PLACING NO-ARBITRAGE BOUNDS ON THE VALUE OF NONMARKETABLE AND THINLY-TRADED SECURITIES

Francis A. Longstaff*

I just remember losing money and staying up nights trying to peddle things, and being unable to sell anything at the market prices," recalls money-manager Michael Steinhardt. . . . To staunch losses, Mr. Steinhardt told his strategist to dump large blocks of European bonds; these could only be sold at well below bid prices. . . . Within days, Mr. Steinhardt could hardly find buyers for even the smallest lots. "By the time we sold our last bonds," he says, "the liquidity had essentially disappeared."

—*Wall Street Journal*, May 20, 1994

*University of California, Los Angeles.

Advances in Futures and Options Research,
Volume 8, pages 203–228.

ISBN: 1-55938-852-8

INTRODUCTION

A marketable security is one which can be sold by a security holder at any point in time. Thus, a security is nonmarketable from the perspective of a security holder if it can only be sold after a delay, or if the security cannot be sold at all for some fixed period of time. The question of how marketability affects security prices has become one of the most important and central issues in valuation theory as empirical research reveals a growing number of liquidity-related anomalies which cannot be reconciled with existing asset-pricing models. In addition, the recent collapse or near-collapse of a number of highly-levered financial institutions because of event-related illiquidity of securities has also made this issue of key importance to regulators, rating agencies, exchanges, auditors, and other market participants.

The empirical evidence suggests that less-marketable securities are often valued at large discounts to marketable securities. For example, Daves and Ehrhardt (1993) and Grinblatt and Longstaff (1995) show that thinly-traded stripped coupon payments from Treasury bonds can sell for more than three percent below the price of stripped principal payments, even though both have identical cash flows, tax treatment, and bid-ask spreads. Boudoukh and Whitelaw (1991) find that Japanese government bonds can sell at prices more than eight percent below the price of the intensively-traded benchmark issue with a similar coupon rate and maturity date. Pratt (1989) and Silber (1992) find that Rule 144 or letter stock is typically placed privately at thirty to forty percent discounts to the value of otherwise identical unrestricted stock.

Critics of efficient markets often argue that discounts for lack of marketability are too large to be consistent with rational price setting, and view them as clear evidence that investor sentiment determines security values. In contrast, previous research by Mayers (1972, 1973), Brito (1977), Boudoukh and Whitelaw (1993), and others shows that discounts for nonmarketability can occur in equilibrium models. The size of the discount, however, depends on how closely the optimal portfolio strategy for an investor approximates the buy-and-hold strategy. In fact, Mayers (1976) and Stapleton and Subrahmanyam (1979) show that equilibrium discounts for nonmarketability do not occur for a large class of preference functions. This leaves open the issue of whether observed discounts are too large to be consistent with market rationality.

This paper derives analytical expressions for the upper bound on the value of the discount for lack of marketability in a no-arbitrage framework. If the discount exceeds this bound, then a contingent claim could be written which would guarantee an arbitrage profit from holding the nonmarketable security. Intuitively, this upper bound reflects the present value of the greatest possible loss that an investor could experience by foregoing the right to sell the security at any point in time. This upper bound can also be viewed as the largest spread or market-impact cost an investor would be willing to pay to obtain immediacy.

We first examine the case where the security cannot be sold for some fixed period of time. Examples of this type of marketability restriction include letter stock which cannot be sold for two years after its issuance, securities which have been loaned via reverse repurchase agreements or pledged as collateral since these securities cannot be sold by the original owner until returned, and shares acquired in an IPO where the investor has an implicit obligation not to resell the shares immediately in the aftermarket. Using a continuous-time framework, we derive a no-arbitrage upper bound for this type of marketability restriction. The upper bound is a function of the volatility of the security and can vary significantly across assets. To address the rationality issue, we compare these bounds to the prices of restricted stock. The results suggest that the average discounts reported in the literature for restricted stock approximate or exceed the upper bound.

We then examine the thinly-traded case where the security holder may only be able to sell the security after a delay. Examples of this type of marketability restriction include stock issued by small-capitalization firms and off-the-run Treasury notes and bonds. We derive an upper bound for the thin-trading discount and illustrate its dependence on both the volatility and average selling time for the security. We contrast the thin-trading discounts reported in the literature for less-liquid fixed income securities to the upper bound on the thin-trading discount. We find that the reported discounts appear to be much larger than the upper bound.

It is important to acknowledge that this analysis does not constitute a formal test of the hypothesis that illiquid securities are rationally priced. Nevertheless, these results are difficult to reconcile within the context of traditional asset pricing paradigms. In any event, these results indicate that these no-arbitrage bounds can provide useful tools for studying prices in financial markets. In addition, these bounds also provide a number of important insights into the potential costs of trading restrictions such as price limits, trading halts, and circuit breakers.

The remainder of this paper is organized as follows: the second section reviews the empirical evidence about discounts for lack of marketability, the third section considers the case where securities are completely nonmarketable for a fixed period of time, the fourth section considers the thinly-traded case where the investor generally experiences a delay in selling the security, and the fifth section discusses the implications of the results and presents concluding remarks.

THE EMPIRICAL EVIDENCE

To motivate the derivation of the upper bounds, we first briefly review the evidence in the empirical literature about discounts for different types of marketability restrictions. These estimates can then be compared directly to the analytical bounds derived in later sections.

Discounts for Nonmarketability

Much of the empirical evidence about discounts for nonmarketability is based on the pricing of SEC Rule 144 restricted or letter stock. This is stock issued by a firm which is not registered for public trading but is otherwise identical to publicly traded stock. The primary limitation of Rule 144 stock is that the recipient cannot sell the shares for a two-year period. After two years, the shares become marketable, subject to several minor trading-volume limitations. Restricted shares are typically issued by firms via private placements instead of the usual public offering mechanism. By comparing the price at which the restricted stock is privately placed to the market price for the firm's registered shares, the discount for nonmarketability can be directly measured.

Pratt (1989) summarizes the evidence from eight different studies of restricted stock spanning the 1966–1984 period. The studies include the Institutional Investor Study Report of the Securities and Exchange Commission 1971; Gelman 1972; Moroney 1973; Maher 1976; Trout 1977; and SRC Quarterly Reports 1983. The mean or median percentage discount found in these studies ranges from 25.8–45.0%. This is consistent with the results of a recent study by Silber (1992) who finds that the mean discount for nonmarketability is 33.75% in a sample of 69 private placements of stock during the 1981–1988 period.[1] In contrast, Wruck (1989) studies a sample of 37 privately placed equity issues by NYSE and AMEX firms and reports an average discount of only 13.5%. Taking the average of the discounts reported in these studies suggests an overall average of roughly 35%.

The evidence also suggests that discounts for nonmarketability can be extremely large for some firms. For example, Silber (1992) shows that the discount can exceed 80% for some firms. Moroney (1973) reports the ranges of discounts for a sample of restricted stock issues privately placed with a number of registered investment companies. Of the 11 funds in the sample, 10 purchased restricted stock at discounts in excess of 50%, and four purchased restricted stock at discounts of more than 75%. The practice of acquiring blocks of restricted stock is very common among closed-end mutual funds which have little need of immediate liquidity. These holdings of restricted stock may provide a partial explanation for closed-end mutual fund discounts. For example, Barclay, Holderness, and Pontiff (1994) find that closed-end funds that have restricted securities typically have discounts that are about 2% larger than other funds, although the difference is not statistically significant.

Thin-trading Discounts

Perhaps the most direct evidence about the value of the thin-trading discount comes from the fixed income markets. This is because it is often possible to observe prices for fixed income instruments which have identical cash flows and differ only in terms of their liquidity.

In an important recent paper, Amihud and Mendelson (1991) compare the prices of Treasury notes and bonds with no remaining coupon payments to those of Treasury bills with similar maturities. In the absence of liquidity effects, the prices for the two types of securities should be identical. Amihud and Mendelson, however, find that the less-actively-traded Treasury notes and bonds yield an average of 43 basis points more than the corresponding Treasury bills. Given that the average maturity for the Treasury notes and bonds in their sample is 97 days, this translates into an average price discount for the less-marketable Treasury notes and bonds of about 0.12%. Kamara (1994) finds similar results for the percentage price differences between Treasury notes and bonds and Treasury bills.

Boudoukh and Whitelaw (1991) document the benchmark effect in the Japanese government bond market. In this market, a specific bond issue is designated as the benchmark issue by the market participants. Typically, the benchmark issue has a maturity of 9–10 years. Even though there are more than a hundred different issues of Japanese government bonds outstanding, trading in the benchmark issue constitutes 90–95% of the total trading volume for all Japanese government issues. Boudoukh and Whitelaw find that the difference in yields between the highly-liquid benchmark issue and the more-thinly-traded adjacent-maturity issues averaged about 50 basis points during the 1984–1987 period, and was as high as 120 basis points. This implies that the pricing discount for non-benchmark bonds typically was in the 4–8% range during the sample period.

Daves and Ehrhardt (1993) and Grinblatt and Longstaff (1995) study the pricing of Treasury Strips. These zero-coupon bonds are the stripped coupon and principal payments from specific Treasury issues which are identified as strippable. Currently, there are more than 40 Treasury issues which are eligible to be stripped in book entry form. Since each eligible Treasury bond pays its last coupon simultaneously with the terminal principal payment, the cash flow from the last coupon Strip is identical to that of the principal Strip. The only substantive difference between the two securities is that there is a larger dollar amount of the principal Strip outstanding than of the last coupon Strip. Consequently, the greater supply of principal Strips makes them more liquid or marketable than the coupon Strips. The bid-ask spreads for the two types of Strips are typically the same. Daves and Ehrhardt (1993) show that the average percentage pricing difference between the principal and coupon Strips is 0.86%, but find that some percentage differences are as large as 2.3%. The average maturity of the Strips examined in their study is roughly 15 years. Using a more extensive sample, Grinblatt and Longstaff (1995) find that percentage differences for 20 to 30-year Strips can be as high as 4%.

Finally, the fact that the on-the-run or most-recently-issued Treasury notes and bonds trade at higher prices than those of older and less-liquid bonds with similar maturities has been extensively documented. Typical estimates of the difference between the yields of on-the-run and off-the-run 30-year Treasury bonds are about 3–5 basis points. This implies a percentage thin-trading discount for off-the-run 30-year bonds of roughly 0.30–0.50%. The off-the-run discount can be considerably larger, however. For example, Cornell and Shapiro (1990) find that the yield difference for the 30-year 9.25% Treasury bond issued in February 1986 was about 20 basis points, which implies a percentage thin-trading discount of about 2%.

To summarize, the markets for Treasury obligations imply estimates of the thin-trading discount ranging from 0.12% for three-month Treasury notes and bonds, to 0.3 to 0.5% for 30-year off-the-run bonds, and to 0.86% for Strips with an average maturity of 15 years. The average thin-trading discount in the Japanese bond market for the 10-year benchmark issue is on the order of 4%. Clearly, however, these ranges are sample specific and are only approximate. Hence, comparisons made between these estimates and the upper bounds should be viewed as illustrative rather than as formal tests of the hypothesis of market rationality.

NONMARKETABLE SECURITIES

In this section, we derive bounds on the discount for lack of marketability in the case where securities are completely nonmarketable by the security holder for a fixed period of time. In deriving pricing bounds, our approach is to identify a contingent claim that would compensate an investor for the largest loss he could incur by foregoing the right to sell the security for a fixed period of time.

Let V denote the current or time-zero price of an asset or security that is continuously traded in a frictionless market in which there are a large number of investors, all of whom can buy, sell, or short sell the asset as well as trade contingent claims on the asset. We assume that the equilibrium dynamics of V are given by the stochastic process

$$dV = \mu V dt + \sigma V dZ, \tag{1}$$

where μ and σ are constants and Z is a standard Wiener process. We also assume that all investors can borrow or lend at the constant interest rate r. With these assumptions, standard methods for pricing contingent claims on V are applicable.

Now assume that a specific investor is offered an amount ϕ to forego the right to sell the security prior to some fixed horizon T.[2] Let R denote the investor's reservation price for the security with this marketability restriction. If $\phi > V - R$, then the investor would accept the offer. In general, however, the amount ϕ that would need to be offered would vary from investor to investor. Thus, to obtain an upper bound, we must identify a value ϕ such that the offer would be accepted by any investor who prefers more wealth to less. To do this, we identify a contingent claim that when combined with the restricted security results in a pattern of cash flows that dominates the cash flows the investor would receive by retaining the right to sell the security. The market value of this derivative security then provides the appropriate value of ϕ.[3]

Observe that nonmarketability is investor specific rather than security specific in this framework. This differs from the equilibrium models presented in Amihud and Mendelson (1986) and Boudoukh and Whitelaw (1993) in which the security is assume to be illiquid from the perspective of all investors in the market. We use this approach since our objective is to derive an upper bound rather than to model the equilibrium value of liquidity. It is important to observe, however, that many forms of illiquidity are actually investor specific. For example, Rule 144 restrictions apply only to certain investors, rather than to all stockholders of a firm. This notion of marketability is more in the spirit of the definition of liquidity given by Lippman and McCall (1986).

Table 1 compares the cash flows from two investment strategies. The first strategy represents the cash flows to an unrestricted investor from purchasing the security at time zero for an outlay of V and then selling at some arbitrary time $t \leq T$ for a cash inflow of V_t. In the second strategy, the restricted investor pays R for the security at time zero, borrows V_t at

Table 1. Cash Flows from the Unrestricted and Restricted Strategies When the Unrestricted Security is Sold Before the End of the Marketability Restriction Period

Strategy	*Time-0 Cash Flow*	*Time-t Cash Flow*	*Time-T Cash Flow*
Strategy 1			
Buy Security	$-V$		
Sell Security		$+V_t$	
Total	$-V$	$+V_t$	0
Strategy 2			
Buy Security	$-R$		
Sell Security			$+V_T$
Borrow		$+V_t$	
Repay Loan			$-V_t e^{r(T-t)}$
Buy Claim	$-D$		
Claim Payoff			$\max_\tau(V_\tau e^{r(T-\tau)} - V_T)$
Total	$-R-D$	$+V_t$	≥ 0

Notes: The first strategy represents the cash flows to an unrestricted investor from taking a position in a marketable security at time zero and selling at some arbitrary time t. The second strategy represents the cash flows to an investor from taking a position in the same security but with the restriction that the investor cannot sell the security prior to time T, where $t \leq T$. V denotes the market value of the security. R is the value of the security to the investor given the marketability restriction. D denotes the market value of a contingent claim with the indicated payoff function at time T, where the maximum is taken over all τ, $0 \leq \tau \leq T$.

time t, and then sells the security at time T when the restriction lapses and repays the loan. Since the investor can borrow at time t, the time-t cash flows for both strategies are identical. At time T, there is no residual cash flow associated with the first strategy, while the second strategy results in a cash flow of $V_T - V_t e^{r(T-t)}$.

The residual cash flow for the second strategy illustrates why illiquidity may be costly to an investor. Even though the restricted investor can duplicate the time-t cash flow from a fully-marketable security position by borrowing, the investor remains vulnerable to subsequent price changes. If the value of the security increases from time t to time T, the residual cash flow $V_T - V_t e^{r(T-t)}$ is positive, and vice versa. It is easily shown that the expected value of this cash flow is zero under the risk-neutral measure. Despite this, however, the restricted investor may not be indifferent to bearing this residual price risk. For example, the restricted investor may believe that he has market-timing ability and would consider the mean of the residual cash flow $V_T - V_t e^{r(T-t)}$ to be negative, conditional on his private information. Note that the largest possible loss to the restricted investor from this strategy is $\max_\tau V_\tau e^{r(T-\tau)} - V_T$, where the maximum is taken over all values of τ, $0 \leq \tau \leq T$.

To provide an upper bound on the value of the discount for lack of marketability $V - R$, observe that the sum of the residual cash flow $V_T - V_t e^{r(T-t)}$ and the largest possible loss $\max_\tau V_\tau e^{r(T-\tau)} - V_T$ is always greater than or equal to zero. This means that if the restricted investor were given a cash flow at time T equal to $\max_\tau V_\tau e^{r(T-\tau)} - V_T$, then the total cash flow received by the restricted investor at time T would always be greater than or equal to zero. Intuitively, since this cash flow is equal to the greatest possible loss that the investor could experience; receiving this cash flow fully compensates the investor for any actual realized loss. Let $D(V,T)$ designate the present value of a contingent claim with time-T payoff $\max_\tau V_\tau e^{r(T-\tau)} - V_T$. Table 1 shows that when the contingent claim is incorporated into the second strategy, the cash flows from the second strategy dominate the cash flows from the first strategy. Thus, the inequality $V \leq R + D(V,T)$ must hold in order to avoid arbitrage. Hence, $V - R \leq D(V,T)$ and the value of $D(V,T)$ provides an upper bound on the value of the discount for lack of marketability for any investor. Table 2 illustrates that a similar argument can be used to show that $D(V,T)$ provides an upper bound in the case where $T \leq t$.

Table 2. Cash Flows from the Unrestricted and Restricted Strategies When the Unrestricted Security is Sold After the End of the Marketability Restriction Period

Strategy	*Time-0 Cash Flow*	*Time-T Cash Flow*	*Time-t Cash Flow*
Strategy 1			
Buy Security	$-V$		
Sell Security			$+V_t$
Total	$-V$	0	$+V_t$
Strategy 2			
Buy Security	$-R$		
Sell Security			$+V_t$
Buy Claim	$-D$		
Claim Payoff		$\max_\tau(V_\tau e^{r(T-\tau)} - V_T)$	
Total	$-R-D$	≥ 0	$+V_t$

Notes: The first strategy represents the cash flows to an unrestricted investor from taking a position in a marketable security at time zero and selling at some arbitrary time t. The second strategy represents the cash flows to an investor from taking a position in the same security but with the restriction that the investor cannot sell the security prior to time T, where $T < t$. V denotes the market value of the security. R is the value of the security to the investor given the marketability restriction. D denotes the market value of a contingent claim with the indicated payoff function at time T, where the maximum is taken over all τ, $0 \leq \tau \leq T$.

Proposition 1. The discount for lack of marketability $V-R$ satisfies the following inequality

$$0 \leq V - R \leq D(V,T), \tag{2}$$

where

$$D(V,T) = \left(2V + \frac{\sigma^2 T}{2} V\right) N\left(\frac{\sqrt{\sigma^2 T}}{2}\right) + V\sqrt{\frac{\sigma^2 T}{2\pi}} \exp\left(-\frac{\sigma^2 T}{8}\right) - V,$$

and $N(\cdot)$ is the cumulative normal density function.

Proof. See Appendix.

The lower bound on the discount for lack of marketability is clearly zero. The upper bound is the present value of the maximum possible loss that the investor could incur during the period of the marketability restriction. Alternatively, the upper bound can also be viewed as the present value of the opportunity cost suffered by an investor with perfect market timing ability by only being allowed to liquidate his security position at time T.

Intuitively, this liquidity derivative is similar in many respects to a standard lookback type of option in which the investor is given the difference between the maximum value of the underlying security during the life of the derivative and some fixed amount. What is different, however, is that the payoff of the derivative is the difference between the maximum and the time-T value of the underlying. Furthermore, the maximum involved is not simply the maximum value taken by the underlying asset. In particular, if the investor were free to time the sale optimally, the investor would sell at a time that would maximize the present value of the selling price. This would not necessarily coincide with the time when the asset price reached its maximum value. For example, if $r = 0.05$ and V reached 100 at time 1 and 102 at time 2, and investor with perfect market timing ability would maximize the present value of his information by selling at time 1 rather than time 2. Thus, this contingent claim differs in several ways from lookback types of contingent claims on V such as those considered by Goldman, Sosin, and Shepp (1979) and Goldman, Sosin, and Gatto (1979).

This upper bound represents the maximum amount that an investor would be willing to pay to obtain immediacy in liquidating a security position. Thus, this upper bound provides an endogenous measure of the largest possible bid-ask spread or transaction cost for a security. Equivalently, this upper bound represents the spread that would be charged by a market maker who knew that the investor had perfect information. In contrast, previous research on the valuation of illiquid securities by Amihud and Mendelson (1986) and Boudoukh and Whitelaw (1993) and on the valuation of securities in the present of transaction costs by Constantinides (1986) and Vayanos and Vila (1992) takes the bid-ask spread or transaction costs for the security as exogenous.

The maximum discount for lack of marketability is proportional to V and depends on the volatility parameter σ and the length of the marketability restriction T. Differentiation shows that the maximum discount is an increasing function of the variance parameter. Intuitively, this is because the value of perfect market timing ability increases with the variance of the price process since the ultimate selling price is likely to be larger. The maximum discount is also an increasing function of T. It is interesting to observe that the maximum discount does not depend on the riskless interest rate r. The reason for this is that the expected return on V is r under the risk-neutral pricing measure. This expected return just cancels the discount rate of r applied in deriving the present value $D(V,T)$ of the contingent claim.

To illustrate the relation between the discount for nonmarketability and the length of the restriction period T, Figure 1 graphs the maximum discount as a percentage of the value V for values of T ranging from zero to 48 hours and for $\sigma = 0.10$, 0.20, and 0.30. These percentage values represent upper bounds on the present value loss that an investor would incur by agreeing not to sell the security prior to time T. This situation might occur, for example, if the investor acquired the security as part of a public offering and is committed to holding the security to stabilize the aftermarket performance. Alternatively, the values provide upper bounds for the implicit cost to a broker from lending shares of stock on a short-term basis to customers who wish to sell short.

As shown, the maximum percentage discount can be quite large even for short periods of nonmarketability. For example, the maximum percentage discount for a 24-hour marketability restriction is 0.42, 0.84, and 1.26 for $\sigma = 0.10$, 0.20, and 0.30. The maximum percentage discount is an increasing function of σ, and an increasing concave function of T. Note that the maximum percentage discount for $T = 48$ hours is only

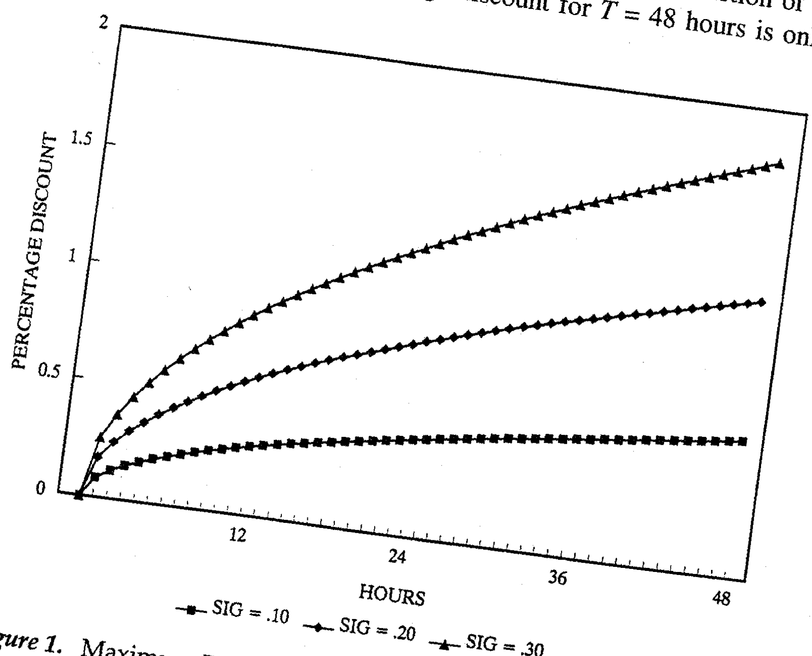

Figure 1. Maximum Percentage Discount for Nonmarketability for Restriction Periods Ranging from Zero to 48 Hours and for Annualized Standard Deviations of Returns of 0.10, 0.20, and 0.30

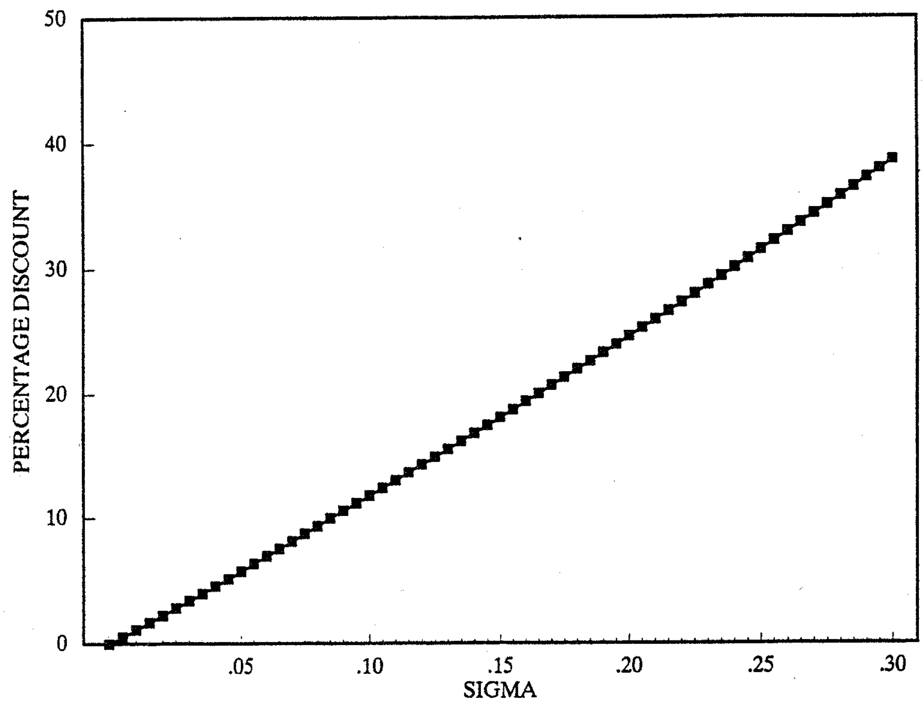

Figure 2. Maximum Percentage Discount for Nonmarketability for a Two-year Restriction Period and for Annualized Standard Deviations of Returns Ranging from Zero to 0.30

about twice that for $T = 12$ hours, and only about seven times that for $T = 1$ hours. Thus, the relation between the length of the restriction period T and the maximum percentage discount is very steep at first, but quickly declines. It is easily shown that the maximum percentage discount converges to zero as $T \rightarrow 0$. As $T \rightarrow \infty$, the maximum percentage discount increases without bound. As a practical matter, however, the maximum percentage discount can never exceed 100%.

To compare the empirical estimates of the discount for nonmarketability observed for restricted stock with the upper bound, Figure 2 graphs the maximum percentage discount for a restriction period of $T = 2$ years as a function of σ. Recall that the average percentage discount for restricted stock is roughly 35%. Figure 2 shows that this empirical estimate lies within the upper bound only if the average value of σ for the firms issuing restricted stock is greater than or equal to 0.28. Unfortunately, data on the average standard deviation of firms included in the restricted stock studies is not available. Typical estimates of the average standard deviation of firms on the NYSE are in the range of 0.15–0.25

during the period covered by the restricted stock studies. Typical estimates of the average standard deviation for the smallest 10% of the firms on the NYSE during the same period are in the range of 0.25–0.30. Although we cannot draw definitive conclusions from these informal comparisons, these results seem to suggest that the observed discounts for restricted stock approximate or perhaps exceed their upper bounds.[4]

THINLY-TRADED SECURITIES

In the previous section, we focused on securities which could not be sold for a fixed period of time. In this section, we consider the more typical case where the investor can choose to sell at any time, but where the thinness of the market may lead to a delay before the sale can be executed.

Our approach is again to identify a contingent claim which would fully compensate an investor for foregoing the right to sell the security at any point in time. To model thin trading, we assume that the restriction is that the investor can only sell the security at discrete points in time, say every 10 minutes. Thus, if the investor chooses to sell, the investor must wait between 0 and 10 minutes before the security can be sold.[5] The implicit cost to the investor is that the market price of the security may decline before the security can be sold. Consequently, the value of the security to the investor may be less than if the security could be sold continuously in the market. We term this the thin-trading discount.

Let L denote the length of the interval between points in time at which the investor can sell his security position. Let T denote the horizon or life of the security. Thus, $I = T/L$ is the number of discrete points in time during the life of the security at which the investor can sell.[6] The upper bound on the thin-trading discount is obtained by again specifying a contingent claim which fully compensates the investor for any possible loss due to the marketability restriction. Consider the two investment strategies shown in Table 3. The first is again the unrestricted case in which the security is purchased at time zero for a cash outlay of V, and then sold at some arbitrary time t for V_t. The second strategy is the restricted case in which the investor pays R for the security, borrows V_t at time t, liquidates the security at the next opportunity, and uses the proceeds to pay down the balance of the borrowings. Let N denote the next time after t at which the security can be sold by the investor. Clearly, N is an integral multiple of L. If the balance due on the borrowings at time N, $V_t e^{r(N-t)}$ is less than the proceeds from selling the security V_N, the balance is assumed to be invested at the riskless rate until time T.

Table 3. Cash Flows from the Unrestricted and Thin-trading Strategies

Strategy	*Time-0 Cash Flow*	*Time-t Cash Flow*	*Cash Flow at Next N*	*Time-T Cash Flow*
Strategy 1				
Buy Security	$-V$			
Sell Security		$+V_t$		
Total	$-V$	$+V_t$	0	0
Strategy 2				
Buy Security	$-R$			
Sell Security			$+V_N$	
Borrow		$+V_t$		
Pay Down Loan			$-V_N$	
Loan Balance				$e^{rT}(V_N e^{-rN} - V_t e^{-rt})$
Buy Claim	$-Q$			
Claim Payoff				$\max_I(e^{rT}\max_\tau(V_\tau e^{-r\tau} - V_N e^{-rN}))$
Total	$-R-Q$	$+V_t$	0	≥ 0

Notes: The first strategy represents the cash flows to an unrestricted investor from taking a position in a marketable security at time zero and selling at some arbitrary time $t \leq T$, where T is the life or horizon of the security. The second strategy represents the cash flows to an investor from taking a position in the same security but with the thin-trading restriction that the investor can only sell at I regularly-spaced discrete points in time $1, 2, 3,...N-1, N,...I$, where N denotes the general term. V denotes the market value of the security. R is the value of the security to the investor given the thin-trading restriction. Q denotes the market value of a contingent claim with the indicated payoff function at time T, where the inner maximum is taken over all τ, $N-1 \leq \tau \leq N$, and the outer maximum is taken over all I discrete periods.

The second strategy exactly matches the cash flow at time t available to an investor without marketability restrictions, but has a residual cash flow at time T of $e^{rT}(V_N e^{-rN} - V_t e^{-rt})$. A contingent claim that guarantees the restricted investor a positive net cash flow at time T can be constructed as follows. First, for each of the I periods of length L during $[0, T]$, find the τ that maximizes $V_\tau e^{r(T-\tau)} - V_N e^{r(T-N)}$, where $N - L \leq \tau \leq N$. Second, of these I values, find the maximum. Finally, create a contingent claim which makes a cash flow at time T equal to this maximum. It is easily shown that the sum of this cash flow and the residual cash flow from the second strategy is greater than or equal to zero. Thus, the sum of the R and the present value of this contingent claim $Q(V,T,L)$ must be less than or equal to V in order to avoid arbitrage. This implies the upper bound on the discount for thin trading.

Proposition 2. The thin-trading discount $V - R$ satisfies the following inequality

$$0 \le V - R \le Q(V,T,L), \tag{5}$$

where

$$Q(V,T,L) = (V + D(V,T))\, I \int_0^{\infty} v F^{I-1}(v) f(v)\, dv,$$

$$F(v) = N\left(\frac{v - \sigma^2 L/2}{\sqrt{\sigma^2 L}}\right) - \exp(v) N\left(\frac{-v - \sigma^2 L/2}{\sqrt{\sigma^2 L}}\right),$$

$$f(v) = \frac{2}{\sqrt{2\pi\sigma^2 L}} \exp(v) \exp\left(\frac{-(v + \sigma^2 L/2)^2}{2\sigma^2 L}\right)$$

$$- \exp(v) N\left(\frac{-v - \sigma^2 L/2}{\sqrt{\sigma^2 L}}\right).$$

Proof. See Appendix.

Intuitively, the upper bound on the thin-trading discount is the present value of the largest possible loss during any of the discrete periods during which the investor cannot sell the security. This upper bound differs fundamentally from that for the discount for nonmarketability in Proposition 1. As an example, let L equal 24 hours, let T equal one year, and let σ equal 0.20. The upper bound for a 24-hour nonmarketability discount is 0.84%. In contrast, the upper bound for the thin-trading discount is 3.86%. The reason for the difference is that the upper bound for the discount for nonmarketability is related to the expected maximum loss over a 24-hour period. The upper bound on the thin-trading discount is also related to this maximum loss. However, the upper bound for the thin-trading discount also takes into account that there are 365 separate 24-hour trading periods during the life of the security. Thus, the upper bound for the thin-trading discount is related to the expected maximum or first-order statistic for 365 losses rather than simply a single 24-hour loss. Note also that the upper bound for the thin-trading discount is much smaller that 365 times the upper bound for the 24-hour nonmarketability discount. In addition, it is much smaller than the 16.98% upper bound for a one-year nonmarketability discount.

To illustrate the upper bound on the thin-trading discount, Figure 3 graphs the maximum percentage thin-trading discounts for three-month, six-month, and one-year Treasury securities as functions of the length of the thin-trading interval L, where L ranges from zero to 60 minutes. The values of σ used to compute the upper bounds correspond to the average annual standard deviation of returns on these securities during the 1964–1993 period. In particular, the values of σ for the three-month, six-month, and one-year Treasury securities are 0.00625, 0.0125, and 0.0250, respectively. As shown, the maximum thin-trading discount is an increasing function of L. In addition, the maximum thin-trading discount increases as we move from shorter-term to longer-term Treasury securities.

Using these values, we can compare the theoretical upper bounds with the average discounts observed for less-liquid Treasury securities. Recall that the results in Amihud and Mendelson (1991) imply an average difference between the prices of liquid Treasury bills and less-liquid Treasury notes and bonds with an average maturity of 97 days of about 0.12%. In contrast, the upper bound shown in Figure 3 for 3-month Treasury securities is always less than 0.03%. In fact, an average thin-

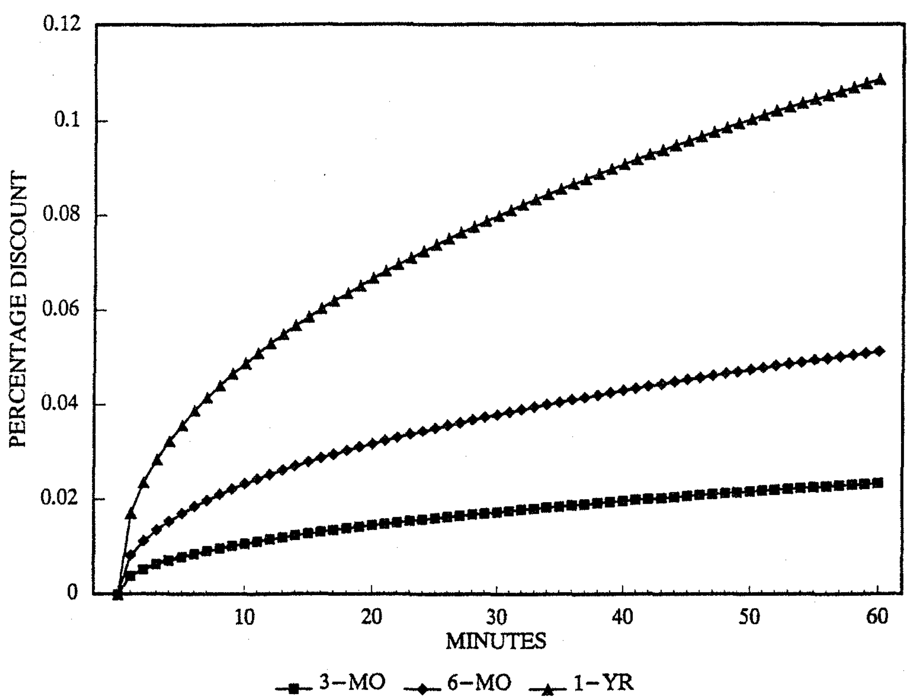

***Figure* 3.** Maximum Percentage Thin-trading Discount for Three-month, Six-month, and One-year Treasury Securities for Thin-trading Intervals Ranging from Zero to 60 Minutes trading discount of 0.12% only lies within the upper bound if the value of L is greater than about 48 hours. While Treasury bills can be sold almost instantaneously, selling less-liquid Treasury notes and bonds with comparable maturities may involve a delay. Nevertheless, conversations with traders suggest that this delay can be measured in minutes rather than hours or days. Thus, these results suggest that the discounts reported by Amihud and Mendelson are much larger than the upper bound.

As another comparison, Figure 4 graphs the maximum percentage thin-trading discounts for 10-year, 15-year, and 30-year Treasury securities, again using historical estimates of σ of 0.0375, 0.0750, and 0.1000, respectively. The results from comparing these bounds to reported estimates of the thin-trading discount are similar to those for the short-term securities. The upper bound for a 15-year Treasury bond when $L = 60$ minutes is only about 0.5%, which is well below the average empirical estimate of 0.86 observed for Treasury Strips. The upper bound for a 30-year Treasury bond when $L = 60$ minutes is about 0.8%. As shown by Cornell and Shapiro (1990), off-the-run discounts can be several

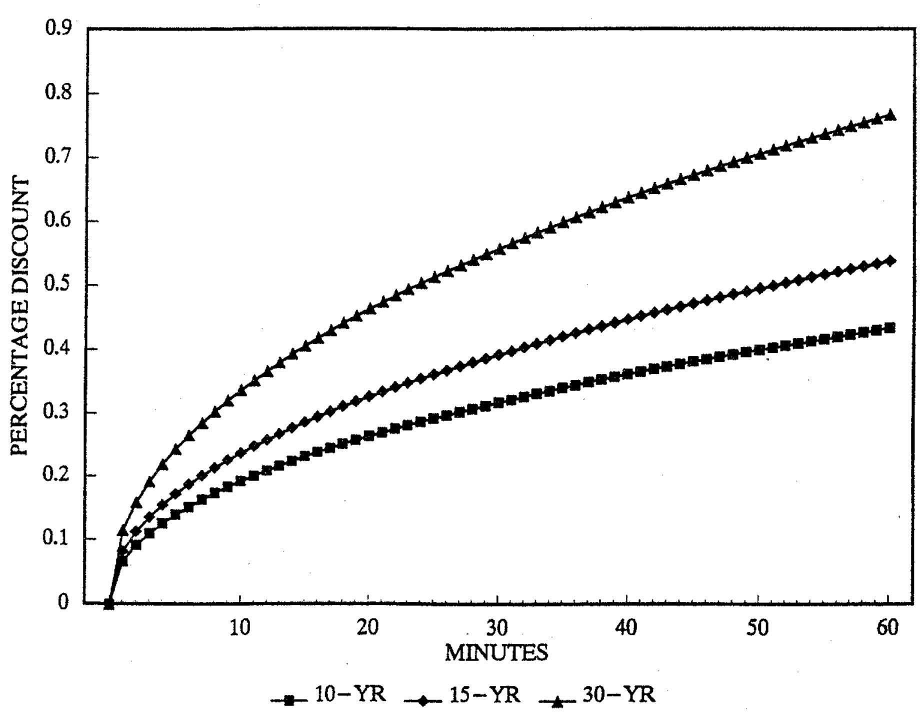

Figure 4. Maximum Percentage Thin-trading Discount for 10-year, 15-year, and 30-year Treasury Securities for Thin-trading Intervals Ranging from Zero to 60 Minutes

times this value. Furthermore, the average benchmark effect during 1984–1987 in the Japanese government bond market is only consistent with the upper bound for values of L on the order of a week. Taken together, these comparisons indicate that thin-trading discounts observed in fixed income markets often exceed the upper bound by substantial amounts.

CONCLUSION

We have derived analytical upper bounds on the discount for lack of marketability in a rational no-arbitrage framework. If the discount exceeds this bound, then a contingent claim could be constructed that would guarantee the holder of the nonmarketable or thinly-traded security an arbitrage profit. This contingent claim represents the present value of the maximum possible loss that the investor might experience by being unable to sell the security at his discretion. These result suggest that return volatility can be a key determinant of discounts for lack of marketability and that discounts can be quite large even if the security can be sold within a relatively short time period.

These analytical bounds on the value of marketability can provide useful information for making policy decisions. For example, these bounds allow the policy maker to assess the maximum effect of imposing a restriction on trading such as a circuit breaker, trading halt, or price limit. Typically, these types of trading restrictions have been proposed by regulators without quantifying the costs imposed on traders and investors because of the temporary impairment of liquidity. Since these restrictions are generally imposed during periods of high volatility, this analysis suggests that the implicit cost to traders and investors could be much higher than is commonly believed.

Since these upper bounds are based on a worst-case scenario, it might be expected that observed discounts would be much less than the upper bound. Surprisingly, many of the estimates given in the literature approximate or exceed the no-arbitrage bound. It is important to acknowledge, however, that these results are based on informal comparisons rather than on formal tests. Clearly, future research should explore more carefully the issue of whether nonmarketability and thin-trading discounts are too large to be consistent with a rational or equilibrium asset pricing model.

APPENDIX

Proof of Proposition 1. The lower bound for $V-R$ cannot be greater than zero since it may be optimal for the investor to buy and hold. Thus, the value of the rights foregone because of the restriction on marketability could be zero. The lower bound cannot be less than zero because the restriction on marketability results in a strictly smaller set of portfolio opportunities for the investor. Tables 1 and 2 show that the cash flows from the second strategy are equal to or greater than the cash flows from the first strategy in every state. Hence, the inequality $V-R \leq D(V,T)$ must hold in order to avoid arbitrage. Since V is a traded security and r is constant, the value of $D(V,T)$ is given by applying the risk-neutral valuation model,

$$D(V,T) = e^{-rT}E[\max_{\tau} V_{\tau}e^{r(T-\tau)} - V_T], \tag{A1}$$

where the expectation is taken with respect to the risk-neutral process

$$dV = rVdt + \sigma VdZ, \tag{A2}$$

and the maximum is taken over all τ, $0 \leq \tau \leq T$. This implies

$$D(V,T) = E[\max_{\tau} V_{\tau}e^{-r\tau}] - e^{-rT}E[V_T], \tag{A3}$$

$$D(V,T) = E[\max_{\tau} V_{\tau}e^{-r\tau}] - V. \tag{A4}$$

Define the process $X_t = \ln(V_te^{-rt}/V)$. This process has time-zero value $X = 0$ and dynamics

$$dX = -\sigma^2/2dt + \sigma dZ. \tag{A5}$$

Define H_T *as* $\max_{\tau}X_{\tau}$, where the maximum is taken over all τ, $0 \leq \tau \leq T$. Thus,

$$D(V,T) = VE[\exp(H_T)] - V. \tag{A6}$$

The density function for H_T follows from Harrison (1985) p. 14 equation (11),

$$\frac{2}{\sqrt{2\pi\sigma^2T}}\exp\left(-\frac{(H_T+\sigma^2T/2)^2}{2\sigma^2T}\right) + \exp(-H_T)N\left(\frac{-H_T+\sigma^2T/2}{\sqrt{\sigma^2T}}\right), \tag{A7}$$

where $0 \le H_T < \infty$. The expectation $E[\exp(H_T)]$ equals

$$\frac{2}{\sqrt{2\pi\sigma^2 T}} \int_0^\infty \exp\left(\frac{-(H_T + \sigma^2 T/2)^2 + 2\sigma^2 T H_T}{2\sigma^2 T}\right) dH_T$$

$$+ \int_0^\infty N\left(\frac{-H_T + \sigma^2 T/2}{\sqrt{\sigma^2 T}}\right) dH_T. \tag{A8}$$

The first integral in equation (A8) equals

$$2\int_0^\infty \frac{1}{\sqrt{2\pi\sigma^2 T}} \exp\left(\frac{-(H_T - \sigma^2 T/2)^2}{2\sigma^2 T}\right) dH_T, \tag{A9}$$

which equals

$$2N\left(\frac{\sqrt{\sigma^2 T}}{2}\right). \tag{A10}$$

The second integral in equation (A8) reduces to

$$\int_0^\infty \frac{H_T}{\sqrt{2\pi\sigma^2 T}} \exp\left(\frac{-(H_T - \sigma^2 T/2)^2}{2\sigma^2 T}\right) dH_T, \tag{A11}$$

after integration by parts. This integral can be evaluated directly by a change of variables resulting in

$$\left(\frac{\sigma^2 T}{2}\right) N\left(\frac{\sqrt{\sigma^2 T}}{2}\right) + \sqrt{\frac{\sigma^2 T}{2\pi}} \exp\left(-\frac{\sigma^2 T}{8}\right). \tag{A12}$$

Substituting the sum of equations (A10) and (A12) for $E[\exp(H_T)]$ in equation (A6) and rearranging terms gives the result.

Proof of Proposition 2. The lower bound for $V - R$ cannot be greater than zero since it may be optimal for the investor to buy and hold. Thus, the value of the rights foregone because of the restriction on marketability could be zero. The lower bound cannot be less than zero because the restriction on marketability results in a strictly smaller set of portfolio opportunities for the investor. Denote the length of time between the discrete periods at which the investor can trade by L. The I times at which the investor can sell the security are $L, 2L, 3L, ..., (N-1)L, NL, ... IL = T$, where T is the life or horizon of the security. Table 3 shows that the cash flows from the second strategy are equal to or greater than the cash flows from the first strategy in every state. Hence, the inequality $V - R \leq Q(V,T,L)$ must hold in order to avoid arbitrage. The value of $Q(V,T,L)$ is again given by applying the risk-neutral valuation model,

$$Q(V,T,L) = e^{-rT} E[\max_I (e^{rT} \max_\tau (V_\tau e^{-r\tau} - V_{NL} e^{-rNL}))], \qquad \text{(A13)}$$

where the inner maximum is taken over all τ, $(N-1)L \leq \tau \leq NL$ and the outer maximum is taken over all I discrete trading times. As in the previous proof, let $X_t = \ln(V_t e^{-rt}/V)$. Define M_{NL} as the maximum value of X_t during the Nth period $[(N-1)L, NL]$. From equation (A13),

$$Q(V,T,L) = VE[\max_I(\max_\tau(e^{X_\tau} - e^{X_{NL}}))], \qquad \text{(A14)}$$

$$Q(V,T,L) = VE[\max_I(e^{M_{NL}} - e^{X_{NL}})], \qquad \text{(A15)}$$

$$Q(V,T,L) = VE[\max_I(e^{M_{NL}}(1 - e^{X_{NL}-M_{NL}}))]\,. \qquad \text{(A16)}$$

Define H_T as the maximum value of X during the life of the security or the security restriction period $[0,T]$. Since $H_T \in \{M_L, M_{2L}, \ldots, M_{IL}\}$,

$$Q(V,T,L) \leq VE\,[e^{H_T} \max_I(1 - e^{X_{NL}-M_{NL}})], \qquad \text{(A17)}$$

$$Q(V,T,L) \leq VE\,[e^{H_T}(1 - \min_I e^{X_{NL}-M_{NL}})]. \qquad \text{(A18)}$$

However, since $X_{NL} \leq M_{NL}$ for all N, $e^{X_{NL}-M_{NL}} \geq 1 + X_{NL} - M_{NL}$. Thus,

$$\min_I(1 + X_{NL} - M_{NL}) \leq \min_I(e^{X_{NL}-M_{NL}}), \qquad \text{(A19)}$$

which implies

$$Q(V,T,L) \leq VE\,[e^{H_T} \max_I(M_{NL} - X_{NL})]. \tag{A20}$$

Because the increments of X are independent and their distribution does not depend on the level of X, and because $L, 2L, ..., IL$ are stopping times, the distribution of each $M_{NL} - X_{NL}$ is the same as $M_L - X_L$. Changing variables to $u = M_L$ and $v = M_L - X_L$ and using the joint density for M_L and X_L given by Harrison (1980) 12, equation (3) implies the joint density of u and v,

$$\frac{2(u+v)}{\sigma^2 L\sqrt{2\pi\sigma^2 L}}\exp(v)\exp\left(-\frac{(u+(v+\sigma^2 L/2))^2}{2\sigma^2 L}\right). \tag{A21}$$

Integrating out u gives the density $f(v)$ for v,

$$f(v) = \frac{2}{\sqrt{2\pi\sigma^2 L}}\exp(v)\exp\left(-\frac{(v+\sigma^2 L/2)^2}{2\sigma^2 L}\right) - \exp(v)N\left(\frac{-v-\sigma^2 L/2}{\sqrt{\sigma^2 L}}\right). \tag{A22}$$

Integrating equation (A22) gives the distribution function F(v) for v,

$$F(v) = N\left(\frac{v-\sigma^2 L/2}{\sqrt{\sigma^2 L}}\right) - \exp(v)N\left(\frac{-v-\sigma^2 L/2}{\sqrt{\sigma^2 L}}\right). \tag{A23}$$

Using the joint density for u and v in equation (A21) and solving for $E[\exp(u)v]$, $E[\exp(u)]$, and $E[v]$, it can be shown that $e^{M_{NL}}$ and $M_{NL} - X_{NL}$ are negatively correlated. Since M_{JL} and $M_{KL} - X_{KL}$ are independent when $J \neq K$, and $H_T \in \{M_L, M_{2L}, ..., M_{IL}\}$, it follows that e^{H_T} and $\max_I(M_{NL} - X_{NL})$ are negatively correlated. From equation (A20) and the definition of the covariance, this implies

$$Q(V,T,L) \leq VE[\exp(H_T)]E\,[\max_I(M_{NL} - X_{NL})]. \tag{A24}$$

However, the distribution of $\max_I(M_{NL} - X_{NL})$ is the same as the distribution of the first order statistic for I observations of $M_L - X_L$. Thus,

$$E[\max_I(M_{NL} - X_{NL})] = I\int_0^{\infty} vF^{I-1}(v)f(v)dv. \tag{A25}$$

Substituting this expectation into equation (A24) and recalling the definition of $D(V,T)$ gives the result.

ACKNOWLEDGMENTS

I am grateful for the comments of Michael Brennan, Julian Franks, Mark Grinblatt, Raymond Kan, Jeffrey Pontiff, Eduardo Schwartz, Jean-Luc Vila, and Pradeep Yadav and seminar participants at the AMEX Options Colloquium, Case Western Reserve University, the London School of Business, McGill University, the University of Strathclyde, the University of Toronto, and the University of Vienna. All errors are my responsibility.

NOTES

1. In his sample, Silber (1992) finds that the firms with the largest discounts tend to be smaller in size. In general, however, little is known about the cross-sectional properties of discounts for letter stock.
2. By foregoing the right to sell, we include strategies which synthetically sell the security such as short sales, short futures positions, short call positions, and so forth.
3. We are implicitly making the standard no-arbitrage assumption that the market price V of the underlying asset is exogenously determined and is unaffected by whether the offer to forego the right to sell is made or not. However, nothing in this framework precludes V from being the equilibrium price determined in a market in which this type of an offer could be made. Note that ϕ can be viewed as the upper bound on the shadow price of the nonmarketability restriction.
4. The upper bounds could potentially be slightly different under different assumptions about the dynamics of the underlying process. For example, the bounds would be tighter if the pricing process was mean reverting. Alternatively, the bounds could be wider if the pricing process displayed stochastic volatility or exhibited jumps.
5. There are other possible ways of modeling thin trading. For example, the investor could be required to wait some fixed period of time after placing a sell order until the sale is executed. The model we use incorporates the more realistic feature that the time until the sale is executed may be random. Alternatively, thin-trading could also be modelled as a situation in which the investor is willing to accept a price concession in order to obtain an immediate sale. The upper bound derived in this section represents an upper bound on the price concession that the investor would pay for immediacy.
6. For notational simplicity, we assume that I is an integer.

REFERENCES

Amihud, Y., and H. Mendelson. 1986. "Asset Pricing and the Bid-Ask Spread." *Journal of Financial Economics* 17: 223–249.

———. 1991. "Liquidity, Maturity, and the Yields on U. S. Treasury Securities." *The Journal of Finance* 46: 1411–1425.

Barclay, M., C. Holderness, and J. Pontiff. 1994. "Private Benefits from Block Ownership and Discounts on Closed-End Funds." *Journal of Financial Economics*.

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

Boudoukh, J., and R. F. Whitelaw. 1991. "The Benchmark Effect in the Japanese Government Bond Market." *The Journal of Fixed Income* (September): 52–59.

———. 1993. "Liquidity as a Choice Variable: A Lesson from the Japanese Government Bond Market." *The Review of Financial Studies* 6: 265–292.

Brito, N. O. 1977. "Marketability Restrictions and the Valuation of Capital Assets Under Uncertainty." *Journal of Finance* 32: 1109–1123.

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

Cornell, B. 1993. "Adverse Selection, Squeezes, and the Bid-Ask Spread on Treasury Securities." *The Journal of Fixed Income* (June): 39–47.

Cornell, B., and A. C. Shapiro. 1990. "The Mispricing of U. S. Treasury Bonds: A Case Study," *The Review of Financial Studies* 2: 297–310.

Daves, P. R., and M. C. Ehrhardt. 1993. "Liquidity, Reconstitution, and the Value of U.S. Treasury Strips." *The Journal of Finance* 48: 315–329.

Gelman, M. 1972. "An Economist-Financial Analyst's Approach to Valuing Stock of a Closely Held Company." *Journal of Taxation* (June): 353–354.

Goldman, M. B., H. B. Sosin, and M. A. Gatto. 1979. "Path Dependent Options: 'Buy at the Low, Sell at the High.'" *Journal of Finance* 34: 1111–1127.

Goldman, M. B., H. B. Sosin, and L. A. Shepp. 1979. "On Contingent Claims that Insure Ex-Post Optimal Stock Market Timing." *Journal of Finance* 34: 401–414.

Grinblatt, M., and F. L. Longstaff. 1995. "Liquidity and the Pricing of Treasury Securities." Working paper, UCLA.

Harrison, J. M. 1985. *Brownian Motion and Stochastic Flow Systems*. New York: John Wiley.

Institutional Investor Study Report of the Securities and Exchange Commission. 1971. Discounts Involved in Purchases of Common Stock (1966–1969), H.R. Doc. No. 64, Part 5, 92nd Congress, 2444–2456.

Kamara, A. 1994. "Liquidity, Taxes, and Short-Term Treasury Yields." *Journal of Financial and Quantitative Analysis*.

Lippman, S., and J. McCall. 1986. "An Operational Measure of Liquidity." *American Economic Review* 76: 43–55.

Maher, J. M. 1976. "Discounts for Lack of Marketability for Closely-Held Business Interests." *Taxes* (September): 562–571.

Mayers, D. 1972. "Nonmarketable Assets and Capital Market Equilibrium Under Uncertainty." In *Studies in the Theory of Capital Markets*, edited by M. C. Jensen. New York: Praeger.

———. 1973. "Nonmarketable Assets and the Determination of Capital Asset Prices in the Absence of Riskless Asset." *Journal of Business* 46: 258–267.

———. 1976. "Nonmarketable Assets, Market Segmentation and the Level of Asset Prices." *Journal of Financial and Quantitative Analysis* 11: 1–12.

Moroney, R. E. 1973. "Most Courts Overvalue Closely Held Stock." *Taxes* (March): 144–154.

Pratt, S. P. 1989. *Valuing a Business*. Homewood, IL: Irwin.

Rogg, O. 1991. "Repurchase Agreements." In *Handbook of Fixed Income Securities*, edited by F. Fabozzi. Homewood, IL: Irwin.

Silber, W. L. 1992. "Discounts on Restricted Stock: The Impact of Illiquidity on Stock Prices." *Financial Analysts Journal* (July/August): 60–64.

SRC Quarterly Reports. 1983. "Revenue Ruling 77-287 Revisited." (Spring): 1–3.

Stapleton, R. C., and M. G. Subrahmanyam. 1979. "Marketability of Assets and the Price of Risk." *Journal of Financial and Quantitative Analysis* 14: 1–10.

Trout, R. R. 1977. "Estimation of the Discount Associated with the Transfer of Restricted Securities." *Taxes* (June): 381–385.

Vayanos, D., and J. L. Vila. 1992. "Equilibrium Interest Rate and Liquidity Premium Under Proportional Transactions Costs." Working paper, Sloan School of Management, MIT.

Wruck, K. H. 1989. "Equity Ownership Concentration and Firm Value: Evidence from Private Equity Financings." *Journal of Financial Economics* 23: 3–28.