked. More generally, we find that the value of the prepayment option, and hence of the mortgage, is significantly affected by both its contract terms and by the dynamics of the index underlying the mortgage. This includes both the average level of the index[14] and its speed of adjustment to interest rate shocks.[15] Our valuation methodology allows us for the first time to quantify these interacting effects. Previous approaches, which ignore the time series properties of the index with respect to current rates, will systematically misprice adjustable rate mortgages.

[14]For example, the mean level of the FHFB rate is higher than that of EDCOFI, leading to higher expected payments on a loan backed by the FHFB rate, all else being equal.

[15]For example, while the mean level of EDCOFI is not very different from that of the one year Treasury rate, its slow adjustment to changes in interest rates mean that the prepayment option embedded in EDCOFI based ARMS are more valuable than those embedded in ARMs indexed to the one year Treasury rate, all else being equal.

References

[1] D. Andruhonis. ARMs surge reveals continuing trend toward standardization. Report 3, Freddie Mac, 1985.

[2] T. S. Breusch and A. R. Pagan. A simple test for heteroskedasticity and random coefficient variation. *Econometrica*, 47:1287–1294, 1979.

[3] B. Cornell. Forecasting the eleventh district cost of funds. *Housing Finance Review*, 6:123–135, 1987.

[4] J. C. Cox, J. E. Ingersoll, and S. A. Ross. A theory of the term structure of interest rates. *Econometrica*, 53:385–467, 1985.

[5] J. H. Crockett, F. E. Nothaft, and G. H. K. Wang. Temporal relationships among adjustable rate mortgage indexes. *Journal of Real Estate Finance and Economics*, 4:409–419, 1991.

[6] D. A. Dickey and W. A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Society*, 74:427–431, 1979.

[7] D. Duffie. *Security Markets: Stochastic Models*. Academic Press, Boston, 1988.

[8] J. Faust. Near observational equivalence and unit root processes: Formal concepts and implications. Working paper, International Finance Division, Board of Governors of the Federal Reserve System, Washington D.C., March 1993.

[9] L. S. Hayre, V. Lodato, and E. Mustafa. Eleventh district cost of funds index composition and modeling. *Journal of Fixed Income*, 3:38–44, 1991.

[10] D. Jacobs, N. Fahey, and M. Wilson. The 1985 study of adjustable rate mortgages being made at savings institutions. Technical report, United States League of Savings Institutions, 1985.

[11] S. Johansen. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12:231–255, 1988.

[12] J. B. Kau, D. C. Keenan, W. J. Muller, III, and J. F. Epperson. The valuation and analysis of adjustable rate mortgages. *Management Science*, 36:1417–1431, 1990.

[13] N. Kishimoto. A simplified approach to pricing path dependent securities. Working paper, Duke University, 1990.

[14] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin. Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54:159–178, 1992.

[15] J. J. McConnell and M. K. Singh. Valuation and analysis of collateralized mortgage obligations. Working paper, Purdue University, 1990.

[16] D. McCracken and W. Dorn. *Numerical Methods and FORTRAN Programming*. John Wiley, New York, 1969.

[17] W. K. Newey and K. D. West. A simple, positive definite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica*, 55:703–708, 1987.

[18] F. E. Nothaft. COFI ARMs: Seeing into the future. *Mortgage Banking*, 50:43–47, 1990.

[19] F. E. Nothaft and G. H. K. Wang. Seasonal variations in cost-of-funds at thrift institutions. *AREUEA Journal*, 20:573–582, 1992.

[20] R. A. Ott. The duration of an adjustable-rate mortgage and the impact of the index. *Journal of Finance*, 41:923–933, 1986.

[21] S. W. Passmore. Econometric models of the eleventh district cost of funds index. *Journal of Real Estate Finance and Economics*, 6:175–188, 1993.

[22] N. D. Pearson and T. Sun. A test of the Cox, Ingersoll, Ross model of the term structure of interest rates using the method of maximum likelihood. Working paper, MIT, 1989.

[23] P. C. B. Phillips and P. Perron. Testing for a unit root in time series regression. *Biometrika*, 75:335–346, 1988.

[24] R. Roll. Adjustable rate mortgages: The indexes. *Housing Finance Review*, 6:137–152, 1987.

[25] R. H. Stanton and N. Wallace. ARM wrestling: Valuing adjustable mortgages indexed to the eleventh district cost of funds. Working paper, U. C. Berkeley, 1993. Forthcoming, AREUEA Journal.

Independent Variables	Dependent Variable		
	EDCOFI	FHFB Average Contract Rate	One Year LIBOR (Weekly Avg.)
Constant	.3306** (2.065)	.366*** (6.062)	.6688** (2.068)
January dummy	−.0632** (−2.094)		
February dummy	.1517*** (3.441)		
First lag of EDCOFI	−.8430*** (17.415)		
First lag of FHFB Avg. Contract Rate		.8966*** (91.979)	
First lag of One Year LIBOR			.1361** (2.622)
Three month T-Bill rate	.1263*** (3.539)	.0928*** (11.340)	.9148*** (16.002)
Autoregressive Parameters			$(1 - .969\,B + .214\,B^2)$ (−11.56) (2.55)
R^2	.994	.997	.857
Breusch-Pagan test for heteroscedasticity, χ^2_m	9.2	4.2	8.7
Durbin test for AR(1) (t-statistic)	.684	−.195	.803

* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Table 1: Estimates for Adjustment Models of ARM Indices, July 1981 – May 1993 (t-statistics in parentheses).

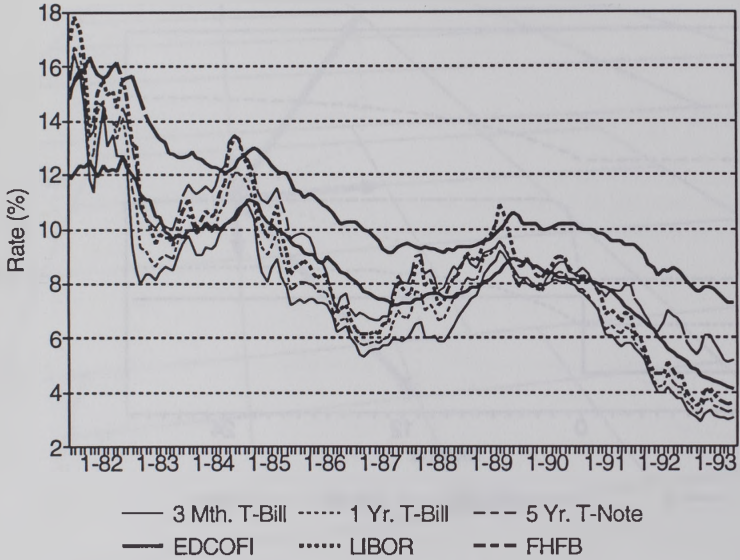


Figure 1: EDCOFI and other indices, July 1981 – May 1993.

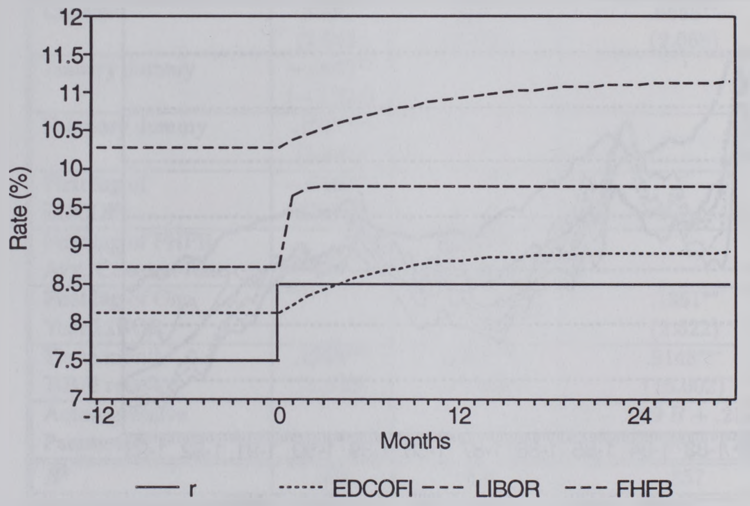


Figure 2: Example of the lags in the different indices' responses to movements in the term structure. The graph shows the movement in the different indices resulting from a jump in the short term riskless interest rate from 7.5% to 8.5%.

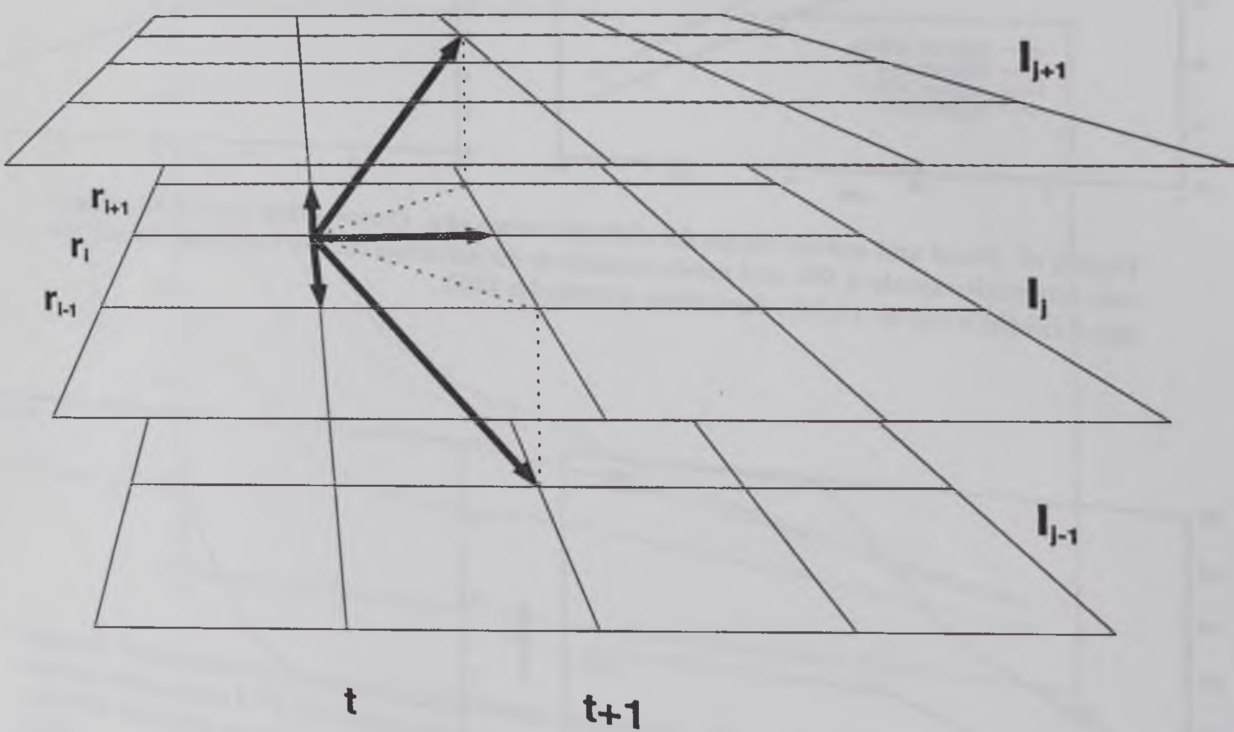


Figure 3: Extended Crank Nicholson algorithm.

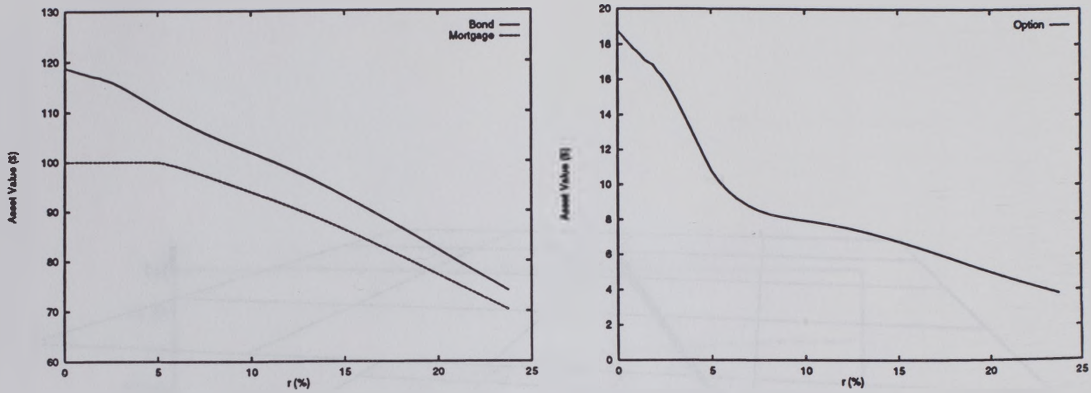


Figure 4: Bond and option values for different values of r. Current FHFB rate is 8.5%. Coupon rate currently equals 8.5%, and resets annually to the prevailing value of FHFB rate. Coupon rate has a lifetime cap of 13.5%. Remaining principal is \$100.

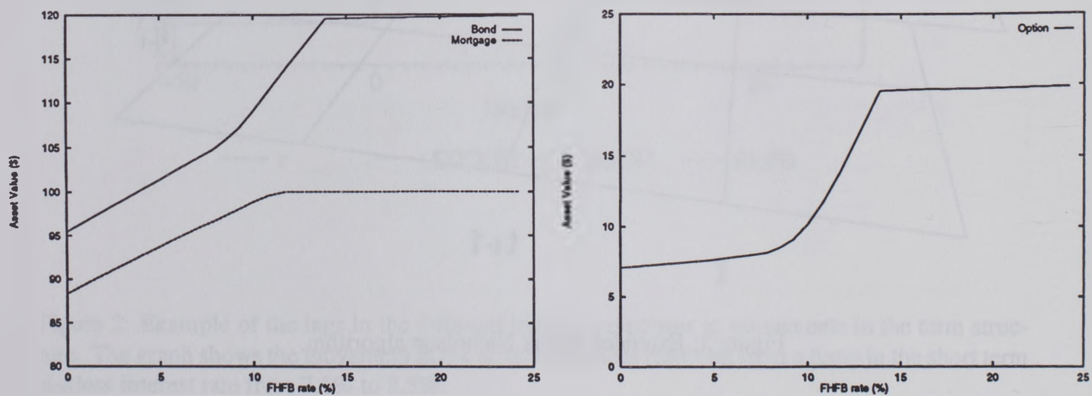


Figure 5: Bond and option values for different values of FHFB rate. Current short term riskless interest rate is 7.5%. Coupon rate equals current value of FHFB rate, and resets annually to the prevailing value of FHFB rate. Coupon rate has a lifetime cap of 13.5%. Remaining principal is \$100.

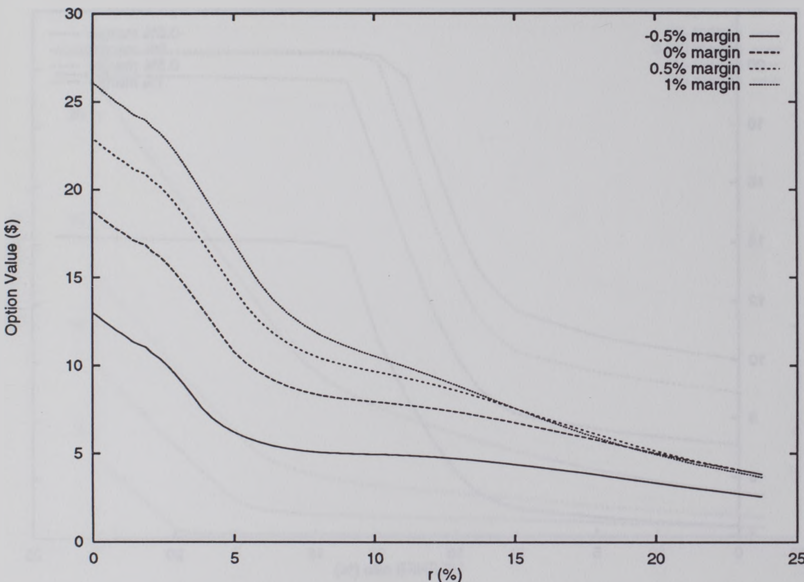


Figure 6: Prepayment option values for different values of r. Current FHFB rate is 8.5%. Coupon rate currently equals 8.5% plus appropriate margin. Coupon resets annually to the prevailing value of FHFB rate, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

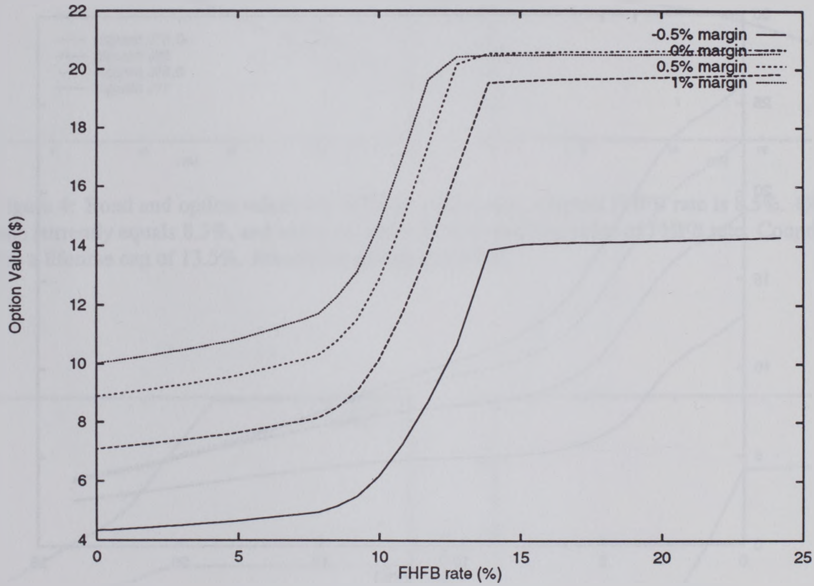


Figure 7: Prepayment option values for different values of FHFB rate. Current short term riskless interest rate is 7.5%. Coupon rate equals current value of FHFB rate, plus appropriate margin. Coupon resets annually to the prevailing value of FHFB rate, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

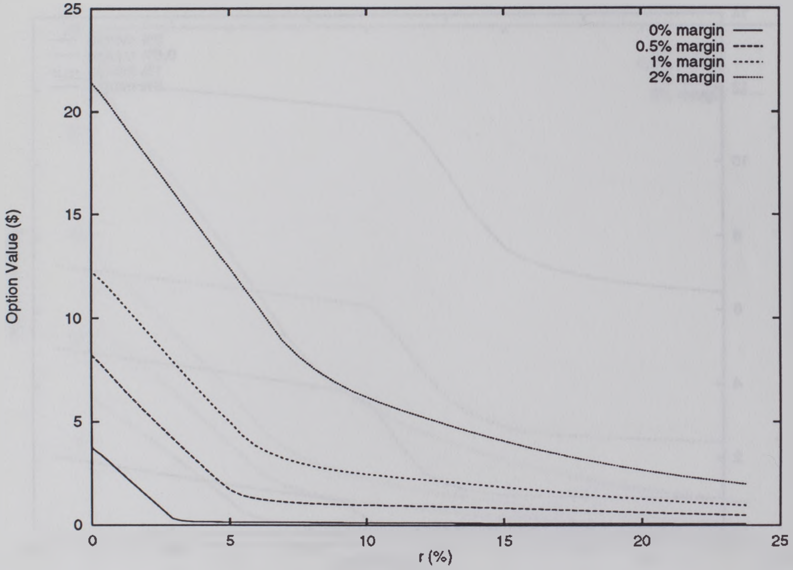


Figure 8: Prepayment option values for different values of r. Current EDCOFI is 8.5%. Coupon rate currently equals 8.5% plus appropriate margin. Coupon resets annually to the prevailing value of EDCOFI, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

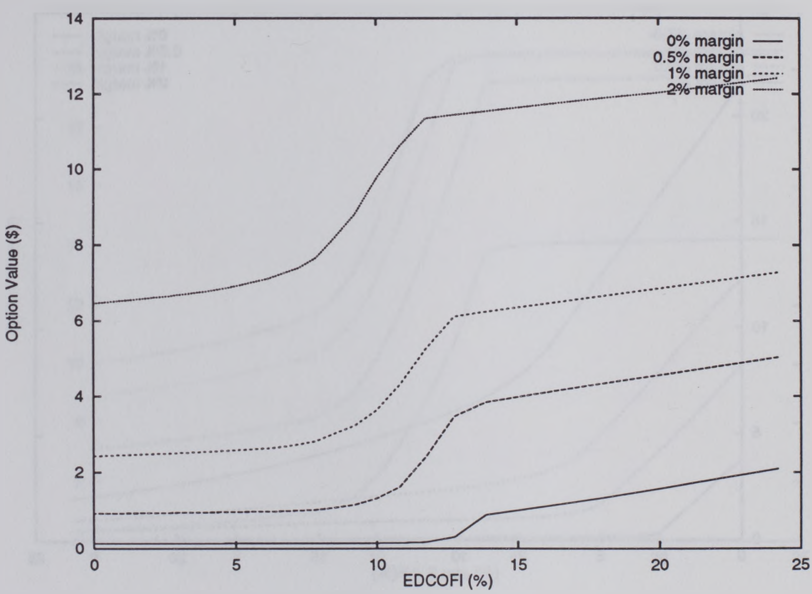


Figure 9: Prepayment option values for different values of EDCOFI. Current short term riskless interest rate is 7.5%. Coupon rate equals current value of EDCOFI, plus appropriate margin. Coupon resets annually to the prevailing value of EDCOFI, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

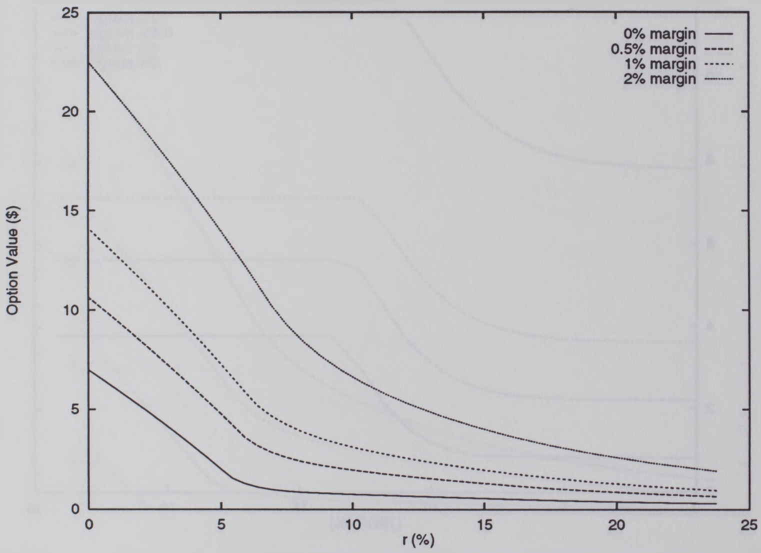


Figure 10: Prepayment option values for different values of r. Current LIBOR is 8.5%. Coupon rate currently equals 8.5% plus appropriate margin. Coupon resets annually to the prevailing value of LIBOR, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

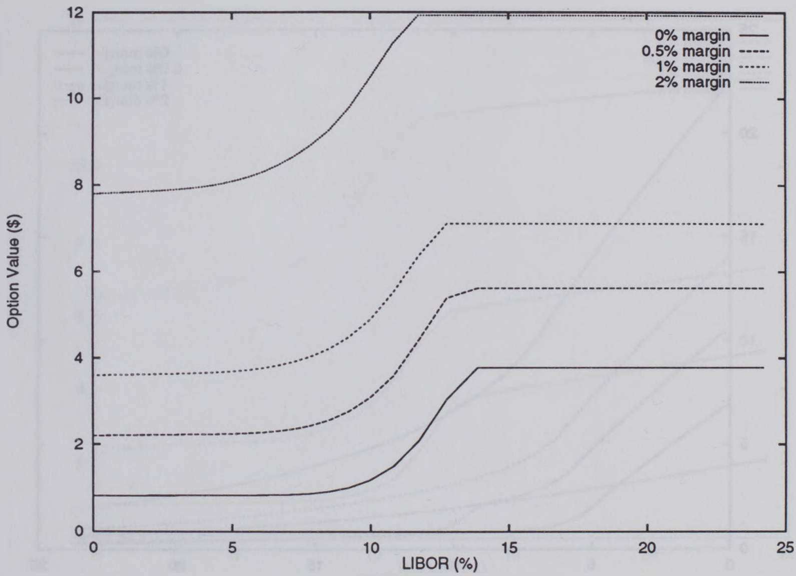


Figure 11: Prepayment option values for different values of LIBOR. Current short term riskless interest rate is 7.5%. Coupon rate equals current value of LIBOR, plus appropriate margin. Coupon resets annually to the prevailing value of LIBOR, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

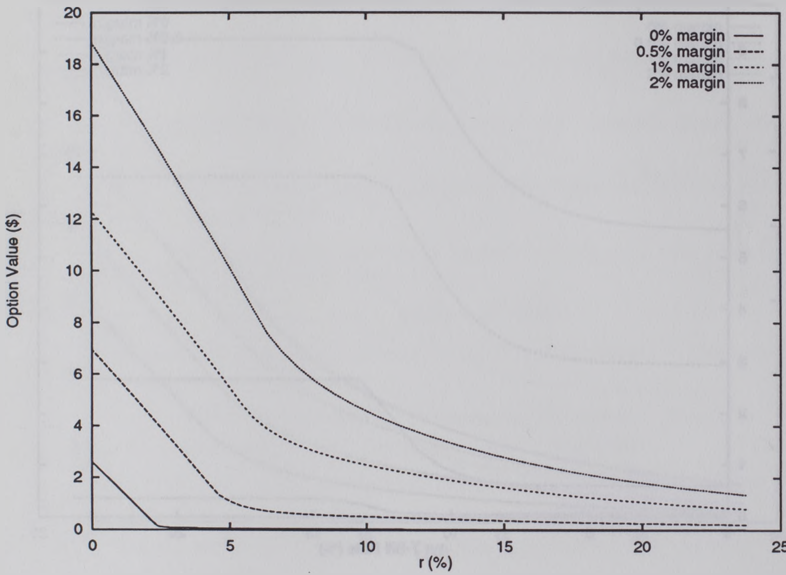


Figure 12: Prepayment option values for different values of r. Current coupon rate is 8.5% plus appropriate margin. Coupon resets annually to the prevailing value of one year T-Bill rate, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

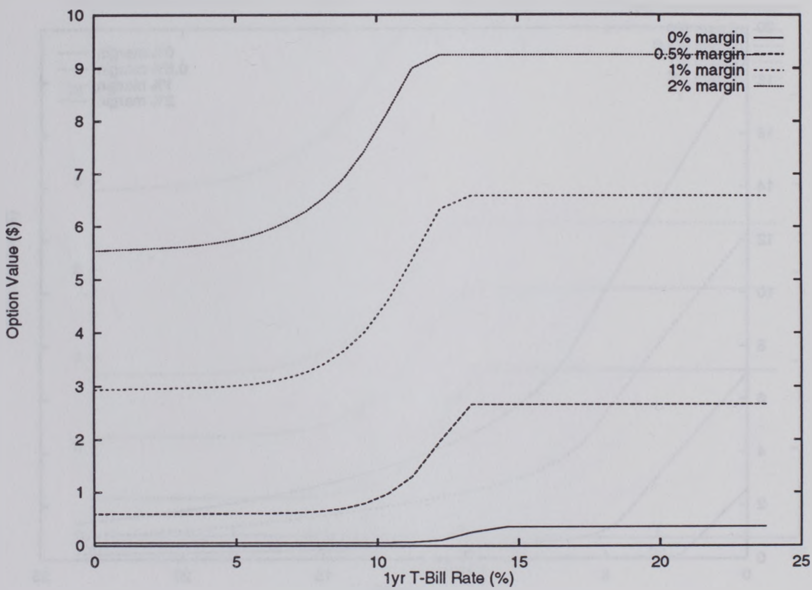


Figure 13: Prepayment option values for different values of one year T-Bill rate. Coupon rate equals current value of one year T-Bill rate, plus appropriate margin. Coupon resets annually to the prevailing value of one year T-Bill rate, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

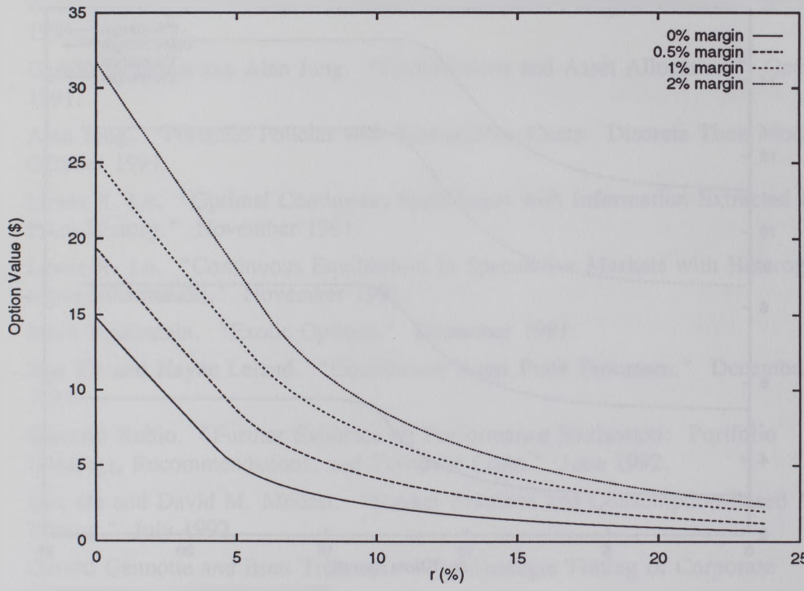


Figure 14: Prepayment option values for different values of r. Current coupon rate is 8.5% plus appropriate margin. Coupon resets annually to the prevailing value of five year T-Note rate, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is \$100.

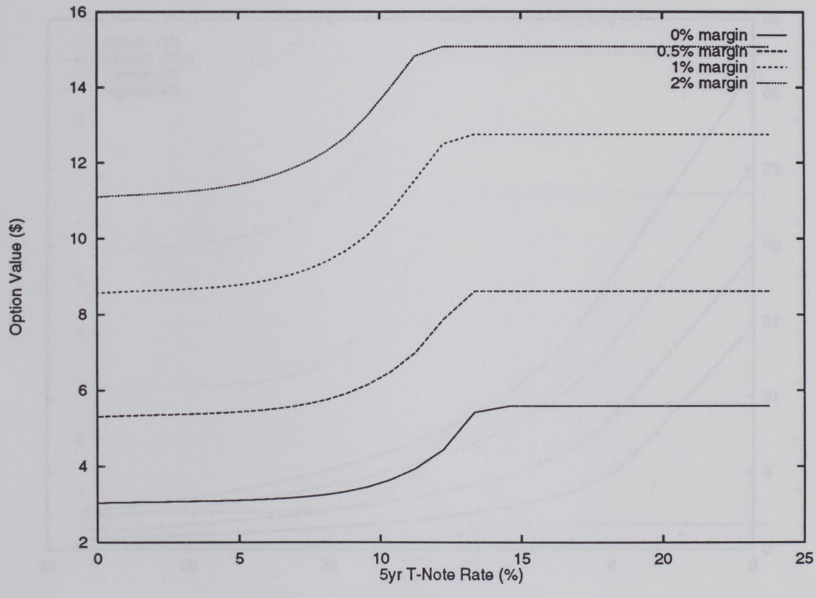


Figure 15: Prepayment option values for different values of five year T-Note rate. Coupon rate equals current value of five year T-Note rate, plus appropriate margin. Coupon resets annually to the prevailing value of five year T-Note rate, plus appropriate margin, subject to a lifetime cap of 13.5%. Remaining principal is $100.

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" -- by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

244. Hayne Leland and Gregory Connor. "Optimal Cash Management for Investment Funds." March 1995.

245. Marcus Berliant and Sankar De. "On Revelation of Private Information in Stock Market Economies." April 1995.

246. Sankar De, Mark Fedenia, and Alexander J. Triantis. "Effects of Competition on Bidder Returns." April 1995.

247. Richard Stanton and Nancy Wallace. "Anatomy of an ARM: Index Dynamics and Adjustable Rate Mortgage Valuation." April 1995.

248. Hua He, William P. Keirstead, and Joachim Rebholz. "Double Lookbacks." May 1995.

WALTER A. HAAS SCHOOL OF BUSINESS

Research Program in Finance

WORKING PAPER SERIES

FINANCE WORKING PAPER NO. 248

Double Lookbacks

by

Hua He

William P. Keirstead

and

Joachim Rebholz

May 1995

Institute of Business and Economic Research

University of California at Berkeley

RESEARCH PROGRAM IN FINANCE AT THE WALTER A. HAAS SCHOOL OF BUSINESS, UNIVERSITY OF CALIFORNIA AT BERKELEY

The Research Program in Finance in the Walter A. Haas School of Business at the University of California has as its purpose the conduct and encouragement of research in finance, investments, banking, securities markets, and financial institutions. The present reprint and working paper series were established in 1971 in conjunction with a grant from the Dean Witter Foundation.

INSTITUTE OF BUSINESS AND ECONOMIC RESEARCH

Richard Sutch, Director

The Institute of Business and Economic Research is a department of the University of California with offices on the Berkeley campus. It exists for the purpose of stimulating and facilitating research into problems of economics and of business with emphasis on problems of particular importance to California and the Pacific Coast, but not to the exclusion of problems of wider import.

Double Lookbacks

Hua He, William P. Keirstead
and Joachim Rebholz

May 1995

Finance Working Paper #248

An earlier version of the paper was written while Joachim Rebholz was at the Department of Mathematics, University of California, Berkeley. Financial support from the Sloan Foundation for (Hua He) is gratefully acknowledged.

Abstract

A new class of options, *double lookbacks*, where the payoffs depend on the maximum and/or minimum prices of one or two traded assets is introduced and analyzed. This class of double lookbacks includes calls and puts with the underlying being the difference between the maximum and minimum prices of one asset over a certain period, and calls or puts with the underlying being the difference between the maximum prices of two correlated assets over a certain period. Analytical expressions of the joint probability distribution of the maximum and minimum values of two correlated geometric Brownian motions are derived and used in the valuation of double lookbacks. Numerical results are shown, and prices of double lookbacks are compared to those of standard lookbacks on a single asset.

Hua He
Haas School of Business
University of California
Berkeley, California

William P. Keirstead
Haas School of Business
University of California
Berkeley, California

Joachim Rebholz
Fixed Income Trading Services
BARRA Inc.
Berkeley, California

Abstract

A new class of options, *double lookbacks*, where the payoffs depend on the maximum and/or minimum prices of one or two traded assets is introduced and analyzed. This class of double lookbacks includes calls and puts with the underlying being the difference between the maximum and minimum prices of one asset over a certain period, and calls or puts with the underlying being the difference between the maximum prices of two correlated assets over a certain period. Analytical expressions of the joint probability distribution of the maximum and minimum values of two correlated geometric Brownian motions are derived and used in the valuation of double lookbacks. Numerical results are shown, and prices of double lookbacks are compared to those of standard lookbacks on a single asset.

Hua He	*William P. Keirstead*	*Joachim Rebholz*
Haas School of Business	Haas School of Business	Fixed Income Trading Services
University of California	University of California	BARRA Inc.
Berkeley, California	Berkeley, California	Berkeley, California

1 Introduction

Exotic options designed as contingent claims on equity indices, currencies, and the term structure of interest rates, have achieved enormous success in global financial markets during the past decade. Such exotic products, while seemingly complicated to small investors, have provided institutional investors new vehicles to meet their various financial needs which include hedging, risk management and speculation (when investors have specific views on future market movements). Theoretical advances by academics as well as practitioners have helped market participants create more exotic products and understand the economic benefits of such products, and consequently, contributed in an important way to the surge in popularity of exotic options.

In this paper we introduce a new class of exotic options: lookback options based on two traded assets. A standard one-asset lookback call (or put) gives its holder the right to buy (or sell) the underlying stock at its historical minimum (or maximum) price over a certain period. Analytical solutions for standard lookback options have been found by Goldman, Sosin, and Gatto (1979) and Goldman, Sosin, and Shepp (1979). Lookback options are appealing because they offer investors the opportunity (at a price, of course) of buying a stock at its lowest price and selling a stock at its highest price.

In our two-asset generalization, we consider options whose payoffs depend on the extremal (i.e., maximum and/or minimum) prices of one and/or two stocks over a given period. For example, we consider call or put options on the spread between the maximum and minimum price of Xerox stock over a given interval of time; an option to receive the maximum of General Motors' stock price (or return) at the maximum of Ford's stock price (or return) over a given period; an option to receive the minimum of IBM's stock price (or return) at the minimum of Digital's stock price (or return) over a given period. We refer to these options as *double lookbacks.*

The economic motivation for double lookbacks is not difficult to perceive. An option on the spread between the maximum and minimum price of a single stock over a given interval of time captures in part the idea of an option on price volatility, and is conceptually simpler. Such an option might be of interest to traders who want to bet on price volatility or hedge an existing position which is sensitive to price volatility. Double lookbacks involving two assets allow investors to bet on the difference between the extreme values of two assets or two indices. Since the double lookback is an exchange between two extreme values, it is cheaper and therefore more attractive than a lookback option that exchanges one asset for the extreme value of that asset. If an investor wants to take a long position in the maximum of one asset and a short position in the value of another asset, then a *semi* double lookback, which is an option on the difference between the maximum price of one asset over a given period and the terminal price of another asset, would be an appropriate investment vehicle.

The main contribution of the paper is to derive analytical expressions for the distributions necessary to price double lookback options. In the case of a single asset, we derive the analytical expression for the probability distribution of the maximum and the minimum prices of one asset following a geometric Brownian motion. In the case of two assets, we derive the analytical expression for the probability distribution of the maximum and/or minimum prices of two assets following two correlated geometric Brownian motions.[1] Numerical procedures are readily available for evaluating all of the double lookback options discussed above.

It is important to point out that the analysis of this paper can also be used to value knockout options based on two traded assets. The analytical solution for the one-asset knockout is well known, see Merton (1973) and Rubinstein (1992). For a two-asset example, we may consider an option on the difference of two asset prices subject to a knockout condition based on either one or both assets' prices not reaching the boundary. We refer to these options as *double knockouts.*

The paper is organized as follows. Section 2 sets up the economy and lists the options in which we are interested. In Section 3, we derive the relevant distribution functions for lookbacks on a single asset, and in Section 4, we derive the distributions needed for lookbacks on two assets. Section 5 contains results on semi-lookbacks. Numerical examples are provided and compared in Sections 3-5. We discuss how our analysis can be used to value double knockouts in Section 6. Section 7 concludes.

2 Arbitrage-free Pricing of Lookbacks

Consider a Black-Scholes economy in which stock prices are log-normal, the interest rate is constant, and continuous trading without transaction costs, taxes, or other market frictions is permitted. There are three assets: one riskfree bond and two risky stocks. The prices of the bond and the stocks are:

$$\begin{aligned} B(t) &= e^{rt} \\ S_1(t) &= S_1(0)e^{(\mu_1 - q_1 - \frac{\sigma_1^2}{2})t + \sigma_1 w_1(t)} \\ S_2(t) &= S_2(0)e^{(\mu_2 - q_2 - \frac{\sigma_2^2}{2})t + \sigma_2 w_2(t)} \end{aligned}$$

where r is the riskless rate, μ_i the expected instantaneous return of stock i, q_i the dividend yield of stock i, σ_i the volatility of stock i, and w_i a standard Brownian motion with $\text{cov}(dw_1, dw_2) = \rho dt$.

[1] As far as we know, this distribution function has not appeared in the literature.

Throughout this paper, we assume that we are pricing the lookback options at date 0, the options expire at date T, and that the lookback period runs from t^* to T. Note that t^* may be either positive or negative. For $i = 1, 2$ and $t \geq t^*$, define the running minimum and maximum of stock price S_i by

$$\begin{aligned} \underline{S}_i(t) &= \min_{t^* \leq s \leq t} S_i(s) \\ \overline{S}_i(t) &= \max_{t^* \leq s \leq t} S_i(s). \end{aligned}$$

We are interested in studying several types of European double lookback options. The payoffs of these options at expiration date T are:

- (*Lookback Spread*) call or put on the spread between the maximum and minimum of a single stock price:

$$\max\left[0, \left(\overline{S}_1(T) - \underline{S}_1(T)\right) - K\right]$$
$$\max\left[0, K - \left(\overline{S}_1(T) - \underline{S}_1(T)\right)\right]$$

- (*Double Maxima*) call or put on the difference between the maximum of S_1 and the maximum of S_2:

$$\max\left[0, \left(a\overline{S}_1(T) - b\overline{S}_2(T)\right) - K\right]$$
$$\max\left[0, K - \left(a\overline{S}_1(T) - b\overline{S}_2(T)\right)\right]$$

  where $a > 0$ and $b > 0$ are parameters to be chosen by investors. In practice, if $t^* = 0$, it may make sense to pick a and b such that $aS_1(0) = bS_2(0)$. For example, $a = \frac{1}{S_1(0)}$ and $b = \frac{1}{S_2(0)}$. When $K = 0$, the double maxima call is equivalent to an option to buy the maximum of S_1 at the maximum of S_2.

- (*Double Minima*) call or put on the difference between the minimum of S_1 and the minimum of S_2:

$$\max\left[0, (a\underline{S}_1(T) - b\underline{S}_2(T)) - K\right]$$
$$\max\left[0, K - (a\underline{S}_1(T) - b\underline{S}_2(T))\right]$$

  When $K = 0$, the double minima call is equivalent to an option to sell the minimum of S_1 for the minimum of S_2.

- (*Double Lookback Spread*) call or put on the spread between the maximum S_1 and the minimum of S_2:

$$\max\left[0, \left(\overline{S}_1(T) - \underline{S}_2(T)\right) - K\right]$$
$$\max\left[0, K - \left(\overline{S}_1(T) - \underline{S}_2(T)\right)\right].$$

For comparison purposes, we will also look at the following options:

- option to buy the maximum of S_1 at $S_1(T)$ or $S_2(T)$:

$$\max\left[0, \overline{S}_1(T) - S_1(T)\right]$$

$$\max\left[0, a\overline{S}_1(T) - bS_2(T)\right]$$

  with $a = 1$ and $b = S_1(0)/S_2(0)$ if $t^* = 0$.

- option to sell S_1 or S_2 for the minimum of S_2:

$$\max\left[0, S_2(T) - \underline{S}_2(T)\right]$$

$$\max\left[0, aS_1(T) - b\underline{S}_2(T)\right]$$

  where $a = S_2(0)/S_1(0)$ and $b = 1$ if $t^* = 0$.

The premiums of these options should be compared with those of double maxima and double minima calls with $K = 0$.

Following Harrison and Kreps (1979) and Harrison and Pliska (1981), the Black-Scholes economy is known to be viable and dynamically complete, Thus, each of the above derivative securities can be replicated through dynamic trading in the stock(s) and bond. Furthermore, there exists a probability measure Q (the *equivalent martingale* or *risk-neutral* measure) under which the discounted price $V^*(t) = V(t)/B(t)$ of any derivative security is a martingale. Under this risk neutral probability, the stock price processes are:

$$\begin{aligned} S_1(t) &= S_1(0)e^{(r-q_1-\frac{\sigma_1^2}{2})t+\sigma_1 w_1^*(t)} \\ S_2(t) &= S_2(0)e^{(r-q_2-\frac{\sigma_2^2}{2})t+\sigma_2 w_2^*(t)} \end{aligned}$$

where w_1^* and w_2^* are standard Brownian motions under Q with the same constant correlation ρ as under the original probability measure. Because the discounted derivative price $V^*(t)$ is a martingale, its value at date 0 can be determined by taking the conditional expectation of its terminal value:

$$V^*(0) = \mathrm{E}^Q[V^*(T)].$$

For a Black-Scholes economy with constant interest rates, this equation can be rewritten as

$$V(0) = e^{-rT}\mathrm{E}^Q[V(T)].$$

Thus we see that each of the valuation problems we face consists of evaluating a conditional expectation, which in turn is simply a matter of integrating the payout over an appropriate density function.

If the lookback period begins at $t^* > 0$, then by iterating the expectation we find

$$V(0) = e^{-rT}\mathrm{E}^Q\left[\mathrm{E}^Q_{t^*}[V(T)]\right].$$

The inner expectation is of the same form as when $t^* \leq 0$, and hence can be evaluated using the same densities derived in the next section. The outer expectation is simply an integral over the distribution of $S(t^*)$, and can be evaluated using standard techniques. For the rest of this paper, we concentrate on the case where $t^* \leq 0$.

For ease of notation, we define $X_i(t)$, the continuously-compounded return of stock i, by

$$X_i(t) = \log S_i(t)/S_i(0) = \alpha_i t + \sigma_i w_i^*(t), \qquad (t^* \leq 0 \leq t)$$

where $\alpha_i = r - q_i - \sigma_i^2/2$. Also, define the running minima and maxima of X_i by

$$\begin{aligned}
m_i &= \min_{t^* \leq s \leq 0} X_i(s) \\
M_i &= \max_{t^* \leq s \leq 0} X_i(s) \\
\underline{X}_i(t) &= \min_{0 \leq s \leq t} X_i(s) \\
\overline{X}_i(t) &= \max_{0 \leq s \leq t} X_i(s).
\end{aligned}$$

Then the payoffs and valuation integrals for the double lookback options become:

- (*Lookback Spread*)

$$\begin{aligned}
V(T) &= \max\left[0, \left(S_1(0)e^{\max(M_1,\overline{X}_1(T))} - S_1(0)e^{\min(m_1,\underline{X}_1(T))}\right) - K\right] \\
V(0) &= e^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{P}(\underline{X}_1(T) \in d\underline{x}_1, \overline{X}_1(T) \in d\overline{x}_1)\, V(T)
\end{aligned}$$

  with corresponding formulas for the put option

- (*Double Maxima*)

$$\begin{aligned}
V(T) &= \max\left[0, aS_1(0)e^{\max(M_1,\overline{X}_1(T))} - bS_2(0)e^{\max(M_2,\overline{X}_2(T))} - K\right] \\
V(0) &= e^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{P}(\overline{X}_1(T) \in d\overline{x}_1, \overline{X}_2(T) \in d\overline{x}_2)\, V(T)
\end{aligned}$$

  with corresponding formulas for the put option

- (*Double Minima*)

$$\begin{aligned} V(T) &= \max\Big[0, aS_1(0)e^{\min(m_1,\underline{X}_1(T))} - bS_2(0)e^{\min(m_2,\underline{X}_2(T))} - K\Big] \\ V(0) &= e^{-rT}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{P}(\underline{X}_1(T) \in d\underline{x}_1, \underline{X}_2(T) \in d\underline{x}_2)\, V(T) \end{aligned}$$

  with corresponding formulas for the put option

- (*Double Lookback Spread*)

$$\begin{aligned} V(T) &= \max\Big[0, \Big(aS_1(0)e^{\max(M_1,\overline{X}_1(T))} - bS_2(0)e^{\min(m_1,\underline{X}_2(T))}\Big) - K\Big] \\ V(0) &= e^{-r(T-t)}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathcal{P}(\overline{X}_1(T) \in d\overline{x}_1, \underline{X}_2(T) \in d\underline{x}_2)\, V(T) \end{aligned}$$

  with corresponding formulas for the put option.

In the next section, we derive the necessary density functions. Performing the integrations indicated above, using numerical techniques if necessary, leads directly to the option prices.

3 Lookbacks on One Asset

In this section we consider lookback spread options where the payoffs depend upon the maximum and minimum of an asset whose price follows a geometric Brownian motion. Specifically, we derive the joint distribution function of the maximum and minimum of a Brownian motion with constant drift.

Before proceeding we state below as a lemma the distribution functions of the maximum and minimum of a Brownian motion with constant drift. We also record below as a theorem the analytical formulae for standard lookbacks based on either the maximum or the minimum of an asset over a given period. The proofs of these lemma and theorem can be found in Conze and Viswanathan (1991). Let us define

$$G(x_1, t; \alpha_1) \equiv N\left(\frac{x_1 - \alpha_1 t}{\sigma_1\sqrt{t}}\right) - e^{\frac{2\alpha_1 x_1}{\sigma_1^2}} N\left(\frac{-x_1 - \alpha_1 t}{\sigma_1\sqrt{t}}\right).$$

Lemma 1 *The distribution functions for the maximum and the minimum of a Brownian motion with constant drift α_1 is given by:*

$$\begin{aligned} \mathcal{P}(\overline{X}_1(t) \le x_1) &= G(x_1, t; \alpha_1), \quad x_1 \ge 0 \\ \mathcal{P}(\underline{X}_1(t) \ge x_1) &= G(-x_1, t; -\alpha_1), \quad x_1 \le 0 \end{aligned}$$

Theorem 1 *Let C_{LB} (or P_{LB}) be the price at time 0 of a standard lookback call (or put) which pays $[S_1(T) - \underline{S}_1(T)]$ (or $[\overline{S}_1(T) - S_1(T)]$) at the expiration date T. Then,*

$$\begin{aligned} C_{LB} &= Se^{-qT}N(d_1) - e^{-rT}mN(d_1 - \sigma\sqrt{T}) \\ &\quad + e^{-rT}\frac{\sigma^2}{2(r-q)}S^{1-\frac{2(r-q)}{\sigma^2}}m^{\frac{2(r-q)}{\sigma^2}}N(-d_1 + \frac{2(r-q)}{\sigma}\sqrt{T}) - \frac{\sigma^2}{2(r-q)}Se^{-qT}N(-d_1) \\ P_{LB} &= -Se^{-qT}N(-d_2) + e^{-rT}MN(-d_2 + \sigma\sqrt{T}) \\ &\quad - e^{-rT}\frac{\sigma^2}{2(r-q)}S^{1-\frac{2(r-q)}{\sigma^2}}M^{\frac{2(r-q)}{\sigma^2}}N(d_2 - \frac{2(r-q)}{\sigma}\sqrt{T}) + \frac{\sigma^2}{2(r-q)}Se^{-qT}N(d_2) \end{aligned}$$

where $S = S_1(0)$, $m = \underline{S}_1(0)$, $M = \overline{S}_1(0)$ and

$$\begin{aligned} d_1 &= \frac{\ln(S/m) + (r-q)T + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \\ d_2 &= \frac{\ln(S/M) + (r-q)T + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}} \end{aligned}$$

Similarly, let C_M (or P_m) be the price at time 0 of a call written on the maximum (or a put on the minimum) with a strike K. Then,

$$C_M = \begin{cases} Se^{-qT}N(d) - e^{-rT}KN(d - \sigma\sqrt{T}) & \\ -e^{-rT}\frac{\sigma^2}{2(r-q)}S^{1-\frac{2(r-q)}{\sigma^2}}K^{\frac{2(r-q)}{\sigma^2}}N(d - \frac{2(r-q)}{\sigma}\sqrt{T}) + \frac{\sigma^2}{2(r-q)}Se^{-qT}N(d), & \text{if } K > M \\ e^{-rT}(M-K) + Se^{-qT}N(d_2) - e^{-rT}MN(d_2 - \sigma\sqrt{T}) & \\ -e^{-rT}\frac{\sigma^2}{2(r-q)}S^{1-\frac{2(r-q)}{\sigma^2}}M^{\frac{2(r-q)}{\sigma^2}}N(d_2 - \frac{2(r-q)}{\sigma}\sqrt{T}) + \frac{\sigma^2}{2(r-q)}Se^{-qT}N(d_2), & \text{if } K < M \end{cases}$$

$$P_m = \begin{cases} -Se^{-qT}N(-d) + e^{-rT}KN(-d + \sigma\sqrt{T}) & \\ +e^{-rT}\frac{\sigma^2}{2(r-q)}S^{1-\frac{2(r-q)}{\sigma^2}}K^{\frac{2(r-q)}{\sigma^2}}N(-d + \frac{2(r-q)}{\sigma}\sqrt{T}) - \frac{\sigma^2}{2(r-q)}Se^{-qT}N(-d), & \text{if } K < m \\ e^{-rT}(K-m) - Se^{-qT}N(-d_1) + e^{-rT}mN(-d_1 + \sigma\sqrt{T}) & \\ +e^{-rT}\frac{\sigma^2}{2(r-q)}S^{1-\frac{2(r-q)}{\sigma^2}}m^{\frac{2(r-q)}{\sigma^2}}N(-d_1 + \frac{2(r-q)}{\sigma}\sqrt{T}) - \frac{\sigma^2}{2(r-q)}Se^{-qT}N(-d_1), & \text{if } K > m \end{cases}$$

where

$$d = \frac{\ln(S/K) + (r-q)T + \frac{1}{2}\sigma^2 T}{\sigma\sqrt{T}}$$

We note that standard lookback call and put prices are equal to the corresponding Black-Scholes values, with strikes set at the current maximum or minimum, plus a premium. This premium reflects the opportunity that the minimum (or maximum) can go down (or up) further. The prices for calls on the minimum and puts on the maximum have a similar interpretation.

We now derive a similar set of joint density/distribution functions for the maximum and minimum of a single Brownian motion.

Lemma 2 *(i) The joint density/distribution of the maximum, minimum, and end point of a Brownian motion with a constant drift, for* $x \in [x_1, x_2]$, $x_1 \le 0$, $x_2 \ge 0$, *is given by:*

$$\mathcal{P}(X_1(t) \in dx, \underline{X}_1(t) \ge x_1, \overline{X}_1(t) \le x_2)$$
$$= \exp\left(\frac{\alpha_1 x}{\sigma_1^2} - \frac{\alpha_1^2 t}{2\sigma_1^2}\right) \sum_{n=-\infty}^{\infty} \frac{1}{\sigma_1\sqrt{t}} \left[\phi\left(\frac{x - 2n(x_2 - x_1)}{\sigma_1\sqrt{t}}\right) - \phi\left(\frac{x - 2n(x_2 - x_1) - 2x_1}{\sigma_1\sqrt{t}}\right)\right] dx,$$

where $\phi(z) = \exp(-z^2/2)/\sqrt{2\pi}$ *is the standard normal density. This density can be expressed in the equivalent form*

$$\mathcal{P}(X_1(t) \in dx, \underline{X}_1(t) \ge x_1, \overline{X}_1(t) \le x_2)$$
$$= \frac{2}{x_2 - x_1} \exp\left(\frac{\alpha_1}{\sigma_1^2}x - \frac{\alpha_1^2}{2\sigma_1^2}t\right) \sum_{n=1}^{\infty} \exp\left(-\frac{n^2\pi^2\sigma_1^2 t}{2(x_2 - x_1)^2}\right) \sin n\pi\left(\frac{-x_1}{x_2 - x_1}\right) \sin n\pi\left(\frac{x - x_1}{x_2 - x_1}\right) dx.$$

(ii) The joint distribution of the maximum and minimum of a Brownian motion with constant drift, for $x_1 \le 0$, $x_2 \ge 0$, *is given by:*

$$\mathcal{P}(\underline{X}_1(t) \ge x_1, \overline{X}_1(t) \le x_2) \tag{1}$$
$$= \sum_{n=-\infty}^{\infty} e^{\frac{2n\alpha_1(x_2 - x_1)}{\sigma_1^2}} \left\{\left[N\left(\frac{x_2 - \alpha_1 t - 2n(x_2 - x_1)}{\sigma_1\sqrt{t}}\right) - N\left(\frac{x_1 - \alpha_1 t - 2n(x_2 - x_1)}{\sigma_1\sqrt{t}}\right)\right]\right.$$
$$\left. - e^{\frac{2x_1\alpha_1}{\sigma_1^2}}\left[N\left(\frac{x_2 - \alpha_1 t - 2n(x_2 - x_1) - 2x_1}{\sigma_1\sqrt{t}}\right) - N\left(\frac{x_1 - \alpha_1 t - 2n(x_2 - x_1) - 2x_1}{\sigma_1\sqrt{t}}\right)\right]\right\}.$$

This distribution can be written in the equivalent form

$$\mathcal{P}(\underline{X}_1(t) \ge x_1, \overline{X}_1(t) \le x_2) = e^{-\frac{\alpha_1^2 t}{2\sigma_1^2}} \sum_{n=1}^{\infty} \frac{2n\pi\left[e^{\frac{\alpha_1 x_1}{\sigma_1^2}} - (-1)^n e^{\frac{\alpha_1 x_2}{\sigma_1^2}}\right]}{n^2\pi^2 + \left[\frac{\alpha_1(x_2 - x_1)}{\sigma_1^2}\right]^2} e^{-\frac{n^2\pi^2\sigma_1^2 t}{2(x_2 - x_1)^2}} \sin\frac{n\pi(-x_1)}{x_2 - x_1}. \tag{2}$$

PROOF: To prove (i), the density function can be obtained by using a reflection principle argument. Karatzas and Shreve (1991) give the zero drift result, and our result just shifts by a Girsanov factor. Alternatively, one can obtain this result by solving the Fokker-Planck equation using a method of images procedure (see, for example, Wilmott, Dewynne, and Howison (1993)).

To get the second density function, define $g(x)\,dx = \mathcal{P}\left(\underline{X}_1(t) > x_1, \overline{X}_1(t) < x_2, X_1(t) \in dx\right)$. Then $g(x)$ satisfies the following Fokker-Planck equation with absorbing boundaries:

$$\frac{\partial g}{\partial t} = \frac{1}{2}\sigma_1^2\frac{\partial^2 g}{\partial x^2} - \alpha_1\frac{\partial g}{\partial x}$$

$$g(x, 0) = \delta(x), \quad g(x_1, t) = g(x_2, t) = 0$$

where $\delta(x)$ denotes the Dirac delta function with a spike at $x = 0$. A routine separation of variables technique leads directly to the answer given. See Gardiner (1990) for the zero-drift solution.

Integrating the expressions in part (i) over $x \in [x_1, x_2]$ leads immediately to the joint distributions in (ii) for the minimum and maximum of a Brownian motion. ■

Remark 1 *One should note that the two seemingly different formulas for the joint density/distribution are in fact equal: the second is the Fourier sine transform of the first. In practice, one or the other may lead to more useful numerical approximations depending upon the problem.*

In Figure 1, we present a surface and contour plot of the joint density function

$$-\frac{\partial^2 \mathcal{P}(\underline{X}_1(t) \geq x_1, \overline{X}_1(t) \leq x_2)}{\partial x_1 \partial x_2}.$$

The parameter values chosen are $\sigma = 0.2\ \text{yr}^{-1/2}$, $r = 0.05\ \text{yr}^{-1}$, $q = 0$, and $t = 1$ yr, and we plot the minimum over the range $[-0.7, 0]$ and the maximum over the range $[0, 0.7]$.

Given the joint density function, the price of lookback spread can be obtained by integrating the final payoffs with respect to the density. Specifically, define

$$V_{SP}(x_1, x_2) = \max\left[0, \left(S_1(0)e^{\max(M_1, x_2)} - S_1(0)e^{\min(m_1, x_1)}\right) - K\right].$$

Theorem 2 *The price at time 0 of a lookback spread call, C_{SP}, is given by*

$$C_{SP} = e^{-rT} \int_{-\infty}^{0} dx_1 \int_{0}^{\infty} dx_2\, V_{SP}(x_1, x_2)\, \frac{-\partial^2 \mathcal{P}(\underline{X}_1(t) \geq x_1, \overline{X}_1(t) \leq x_2)}{\partial x_1 \partial x_2} \tag{3}$$

with corresponding formulas for the put option.

Given Theorem 2, we can evaluate the price of a lookback spread option by direct numerical quadrature. In Table 1, we list the prices for lookback spread call and put options for various parameter values. These prices have the right sensitivities, i.e., higher volatility leads to higher option premiums and larger strike price leads to smaller call premium but larger put premium. When the volatility is 20%, the premium for the lookback spread $(\overline{S} - \underline{S})$ is 31.5% of the initial stock price.

One should note that by having an explicit form for the density function, the computational time to evaluate these options is significantly reduced compared to Monte-Carlo or lattice methods. Typically only a few terms in the infinite series of Lemma 2 are needed to obtain convergence.

σ	K	Call	Put
	0	16.21	0.00
0.1	15	3.11	1.17
	20	1.27	4.08
	30	0.18	12.50
	0	31.50	0.00
0.2	30	5.57	2.60
	35	3.56	5.35
	60	0.42	25.99
	0	47.09	0.00
0.3	45	8.46	4.18
	50	6.42	6.89
	100	0.60	48.63

Table 1: Values for call and put lookback spread options with payouts of, respectively, $(\overline{S}-\underline{S}-K)^+$ and $(K-\overline{S}+\underline{S})^+$. We assume that $S(0)=\overline{S}(0)=\underline{S}(0)=100$, $r=0.05$ yr^{-1}, $q=0$, $t=1$ yr.

4 Lookbacks on Two Assets

In this section we analyze double lookback options whose payoffs depend on the extreme values of two assets following correlated geometric Brownian motions. We first derive the joint distribution functions necessary for the evaluation of these options, and then we discuss the use of these distributions in pricing.

We begin with the following lemma for the joint distribution of the extreme values and terminal values of two correlated Brownian motions.

Lemma 3 *(i) For $x_1 \geq m_1$, $x_2 \geq m_2$, where $m_1 \leq 0$, $m_2 \leq 0$,*

$$\mathcal{P}(X_1(t) \in dx_1, X_2(t) \in dx_2, \underline{X}_1(t) \geq m_1, \underline{X}_2(t) \geq m_2) = p(x_1, x_2, t; m_1, m_2, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho)\, dx_1 dx_2,$$

where

$$p(x_1, x_2, t; m_1, m_2, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho) = e^{a_1x_1+a_2x_2+bt} q(x_1, x_2, t; m_1, m_2, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho),$$

$$q(x_1, x_2, t; m_1, m_2, \alpha_1, \alpha_2, \sigma_1, \sigma_2, \rho) = \frac{1}{\sigma_1\sigma_2\sqrt{1-\rho^2}} \frac{2}{\beta t} \sum_{n=1}^{\infty} e^{-\frac{r^2+r_0^2}{2t}} \sin\frac{n\pi\theta_0}{\beta} \sin\frac{n\pi\theta}{\beta} I_{\frac{n\pi}{\beta}}\left(\frac{rr_0}{t}\right),$$

and

$$a_1 = \frac{\alpha_1\sigma_2 - \rho\alpha_2\sigma_1}{(1-\rho^2)\sigma_1^2\sigma_2}$$

$$
\begin{aligned}
a_2 &= \frac{\alpha_2\sigma_1 - \rho\alpha_1\sigma_2}{(1-\rho^2)\sigma_1\sigma_2^2} \\
b &= -\alpha_1 a_1 - \alpha_2 a_2 + \frac{1}{2}\sigma_1^2 a_1^2 + \rho\sigma_1\sigma_2 a_1 a_2 + \frac{1}{2}\sigma_2^2 a_2^2 \\
\tan\beta &= -\frac{\sqrt{1-\rho^2}}{\rho}, \quad \beta \in [0,\pi], \\
z_1 &= \frac{1}{\sqrt{1-\rho^2}}\left[\left(\frac{x_1-m_1}{\sigma_1}\right) - \rho\left(\frac{x_2-m_2}{\sigma_2}\right)\right] \\
z_2 &= \left(\frac{x_2-m_2}{\sigma_2}\right) \\
z_{10} &= \frac{1}{\sqrt{1-\rho^2}}\left[-\frac{m_1}{\sigma_1} + \frac{\rho m_2}{\sigma_2}\right] \\
z_{20} &= -\frac{m_2}{\sigma_2} \\
r &= \sqrt{z_1^2 + z_2^2} \\
\tan\theta &= \frac{z_2}{z_1}, \quad \theta \in [0,\beta], \\
r_0 &= \sqrt{z_{10}^2 + z_{20}^2} \\
\tan\theta_0 &= \frac{z_{20}}{z_{10}}, \quad \theta_0 \in [0,\beta].
\end{aligned}
$$

(ii) For $x_1 \geq m_1$, $x_2 \leq M_2$, where $m_1 \leq 0$, $M_2 \geq 0$, we have

$$\mathcal{P}(X_1(t) \in dx_1, X_2(t) \in dx_2, \underline{X}_1(t) \geq m_1, \overline{X}_2(t) \leq M_2) = p(x_1, -x_2, t; m_1, -M_2, \alpha_1, -\alpha_2, \sigma_1, \sigma_2, -\rho)\, dx_1 dx_2.$$

(iii) For $x_1 \leq M_1$, $x_2 \leq M_2$, where $M_1 \geq 0$, $M_2 \geq 0$, we have

$$\mathcal{P}(X_1(t) \in dx_1, X_2(t) \in dx_2, \overline{X}_1(t) \leq M_1, \overline{X}_2(t) \leq M_2) = p(-x_1, -x_2, t; -M_1, -M_2, -\alpha_1, -\alpha_2, \sigma_1, \sigma_2, \rho)\, dx_1 dx$$

PROOF: For notational convenience, we denote the density as $p(x_1, x_2, t)$. We know that p satisfies the Fokker-Planck equation

$$\frac{\partial p}{\partial t} = -\alpha_1 \frac{\partial p}{\partial x_1} - \alpha_2 \frac{\partial p}{\partial x_2} + \frac{1}{2}\sigma_1^2 \frac{\partial^2 p}{\partial x_1^2} + \rho\sigma_1\sigma_2 \frac{\partial^2 p}{\partial x_1 \partial x_2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 p}{\partial x_2^2}$$

with the following initial condition

$$p(x_1, x_2, t=0) = \delta(x_1)\delta(x_2)$$

and absorbing boundary conditions

$$
\begin{aligned}
p(x_1 = m_1, x_2, t) &= 0 \\
p(x_1, x_2 = m_2, t) &= 0.
\end{aligned}
$$

We proceed to explicitly solve this PDE.

First, we note that we can eliminate the drift terms by the following transformation. Define

$$p(x_1, x_2, t) = e^{a_1 x_1 + a_2 x_2 + bt} q(x_1, x_2, t)$$

where a_1, a_2, and b are defined as above. Then $q(x_1, x_2, t)$ satisfies

$$\frac{\partial q}{\partial t} = \frac{1}{2}\sigma_1^2 \frac{\partial^2 q}{\partial x_1^2} + \rho\sigma_1\sigma_2 \frac{\partial^2 q}{\partial x_1 \partial x_2} + \frac{1}{2}\sigma_2^2 \frac{\partial^2 q}{\partial x_2^2}$$

with boundary conditions

$$\begin{aligned} q(x_1, x_2, t=0) &= \delta(x_1)\delta(x_2) \\ q(x_1 = m_1, x_2, t) &= 0 \\ q(x_1, x_2 = m_2, t) &= 0. \end{aligned}$$

Next, we note that this PDE can be simplified by a suitable transformation of coordinates, to eliminate the cross-partial derivative and normalize the Brownian motions. Explicitly, if we define new coordinates z_1 and z_2, as given above, then

$$q(x_1, x_2, t) = \frac{h(z_1(x), z_2(x), t)}{\sigma_1\sigma_2\sqrt{1-\rho^2}}$$

and $h(z_1, z_2, t)$ satisfies

$$\frac{\partial h}{\partial t} = \frac{1}{2}\left(\frac{\partial^2 h}{\partial z_1^2} + \frac{\partial^2 h}{\partial z_2^2}\right)$$

with boundary conditions

$$\begin{aligned} h(z_1, z_2, t) &= \delta(z_1 - z_{10})\delta(z_2 - z_{20}) \\ h(L_1, t) &= h(L_2, t) = 0, \end{aligned}$$

where

$$\begin{aligned} L_1 &= \{(z_1, z_2) : z_2 = 0\} \\ L_2 &= \left\{(z_1, z_2) : z_2 = -\frac{\sqrt{1-\rho^2}}{\rho} z_1\right\} \end{aligned}$$

These boundary conditions along L_1 and L_2 are more conveniently expressed in polar coordinates. Introducing polar coordinates (r, θ) corresponding to (z_1, z_2) as defined above, we obtain

$$\frac{\partial h}{\partial t} = \frac{1}{2}\left(\frac{\partial^2 h}{\partial r^2} + \frac{1}{r}\frac{\partial h}{\partial r} + \frac{1}{r^2}\frac{\partial^2 h}{\partial \theta^2}\right)$$

with boundary conditions

$$\begin{aligned} h(r,\theta,t=0) &= \frac{1}{r_0}\delta(r-r_0)\delta(\theta-\theta_0) \\ h(r,\theta=0,t) &= 0 \\ h(r,\theta=\beta,t) &= 0. \end{aligned}$$

To solve this PDE for $h(r,\theta,t)$, we look for separable solutions of the form $R(r)\Theta(\theta)T(t)$. Plugging this in to the PDE, we find

$$\frac{T'}{T} = \frac{1}{2}\left(\frac{R''}{R} + \frac{1}{r}\frac{R'}{R} + \frac{1}{r^2}\frac{\Theta''}{\Theta}\right) \equiv -\lambda^2/2$$

where the separation constant is negative because the solutions must decay as $t \to \infty$. Hence, we have

$$T(t) \sim e^{-\lambda^2 t/2}$$

and

$$\left(r^2\frac{R''}{R} + r\frac{R'}{R} + \lambda^2 r^2\right) + \left(\frac{\Theta''}{\Theta}\right) = 0.$$

Defining $\Theta''/\Theta = -k^2$, we find

$$\Theta(\theta) \sim A\sin k\theta + B\cos k\theta.$$

The boundary conditions require that $\Theta(0) = \Theta(\beta) = 0$, and hence k must be real, $B = 0$, and

$$\sin k\beta = 0.$$

This last requirement restricts k to discrete values of the form

$$k_n = \frac{n\pi}{\beta}, \quad n = 1, 2, \ldots.$$

Thus the most general angular solution consistent with the boundary conditions is

$$\Theta(\theta) \sim \sin\frac{n\pi\theta}{\beta}, \quad n = 1, 2, \ldots.$$

Finally, the radial part of the solution is

$$r^2 R'' + rR' + (\lambda^2 r^2 - k_n^2)R = 0.$$

Defining $y = \lambda r$, we can rewrite this in the standard form

$$y^2\frac{d^2R}{dy^2} + y\frac{dR}{dy} + (y^2 - k_n^2)R = 0.$$

This is Bessel's equation, with the well-known fundamental solutions $J_{k_n}(y)$ and $I_{k_n}(y)$. Since $I_{k_n}(0)$ diverges, and we require $R(0)$ to be well-behaved, the $I_{k_n}(x)$ solution is not permitted. Hence the general radial solution is

$$R(r) \sim J_{k_n}(\lambda r).$$

In summary, then, the most general solution to the PDE for $h(r,\theta,t)$ consistent with the absorbing boundary conditions $h(r,0,t) = h(r,\beta,t) = 0$, is given by

$$h(r,\theta,t) = \int_0^\infty \sum_{n=1}^\infty c_n(\lambda) e^{-\frac{\lambda^2 t}{2}} \sin\left(\frac{n\pi\theta}{\beta}\right) J_{\frac{n\pi}{\beta}}(\lambda r)\, d\lambda.$$

Our goal now is to find the coefficients $c_n(\lambda)$ which fit the initial condition $h(r,\theta,0) = r_0^{-1}\delta(r - r_0)\delta(\theta - \theta_0)$.

To find $c_n(\lambda)$, multiply the previous equation at $t = 0$ by $\sin\left(\frac{n\pi\theta}{\beta}\right)$ and integrate over θ. We find

$$r_0^{-1}\delta(r - r_0)\sin\left(\frac{n\pi\theta_0}{\beta}\right) = \frac{\beta}{2}\int_0^\infty d\lambda\, c_n(\lambda) J_{\frac{n\pi}{\beta}}(\lambda r).$$

Next, multiply this equation by $r J_{\frac{n\pi}{\beta}}(\lambda r)$ and integrate over r. Using the well-known completeness relation

$$\int_0^\infty x J_\nu(ax) J_\nu(bx)\, dx = a^{-1}\delta(a-b),$$

we find

$$c_n(\lambda) = \frac{2\lambda}{\beta}\sin\left(\frac{n\pi\theta_0}{\beta}\right) J_{\frac{n\pi}{\beta}}(\lambda r_0).$$

Plugging this expression back into the general formula for h, we find

$$h(r,\theta,t) = \int_0^\infty \left(\frac{2\lambda}{\beta}\right) \sum_{n=1}^\infty e^{-\frac{\lambda^2 t}{2}} \sin\left(\frac{n\pi\theta_0}{\beta}\right) \sin\left(\frac{n\pi\theta}{\beta}\right) J_{\frac{n\pi}{\beta}}(\lambda r_0) J_{\frac{n\pi}{\beta}}(\lambda r)\, d\lambda.$$

The λ integral can be performed explicitly using the fact that [Gradshteyn and Ryzhik, p. 718]

$$\int_0^\infty y e^{-c^2 y^2} J_\nu(ax) J_\nu(bx)\, dx = \frac{1}{2c^2} e^{-\frac{a^2+b^2}{4c^2}} I_\nu\left(\frac{ab}{2c^2}\right).$$

Doing so leads to the final expression that

$$h(r,\theta,t) = \left(\frac{2}{\beta t}\right) \sum_{n=1}^\infty e^{-\frac{r^2+r_0^2}{2t}} \sin\left(\frac{n\pi\theta_0}{\beta}\right) \sin\left(\frac{n\pi\theta}{\beta}\right) I_{\frac{n\pi}{\beta}}\left(\frac{r r_0}{t}\right).$$

This completes the proof of part (i). The proofs of parts (ii) and (iii) follow immediately by symmetry. ■

Remark 2 *For $\alpha_1 = \alpha_2 = 0$ and $\sigma_1 = \sigma_2 = 1$, our result coincides with Caslow (1959, p. 279 or p. 379).*

For certain special values of the correlation coefficient ρ, the distribution functions in Lemma 3 can be simplified. We state our result in the following lemma.

Lemma 4 *Suppose the same assumptions hold as in Lemma 3(i), except that the correlation ρ can take on only the special values*

$$\rho_n = -\cos\left(\frac{\pi}{n}\right), \quad n = 2, 3, \ldots.$$

Then the solution to the Fokker-Planck equation is

$$p(x_1, x_2, t) = e^{a_1 x_1 + a_2 x_2 + bt} \frac{h(z_1, z_2, t)}{\sigma_1 \sigma_2 \sqrt{1 - \rho_n^2}},$$

where h is a finite *sum of bivariate normal densities*

$$h(z_1, z_1, t) = \sum_{k=0}^{n-1} \left[g_k^+(z_1, z_2, t) + g_k^-(z_1, z_2, t) \right].$$

and

$$g_k^{\pm}(z_1, z_2, t) = \pm(2\pi)^{-1} \exp\left(-\frac{1}{2} \left[\left(z_1 - r_0 \cos\left(\frac{2k\pi}{n} \pm \theta_0\right)\right)^2 + \left(z_2 - r_0 \sin\left(\frac{2k\pi}{n} \pm \theta_0\right)\right)^2 \right] \right).$$

Proof of Lemma 4: Follow the proof of Lemma 3(i) until the PDE for $h(z_1, z_2, t)$ is derived. When $\rho_n = -\cos(\pi/n)$, note that the angle between the lines L_1 and L_2 in Lemma 3 takes one of the special values

$$\beta_n = \pi/n, \quad n = 2, 3, \ldots.$$

For these angles, a method of images solution to the PDE is possible, as follows. Note that

$$g^{\pm}(z_1, z_2, t; a_1, a_2) = \pm(2\pi)^{-1} \exp\left(-\frac{1}{2} \left[(z_1 - a_1)^2 + (z_2 - a_2)^2 \right] \right)$$

satisfies the PDE with initial condition

$$g^{\pm}(z_1, z_2, 0; a_1, a_2) = \pm\delta(z_1 - a_1)\delta(z_2 - a_2).$$

Furthermore, since the PDE is linear in h, any linear combination of these $g^{\pm}$'s, with different values of (a_1, a_2) also satisfies the PDE. We want to find that particular linear combination which also satisfies the boundary and initial conditions.

Consider the case $n = 3$, as illustrated in Figure 2. For this correlation value, we have $\beta_3 = \pi/3$. Let a circle enclosing a '+' or '−' represent the solution $g^{\pm}$, with the location of center of the symbol indicating the value of (a_1, a_2). The first hextant (shaded in Fig. 2 and enclosed by solid radii representing the rays L_1 and L_2) is the region $\theta \in [0, \pi/3]$, the region where we want to find a solution to the PDE. The '+' in the first hextant is located at (z_{10}, z_{20}), which makes an angle θ_0 with respect to the z_1-axis, and is at a distance r_0 from the origin. Since this is the only symbol in the first hextant, the delta function initial condition is satisfied. The other symbols are located as follows (each at distance r_0 from the origin): the '+' symbols occur at angles $\theta_0 + 2\pi/3$ and $\theta_0 + 4\pi/3$, and the '−' symbols occur at angles $-\theta_0$, $-\theta_0 + 2\pi/3$, and $-\theta_0 + 4\pi/3$. We claim that $h(z_1, z_2, t)$ is given by the sum of these six densities, each with unit weighting. As already seen, this linear combination satisfies the PDE and initial condition, and hence we only need to show that the absorbing boundary conditions are satisfied. But as is easily seen from the symmetry of the diagram, the six densities cancel in pairs along the rays L_1 and L_2. Hence the absorbing boundary conditions are satisfied, and the sum of these six Gaussians is the unique solution for $h(z_1, z_2, t)$.

The solution for other values of n follows in a similar fashion, leading to the result given.

■

Remark 3 *Note that the special correlation values in Lemma 4 are all negative. Thus for the double minima or double maxima densities, the result may be of limited usefulness, since we expect most assets to be positively correlated. For the density involving the minima of one asset and the maxima of another, the Lemma is more interesting, since it is applicable to assets with* positive *correlations of the form* $\cos(\pi/n)$.

For purposes of evaluating double lookback options, we need only the distribution of the terminal extreme values of the Brownian motions. These distributions are related to the above results by the following corollary:

Corollary 1 *(i) The joint distribution of the minima of two correlated Brownian motions, for* $m_1 \leq 0$, $m_2 \leq 0$, *is given by*

$$\mathcal{P}(\underline{X}_1(t) \geq m_1, \underline{X}_2(t) \geq m_2) = \int_{m_1}^{\infty} \int_{m_2}^{\infty} \mathcal{P}(X_1(t) \in dx_1, X_2(t) \in dx_2, \underline{X}_1(t) \geq m_1, \underline{X}_2(t) \geq m_2).$$

(ii) The joint distribution of the minimum of one Brownian motion and the maximum of another, for $m_1 \leq 0$, $M_2 \geq 0$, *is given by*

$$\mathcal{P}(\underline{X}_1(t) \geq m_1, \overline{X}_2(t) \leq M_2) = \int_{m_1}^{\infty} \int_{-\infty}^{M_2} \mathcal{P}(X_1(t) \in dx_1, X_2(t) \in dx_2, \underline{X}_1(t) \geq m_1, \overline{X}_2(t) \leq M_2).$$

(iii) The joint distribution of the maxima of two correlated Brownian motions, for $M_1 \geq 0$, $M_2 \geq 0$, *is given by*

$$\mathcal{P}(\overline{X}_1(t) \leq M_1, \overline{X}_2(t) \leq M_2) = \int_{-\infty}^{M_1} \int_{-\infty}^{M_2} \mathcal{P}(X_1(t) \in dx_1, X_2(t) \in dx_2, \overline{X}_1(t) \leq M_1, \overline{X}_2(t) \leq M_2).$$

Remark 4 *For the general density given in Lemma 3, we do not believe that the integrals in the Corollary 1 can be further simplified. However, by a simple factorization we can obtain an alternative way of expressing these integrals:*

$$\mathcal{P}(\underline{X}_1(t) \geq m_1, \underline{X}_2(t) \geq m_2) = e^{a_1 m_1 + a_2 m_2 + bt} f(r_0, \theta_0, t)$$

where

$$f(r_0, \theta_0, t) \equiv \frac{2}{\beta t} \sum_{n=1}^{\infty} \sin(\frac{n\pi\theta_0}{\beta}) e^{-\frac{r_0'^2}{2t}} \int_0^{\beta} \sin(\frac{n\pi\theta}{\beta}) g_n(\theta) d\theta$$

with

$$g_n(\theta) = \int_0^{\infty} r e^{-\frac{r^2}{2t}} e^{b_1 r \cos\theta + b_2 r \sin\theta} I_{\frac{n\pi}{\beta}}(\frac{r r_0}{t})\, dr.$$

(Note: All parameters are as defined in Lemma 3.)

For the special correlation values discussed in Lemma 4, the distribution functions of Corollary 1 can be computed explicitly. We state our result in the following Corollary. Note that we consider here only the min-max case where the special correlation values $\rho_n = \cos(\pi/n)$ are positive.

Corollary 2 *The joint distribution of the minimum of one Brownian motion and the maximum of another, with constant correlation $\rho_n = \cos\frac{\pi}{n}$, $n = 2, 3, \ldots$, for $x_1 \geq 0$, $x_2 \leq 0$, is given by*

$$\mathcal{P}(\overline{X}_1(t) \leq x_1, \underline{X}_2(t) \geq x_2) = \sum_{k=0}^{n-1} \left[H(r_0, \frac{2k\pi}{n} + \theta) - H(r_0, \frac{2k\pi}{n} - \theta) \right]$$

where

$$\begin{aligned}
H(r_0, \frac{2k\pi}{n} \pm \theta) &= \exp\left[A_0 + A_1\zeta_1 + A_2\zeta_2 + \left(\frac{1}{2}A_1^2 + \frac{1}{2}A_2^2 + b \right) t \right] \times \\
&\quad N_2\left(\frac{\zeta_2 + A_2 t}{\sqrt{t}}, \frac{-\sqrt{1-\rho_n^2}(\zeta_1 + A_1 t) - \rho_n(\zeta_2 + A_2 t)}{\sqrt{t}}, -\rho_n \right) \\
\zeta_1 &= -r_0 \cos(\frac{2k\pi}{n} \pm \theta) \\
\zeta_2 &= r_0 \sin(\frac{2k\pi}{n} \pm \theta) \\
\tan\theta &= \frac{\sqrt{1-\rho_n^2}}{\rho_n - (x_1\sigma_2/x_2\sigma_1)}
\end{aligned}$$

and

$$\begin{aligned}
A_0 &= a_1 x_1 + a_2 x_2 \\
A_1 &= a_1 \sigma_1 \sqrt{1-\rho_n^2} \\
A_2 &= a_1 \sigma_1 \rho_n + a_2 \sigma_2
\end{aligned}$$

and $N_2(x, y, \rho)$ is the standard bivariate normal distribution for correlation ρ. The definitions of r_0, a_1, a_2, and b are the same as in Lemma 3.

PROOF OF COROLLARY 2: Note that

$$\begin{aligned}
\mathcal{P}(\overline{X}_1(t) \le x_1, \underline{X}_2(t) \ge x_2) &= \int_{-\infty}^{x_1}\int_{x_2}^{\infty} \mathcal{P}(X_1(t) \in dy_1, X_2(t) \in dy_2, \overline{X}_1(t) \le x_1, \underline{X}_2(t) \ge x_2) \\
&= \int_{-\infty}^{x_1} dy_1 \int_{x_2}^{\infty} dy_2\, p(-y_1, y_2, t; -x_1, x_2, -\alpha_1, \alpha_2, \sigma_1, \sigma_2, -\rho) \\
&= \int_{-x_1}^{\infty} dy_1 \int_{x_2}^{\infty} dy_2\, p(y_1, y_2, t; -x_1, x_2, -\alpha_1, \alpha_2, \sigma_1, \sigma_2, -\rho)
\end{aligned}$$

where we have used the results of Lemma 3(ii). The density on the last line is given by Lemma 4. Direct evaluation of the integrals leads to the result given.

■

We now state a theorem that summarizes the evaluation of various double lookback options discussed in Section 2. We define

$$\begin{aligned}
V_{Dmax}(x_1, x_2) &= \max\left[0, aS_1(0)e^{\max(M_1, x_1)} - bS_2(0)e^{\max(M_2, x_2)} - K\right] \\
V_{Dmin}(x_1, x_2) &= \max\left[0, aS_1(0)e^{\min(m_1, x_1)} - bS_2(0)e^{\min(m_2, x_2)} - K\right] \\
V_{DLS}(x_1, x_2) &= \max\left[0, aS_1(0)e^{\max(M_1, x_1)} - bS_2(0)e^{\min(m_1, x_2)} - K\right]
\end{aligned}$$

for some constants a and b.

Theorem 3 *The call prices C_{Dmax}, C_{Dmin} and C_{DLS}, respectively, for double maxima, double minima and double lookback spread options are determined as follows,*

$$C_{Dmax} = e^{-rT}\int_0^{\infty} dx_1 \int_0^{\infty} dx_2\, V_{Dmax}(x_1, x_2)\frac{\partial^2 \mathcal{P}(\overline{X}_1(t) \le x_1, \overline{X}_2(t) \le x_2)}{\partial x_1 \partial x_2} \tag{4}$$

$$C_{Dmin} = e^{-rT}\int_{-\infty}^{0} dx_1 \int_{-\infty}^{0} dx_2\, V_{Dmin}(x_1, x_2)\frac{\partial^2 \mathcal{P}(\underline{X}_1(t) \ge x_1, \underline{X}_2(t) \ge x_2)}{\partial x_1 \partial x_2} \tag{5}$$

$$C_{DLS} = e^{-rT}\int_{-\infty}^{0} dx_1 \int_0^{\infty} dx_2\, V_{DLS}(x_1, x_2)\frac{-\partial^2 \mathcal{P}(\underline{X}_1(t) \ge x_1, \overline{X}_2(t) \le x_2)}{\partial x_1 \partial x_2} \tag{6}$$

Formulas for corresponding puts can be obtained similarly.

This theorem states that the prices of double lookbacks can be obtained by integrating their final payoffs with respect to the corresponding distribution function. In general, evaluation will require a four-dimensional numerical quadrature. There exist standard numerical techniques evaluate such integrals and we expect that this methodology will be more computationally efficient than Monte-Carlo or lattice techniques.

σ_1	σ_2	ρ	K	Call	Put
		any	0	31.50	0.00
0.2	0.2	.50	30	7.14	4.17
		.71	30	6.55	3.58
		.90	30	5.92	2.95
		any	0	44.25	0.00
0.2	0.4	.50	45	7.59	6.15
		.71	45	6.69	5.26
		.90	45	5.73	4.28
		any	0	50.10	0.00
0.4	0.2	.50	50	13.28	10.74
		.71	50	12.63	10.14
		.90	50	12.09	9.57
		any	0	62.84	0.00
0.4	0.4	.50	65	12.58	11.57
		.71	65	11.43	10.46
		.90	65	10.32	9.35

Table 2: Values for call and put double lookback spread options with payouts of, respectively, $(\overline{S}_1 - \underline{S}_2 - K)^+$ and $(K - \overline{S}_1 + \underline{S}_2)^+$. We assume that $S_1(0) = S_2(0) = \overline{S}_1(0) = \underline{S}_2(0) = 100$, $r = 0.05\ \text{yr}^{-1}$, $q_1 = q_2 = 0$, $t = 1$ yr. Note that the correlation coefficients correspond to $\rho = \cos\frac{\pi}{3}$, $\cos\frac{\pi}{4}$, and $\cos\frac{\pi}{7}$.

In Figure 3, we present a surface and contour plot of the joint density function $-\frac{\partial^2}{\partial x_1 \partial x_2}\mathcal{P}(\overline{X}_1(t) \le x_1, \underline{X}_2(t) \ge x_2)$, as given in Corollary 2. The parameter values chosen for the picture are $\sigma_1 = \sigma_2 = 0.2\ \text{yr}^{-1/2}$, $t = 1$ yr, $r = 0.05\ \text{yr}^{-1}$, $q_1 = q_2 = 0$, and $\rho = 0.5$. Using the results of Theorem 3, we employ this density in a numerical quadrature evaluation of double lookback spread options.

In Table 2, we give numerical prices for various parameter values. Once again, these prices have the desired sensitivities with respect to σ_1, σ_2 and K. We note that as the correlation between the two assets increases, the option premiums decrease for both calls and puts. This should be intuitive, as higher correlation leads to lower volatility in $\overline{S}_1 - \underline{S}_2$. Also, if we hold $\sigma_1 = \sigma_2 = 0.2$, then the option premiums for the double lookback spread in Table 2 is more expensive than those of the lookback spread in Table 1 (for fixed strike). But, as ρ increases to 1, the option premiums converge to those in Table 1.

5 Semi-Lookbacks on Two Assets

To make an interesting comparison of various types of lookback options, we now introduce another class of lookback options whose payoffs depend on the extreme value of one asset and the final value of another asset. For example, consider an option to buy the maximum of S_1 at the final value of S_2 or an option to sell the minimum of S_2 at the final value of S_1. We call these options *semi-lookbacks.* In this section, we present the probability density functions necessary for the valuation of semi-lookbacks. We define

$$\mathcal{P}(\overline{X}_1(t) \in dx_1, X_2(t) \in dx_2) \equiv f_{+0}(x_1, x_2, t)\, dx_1\, dx_2$$

$$\mathcal{P}(X_1(t) \in dx_1, \underline{X}_2(t) \in dx_2) \equiv f_{0-}(x_1, x_2, t)\, dx_1\, dx_2$$

The following lemma is useful for the valuation of semi-lookback options.

Lemma 5 *The density function of the maximum of one Brownian motion and the end point of another Brownian motion is given by:*

$$f_{+0}(x_1, x_2, t) = \frac{1}{\sigma_1 \sigma_2 \pi t^2} \exp(A) \left[t\sqrt{1-\rho^2} \exp(-\frac{B^2}{2}) - d\sqrt{2\pi t} N(-B) \right] \tag{7}$$

with constants

$$\begin{aligned} A &= \frac{2\alpha_1 x_1}{\sigma_1^2} - \frac{\alpha_1^2 t}{2\sigma_1^2} - \frac{c^2\rho^2}{2t(1-\rho^2)} + \frac{d^2}{2t(1-\rho^2)} \\ B &= \frac{1}{\sqrt{t(1-\rho^2)}} \left(\frac{x_1}{\sigma_1} + \frac{\alpha_1 t}{\sigma_1}(1-\rho^2) + c\rho^2 \right) \end{aligned}$$

where

$$\begin{aligned} c &= \frac{x_2 - \alpha_2 t}{\sigma_2 \rho} - \frac{2x_1 - \alpha_1 t}{\sigma_1} \\ d &= \frac{\alpha_1 t(1-\rho^2)}{\sigma_1} + c\rho^2 \end{aligned}$$

Similarly,

$$f_{0-}(x_1, x_2) = f_{+0}(-x_2, x_1, t; -\alpha_2, \alpha_1, \sigma_2, \sigma_1, -\rho).$$

PROOF: Set

$$\begin{aligned} B_1(t) &= w_1^*(t) \\ B_2(t) &= -\frac{\rho}{\sqrt{1-\rho^2}} w_1^*(t) + \frac{1}{\sqrt{1-\rho^2}} w_2^*(t) \end{aligned}$$

Then, (B_1, B_2) is a standard (uncorrelated) two-dimensional Brownian motion. Hence we get for the density function f_{+0},

$$\begin{aligned} & f_{+0}(x_1, x_2, t)\, dx_1\, dx_2 \\ & \quad = \mathcal{P}(\overline{X}_1(t) \in dx_1; X_2(t) \in dx_2) \end{aligned}$$

$$= \mathcal{P}(\max_{0\le s\le t}(\alpha_1 s + \sigma_1 w_1^*(s)) \in dx_1; \alpha_2 t + \sigma_2 w_2^*(t) \in dx_2)$$
$$= \mathcal{P}(\max_{0\le s\le t}(\alpha_1 s + \sigma_1 B_1(s)) \in dx_1; \alpha_2 t + \sigma_2(\rho B_1(t) + \sqrt{1-\rho^2}B_2(t)) \in dx_2)$$
$$= \mathcal{P}(\max_{0\le s\le t}(\frac{\alpha_1}{\sigma_1}s + B_1(s)) \in \frac{dx_1}{\sigma_1}; \frac{\alpha_2}{\sigma_2\rho}t + B_1(t) \in \frac{dx_2}{\sigma_2\rho} - \frac{\sqrt{1-\rho^2}}{\rho}B_2(t))$$
$$= \int_{-\infty}^{\infty} \mathcal{P}(\max_{0\le s\le t}(\frac{\alpha_1}{\sigma_1}s + B_1(s)) \in \frac{dx_1}{\sigma_1}; \frac{\alpha_2}{\sigma_2\rho}t + B_1(t) \in \frac{dx_2}{\sigma_2\rho} - \frac{\sqrt{1-\rho^2}}{\rho}x|B_2(t) = x) \times \mathcal{P}(B_2(t) \in dx)$$
$$= \int_{-\infty}^{\infty} \mathcal{P}(\max_{0\le s\le t}(\frac{\alpha_1}{\sigma_1}s + B_1(s)) \in \frac{dx_1}{\sigma_1}; \frac{\alpha_1}{\sigma_1}t + B_1(t) \in \frac{dx_2}{\sigma_2\rho} - \frac{\sqrt{1-\rho^2}}{\rho}x + (\frac{\alpha_1}{\sigma_1} - \frac{\alpha_2}{\sigma_2\rho})t) \times \mathcal{P}(B_2(t) \in dx)$$
$$= \left[\int_a^\infty \frac{2(2\bar{y}-\bar{x})}{\sqrt{2\pi t^3}} \exp(-\frac{(2\bar{y}-\bar{x})^2}{2t}) \exp(\frac{\alpha_1}{\sigma_1}\bar{x} - \frac{1}{2}(\frac{\alpha_1}{\sigma_1})^2 t) \frac{1}{\sqrt{2\pi t}} \exp(-\frac{x^2}{2t}) dx\right] \frac{dx_1\, dx_2}{\rho\sigma_1\sigma_2}$$

where we have set $\bar{y} = \frac{x_1}{\sigma_1}$ and $\bar{x} = \frac{x_2}{\sigma_2\rho} - \frac{\sqrt{1-\rho^2}}{\rho}x + (\frac{\alpha_1}{\sigma_1} - \frac{\alpha_2}{\sigma_2\rho})t)$. In the last equation, we used the density corresponding to the distribution given in Lemma 1. The lower integral bound a is defined by the condition $\bar{y} \ge \bar{x}$ which is equivalent to

$$x \ge \frac{\rho}{\sqrt{1-\rho^2}}(\frac{x_2}{\sigma_2\rho} - \frac{x_1}{\sigma_1} + (\frac{\alpha_1}{\sigma_1} - \frac{\alpha_2}{\sigma_2\rho})t) \equiv a$$

Evaluating the integral yields the result given. The result for $f_{0-}(x_1, x_2, t)$ follows by symmetry. ■

In Figure 4, we plot the semi-lookback density f_{+0} for a specific choice of parameter values. As an example of the use of this density, we consider the evaluation of the following class of semi-lookback options. We define the call options as

$$\begin{aligned} V_{Smax}(x_1, x_2) &= \max\left[0, S_1(0)e^{\max(M_1, x_1)} - S_2(0)e^{x_2} - K\right] \\ V_{Smin}(x_1, x_2) &= \max\left[0, S_1(0)e^{x_1} - S_2(0)e^{\min(m_2, x_2)} - K\right] \end{aligned}$$

with corresponding formulas for the put options. The values of these options are given by the following theorem which makes use of the density derived in Lemma 5.

Theorem 4 *Let C_{Smax} and C_{Smin} be the prices at time 0 of a semi-lookback call options with payoffs $\left[\overline{S}_1(T) - S_2(T) - K\right]^+$ and $[S_1(T) - \underline{S}_2(T) - K]^+$, respectively. Then,*

$$C_{Smax} = e^{-rT}\int_0^\infty dx_1 \int_{-\infty}^\infty dx_2\, V_{Smax}(x_1, x_2)\, f_{+0}(x_1, x_2, T) \tag{8}$$
$$C_{Smin} = e^{-rT}\int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^0 dx_2\, V_{Smin}(x_1, x_2) f_{0-}(x_1, x_2, T). \tag{9}$$

In Table 3, we list values of semi-lookback spread options for various parameter values, obtained by numerical quadrature of the integrals in Theorem 4. Comparing Tables 2 and 3, we find that for a fixed set of volatilities, correlation, and strike parameters, the call premiums for semi-lookback spread options are cheaper than those of the double lookback options. However, the put premiums should be more expensive. As the correlation between two assets increases, the option premiums decrease for both calls and puts.

6 Double Knockouts

The analysis presented in the previous sections can be readily used to value knockout options based on two assets. In a standard one-asset knockout call, the option pays $\max(S_1(T) - K, 0)$, subject to a condition that the stock price $S_1(t)$ never hits a fixed boundary (which could either be larger or smaller than the initial stock price). Merton (1973) was the first to provide a solution to value this option. We refer the reader to Rubinstein (1992) for a complete list of one-asset knockout options.

In the case with two traded assets, we consider a general contingent claim which pays, at the maturity date,

$$V(T) = f(S_1(T), S_2(T))$$

for some function f, subject to the conditions that $S_1(t)$ and/or $S_2(t)$ never hit some pre-determined boundaries. We shall call such options as *double knockouts*. Obviously, the keys to valuing double knockouts are the following probability density functions, which are derived in Lemma 3:

$$\mathcal{P}(X_1(T) \in dx_1, X_2(T) \in dx_2, \overline{X}_1 \leq K_1, \overline{X}_2 \leq K_2)$$

$$\mathcal{P}(X_1(T) \in dx_1, X_2(T) \in dx_2, \underline{X}_1 \geq K_1, \overline{X}_2 \leq K_2)$$

$$\mathcal{P}(X_1(T) \in dx_1, X_2(T) \in dx_2, \underline{X}_1 \geq K_1, \underline{X}_2 \geq K_2)$$

Integrating these functions over the payoff $V(T)$ gives rise to the desired option premium.

In Figures 5 and 6, we present plots of the knockout density functions corresponding to the distributions derived in Lemmas 2 and 3. In both cases, only a few terms in the infinite series are typically needed to obtain good convergence.

Similar to the semi-lookbacks considered in the previous section, we can also evaluate a special class of knockout options which are European calls or puts written on one asset, subject to a condition that the value of another asset never hits a pre-determined boundary. The probability densities necessary for this type of options are

$$\mathcal{P}(X_1(T) \in dx_1, \overline{X}_2 \leq K_2)$$

σ_1	σ_2	ρ	K	Call	Put
		0.1	0	18.44	4.16
		0.1	15	9.52	9.51
		0.3	0	17.78	3.49
		0.3	15	8.63	8.62
0.2	0.2	0.5	0	17.04	2.74
		0.5	15	7.64	7.62
		0.7	0	16.18	1.88
		0.7	15	6.47	6.45
		0.9	0	15.09	0.78
		0.9	15	4.97	4.95
		0.1	0	25.46	11.16
		0.1	20	13.87	18.62
		0.3	0	24.57	10.27
		0.3	20	12.86	17.60
0.2	0.4	0.5	0	23.58	9.28
		0.5	20	11.75	16.48
		0.7	0	22.44	8.13
		0.7	20	10.53	15.25
		0.9	0	21.05	6.73
		0.9	20	9.18	13.90
		0.1	0	35.56	2.67
		0.1	30	16.68	12.35
		0.3	0	34.93	2.04
		0.3	30	15.47	11.14
0.4	0.2	0.5	0	34.28	1.39
		0.5	30	14.14	9.80
		0.7	0	33.64	0.75
		0.7	30	12.62	8.28
		0.9	0	33.07	0.18
		0.9	30	10.83	6.49
		0.1	0	41.01	8.13
		0.1	35	19.36	19.79
		0.3	0	39.75	6.86
		0.3	35	17.56	17.99
0.4	0.4	0.5	0	38.31	5.42
		0.5	35	15.52	15.95
		0.7	0	36.62	3.72
		0.7	35	13.08	13.49
		0.9	0	34.45	1.55
		0.9	35	9.92	10.33

Table 3: Values for call and put semi-lookback spread options with terminal payouts of, respectively, $(\overline{S}_1 - S_2 - K)^+$ and $(K - \overline{S}_1 + S_2)^+$. We assume that $S_1(0) = S_2(0) = \overline{S}_1(0) = 100$, $r = 0.05\ \text{yr}^{-1}$, $q_1 = q_2 = 0$, $t = 1$ yr.

$$\mathcal{P}(X_1(T) \in dx_1, \underline{X}_2 \geq K_2)$$

which can be obtained by integrating the densities in Lemma 5.[2]

7 Conclusions

We have presented a technique for pricing lookback options on two assets following lognormal distribution. The essential part of this technique is to derive the probability distribution function of the extreme values of two correlated Brownian motions. With this technique, the prices of many kinds of lookback and knockout options can be calculated efficiently. We hope that our pricing technology will be useful for future research that involves the extreme values of two correlated geometric Brownian motions.

[2]Heynen and Kat (1994) have derived the same density function and obtained closed-form solutions for semi-knockout options.

References

[1] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables*, 1972, 9th Printing, US Department of Commerce.

[2] R. H. Cameron and W. T. Martin, *Transformation of Wiener integrals under translations*, Ann.Math. 45 (1944), 386-396.

[3] H. S. Caslow, *Conduction of Heat in Solids*, Clarendon Press, Oxford, 1959.

[4] A. Conze and Viswanathan, *Path-Dependent Options: The Case of Lookback Options*, Journal of Finance 46, 1893–1907 (1991).

[5] C. W. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*, 2nd Ed., Springer-Verlag, Berlin, 1990, p.132.

[6] M. B. Goldman, H. B. Sosin, and M. A. Gatto, *Path Dependent Options: "Buy at the Low, Sell at the High"*, Journal of Finance 34, 1111–1127 (1979).

[7] M. B. Goldman, H. B. Sosin, and L. A. Shepp, *On Contingent Claims that Insure Ex-Post Optimal Stock Market Timing*, Journal of Finance 34, 401–413 (1979).

[8] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*, Academic Press, New York, 1980.

[9] J. M. Harrison and D. Kreps, *Martingale and Arbitrage in Multiperiod Security Markets*, Journal of Economic Theory 20, 381-408 (1979).

[10] J. M., Harrison and S. Pliska, *Martingales and Stochastic Integrals in the Theory of Continuous Trading*, Stochastic Processes and Their Applications 11, 215-260 (1981).

[11] R. Heynen and H. Kat, *Crossing Barriers*, Risk, Vol. 7, No. 6, 46-51 (1994).

[12] I. Karatzas and S. Shreve, *Brownian Motion and Stochastic Calculus*, 2nd Ed., Springer-Verlag, Berlin, 1991.

[13] P. Lévy, *Processus stochastiques et mouvement brownien*, Gauthier-Villars, Paris, 1948.

[14] R. Merton, *Theory of Rational Option Pricing*, Bell Journal of Economics and Management Science 4, 141–183 (1973).

[15] M. Rubinstein, Notes on Exotic Options, 1992.

[16] P. Wilmott, J. Dewynne, and S. Howison, *Option Pricing*, Oxford Financial Press, Oxford, 1993.

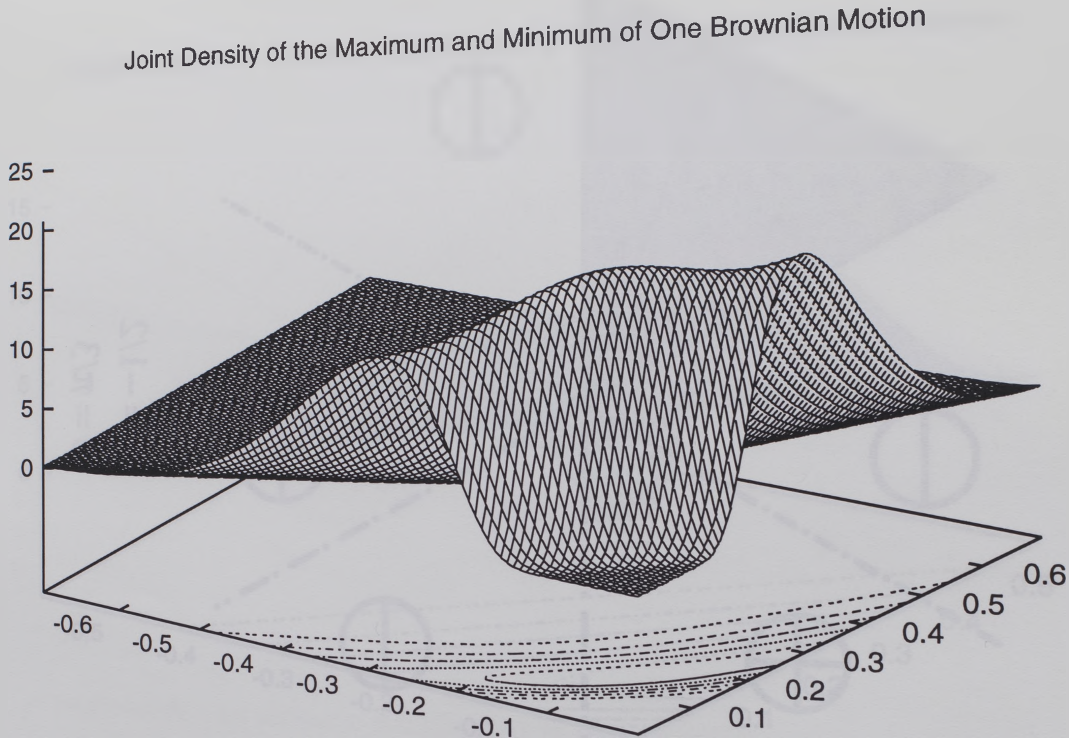
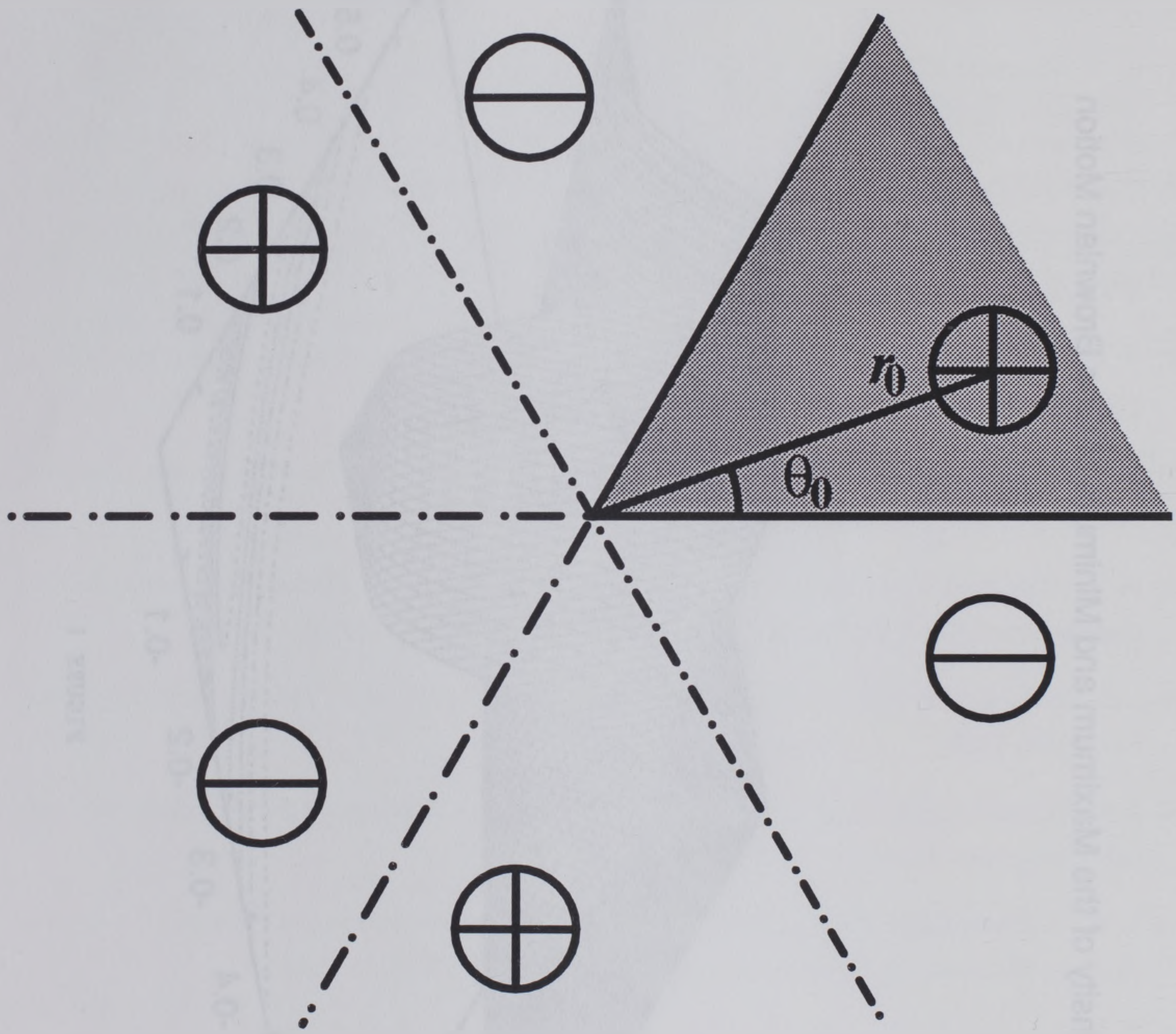

Joint Density of the Maximum and Minimum of One Brownian Motion
25
20
15
10
5
0
-0.6
-0.5
-0.4
-0.3
-0.2
-0.1
0.6
0.5
0.4
0.3
0.2
0.1

FIGURE 1



$\rho = -1/2$
$\beta = \pi/3$

FIGURE 2

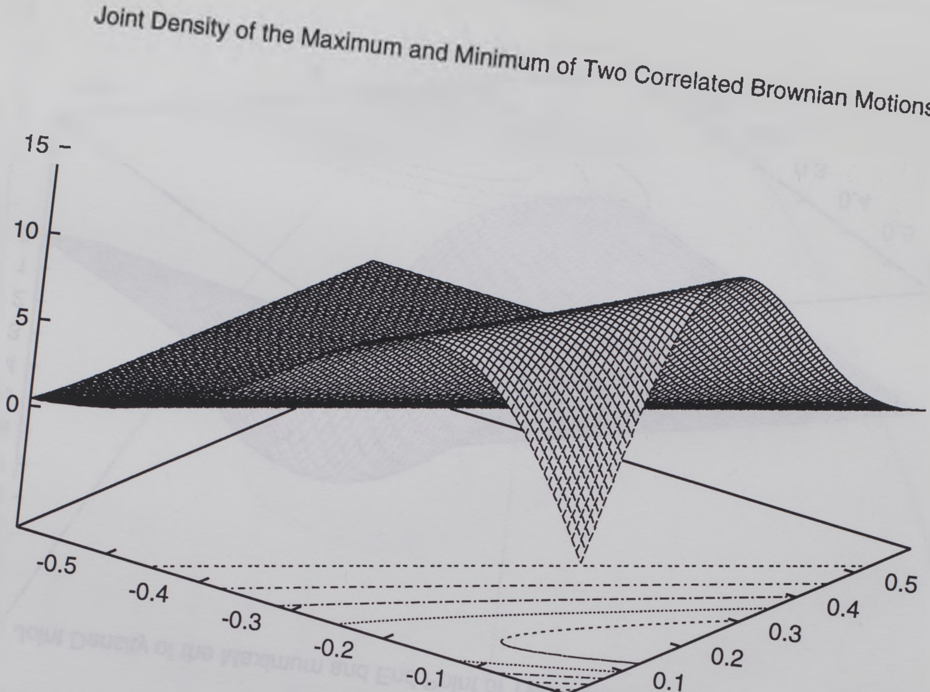

Joint Density of the Maximum and Minimum of Two Correlated Brownian Motions
15
10
5
0
-0.5
-0.4
-0.3
-0.2
-0.1
0.1
0.2
0.3
0.4
0.5

FIGURE 3

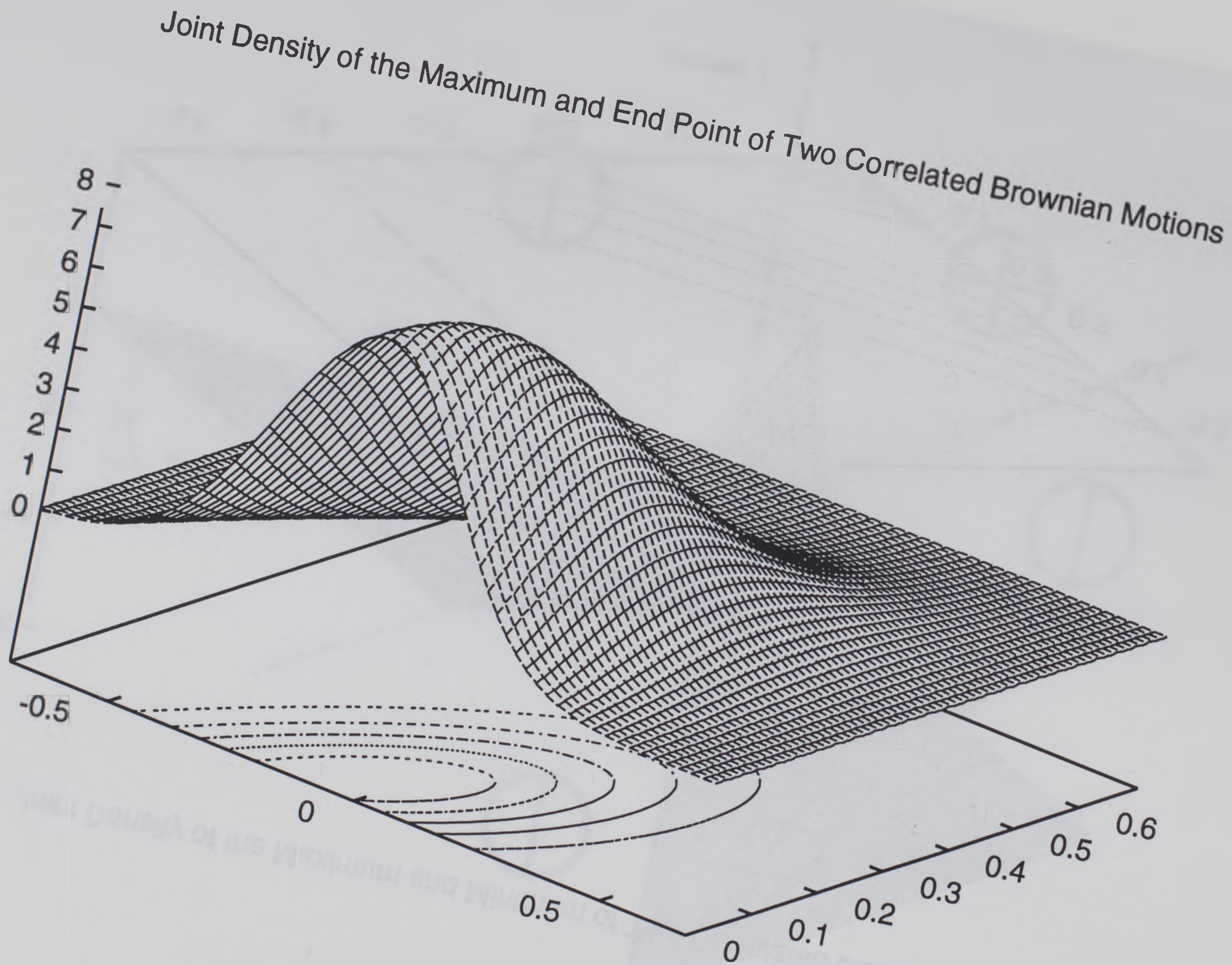

Joint Density of the Maximum and End Point of Two Correlated Brownian Motions
8
7
6
5
4
3
2
1
0
-0.5
0
0.5
0
0.1
0.2
0.3
0.4
0.5
0.6

FIGURE 4

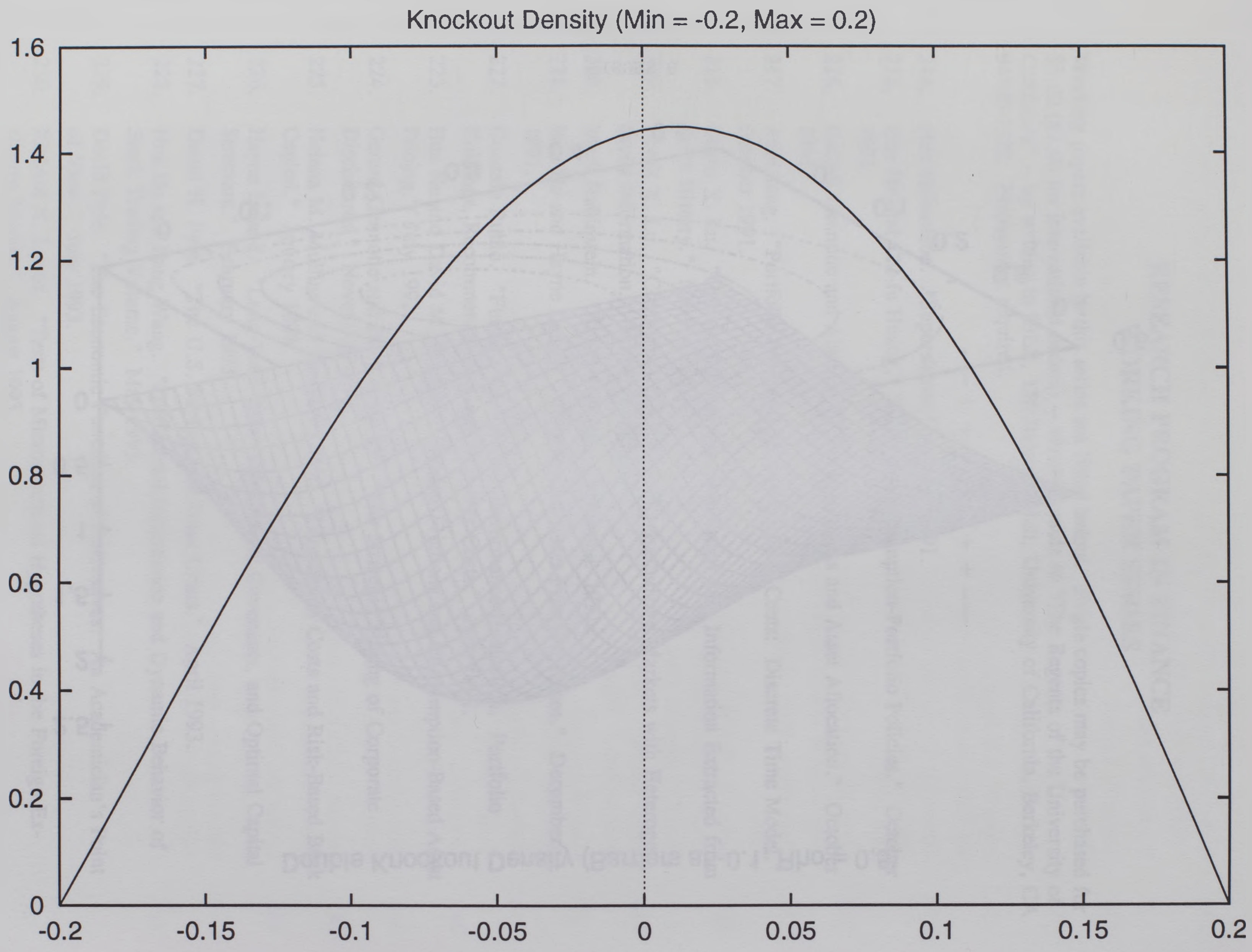

Knockout Density (Min = -0.2, Max = 0.2)
1.6
1.4
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
-0.15
-0.1
-0.05
0
0.05
0.1
0.15
0.2

FIGURE 5

Double Knockout Density (Barriers at -0.1, Rho = 0.5)

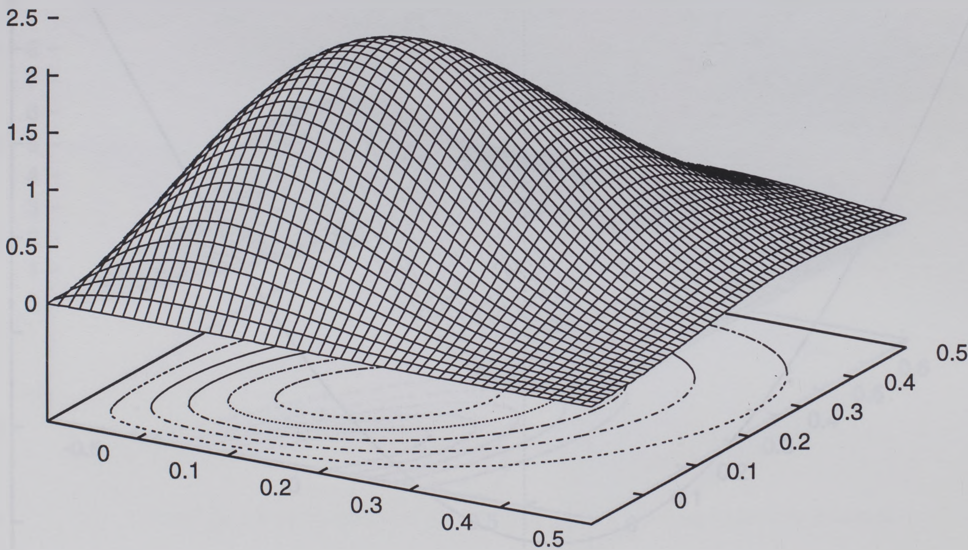


FIGURE 6

RESEARCH PROGRAM IN FINANCE
WORKING PAPER SERIES

Working papers available in this series are listed below. Single copies may be purchased for $5.00 ($7.00 for international orders) -- check payable to "The Regents of the University of California" -- by writing to IBER, 156 Barrows Hall, University of California, Berkeley, CA 94720-1922. Prepayment required.

----- + + + --- + + + -----

214. Nils Hakansson. "Supershares." August 1991.

215. Hua He and Chi-fu Huang. "Efficient Consumption-Portfolio Policies." October 1991.

216. Gerard Gennotte and Alan Jung. "Commissions and Asset Allocation." October 1991.

217. Alan Jung. "Portfolio Policies with Transactions Costs: Discrete Time Model." October 1991.

218. Lewis X. Lu. "Optimal Continuous Speculation with Information Extracted from Price History." November 1991.

219. Lewis X. Lu. "Continuous Equilibrium in Speculative Markets with Heterogeneous Information." November 1991.

220. Mark Rubinstein. "Exotic Options." December 1991.

221. Hua He and Hayne Leland. "Equilibrium Asset Price Processes." December 1991.

222. Gonzalo Rubio. "Further Evidence on Performance Evaluation: Portfolio Holdings, Recommendations, and Turnover Costs." June 1992.

223. Hua He and David M. Modest. "Market Frictions and Consumption-Based Asset Pricing." July 1992.

224. Gerard Gennotte and Brett Trueman. "The Strategic Timing of Corporate Disclosures." November 1992.

225. Helena M. Mullins and David H. Pyle. "Liquidation Costs and Risk-Based Bank Capital." January 1993.

226. Hayne Leland. "Long-Term Debt Value, Bond Covenants, and Optimal Capital Structure." February 1993.

227. David H. Pyle. "The U.S. Savings and Loan Crisis." April 1993.

228. Hua He and Jiang Wang. "Differential Information and Dynamic Behavior of Stock Trading Volume." May 1993.

229. David Pyle. "The Economic Functions of Derivatives: An Academician's Point of View." July 1993.

230. Richard K. Lyons. "Tests of Microstructural Hypotheses in the Foreign Exchange Market." August 1993.

231. Richard K. Lyons. "Optimal Transparency in a Dealership Market with an Application to Foreign Exchange." September 1993.

232. Mark Rubinstein. "Implied Binomial Trees." January 1994.

233. Hayne E. Leland. "Corporate Debt Value, Bond Covenants, and Optimal Capital Structure." January 1994.

234. Bruce N. Lehmann and David M. Modest. "Trading and Liquidity on the Tokyo Stock Exchange: A Bird's Eye View." April 1994.

235. Bruce N. Lehmann and David M. Modest. "Market Structure and Liquidity on the Tokyo Stock Exchange." May 1994.

236. Domenico Cuoco and Hua Hè. "Dynamic Aggregation and Computation of Equilibria in Finite-Dimensional Economies with Incomplete Financial Markets." June 1994.

237. Klaus Bjerre Toft. "Exact Formulas for Expected Hedging Error and Transactions Costs in Option Replication." July 1994.

238. Klaus Bjerre Toft. "Options on Leveraged Equity with Default Risk." July 1994.

239. Robert R. Grauer and Nils H. Hakansson. "Gains from Diversifying into Real Estate: Three Decades of Portfolio Returns Based on the Dynamic Investment Model." October 1994.

240. Hayne Leland. "Bond Prices, Yield Spreads, and Optimal Capital Structure with Default Risk." November 1994.

241. Mark Rubinstein. "On the Accounting Valuation of Employee Stock Options." December 1994.

242. Richard K. Lyons and Andrew K. Rose. "Explaining Forward Exchange Bias...Intraday." January 1995.

243. Richard K. Lyons. "Foreign Exchange Volume: Sound and Fury Signifying Nothing?" January 1995.

244. Hayne Leland and Gregory Connor. "Optimal Cash Management for Investment Funds." March 1995.

245. Marcus Berliant and Sankar De. "On Revelation of Private Information in Stock Market Economies." April 1995.

246. Sankar De, Mark Fedenia, and Alexander J. Triantis. "Effects of Competition on Bidder Returns." April 1995.

247. Richard Stanton and Nancy Wallace. "Anatomy of an ARM: Index Dynamics and Adjustable Rate Mortgage Valuation." April 1995.

248. Hua He, William P. Keirstead, and Joachim Rebholz. "Double Lookbacks." May 1995.