

11 November 2002

## **A CAPM-Based Approach to Calculating Illiquidity Discounts**

*by Dr. David Tabak*

**NERA**  
Economic Consulting



Marsh & McLennan Companies

# A CAPM-Based Approach to Calculating Illiquidity Discounts

David I. Tabak\*

## Contents

<b>I.</b>	<b>INTRODUCTION.....</b>	<b>2</b>
<b>II.</b>	<b>LITERATURE REVIEW .....</b>	<b>4</b>
A.	EMPIRICAL STUDIES.....	4
B.	THEORETICAL SOLUTIONS.....	6
1.	<i>Chaffee Put Analysis, or Accounting for the Possibility of Loss.....</i>	<i>6</i>
2.	<i>Longstaff Upper Bound Analysis, Or Accounting for the Possibility of Missing Out on Upward Gains.....</i>	<i>7</i>
3.	<i>Mercer Required Rate of Return Analysis.....</i>	<i>8</i>
<b>III.</b>	<b>A THEORY BASED ON THE CAPM AND THE EQUITY RISK PREMIUM.....</b>	<b>10</b>
A.	WHAT DOES THE EQUITY RISK PREMIUM MEASURE?.....	10
B.	A THEORY OF APPLYING THE EQUITY RISK PREMIUM TO AN ILLIQUID SECURITY.....	11
C.	A THEORY OF APPLYING THE EQUITY RISK PREMIUM TO A DISCOUNTED CASH FLOW VALUATION OF A FIRM .....	17
D.	DETERMINING THE LIKELY LENGTH OF THE PERIOD OF ILLIQUIDITY .....	19
E.	ADDITIONAL CONSIDERATIONS FOR THE MODEL .....	21
F.	CONCLUSION.....	22

---

\* David Tabak, Vice President, National Economic Research Associates, Inc. I would like to thank Lou Guth for many helpful comments on earlier drafts and Todd Foster and Rashmi Khare for assistance in data collection and analysis.

## I. Introduction

“In many valuations of closely held businesses or business interests, the discount for lack of marketability turns out to be the largest single issue to resolve – and an important issue in almost all such valuations.”

-- Shannon Pratt<sup>1</sup>

A discount for illiquidity or for a lack of marketability is the reduction in the value of an asset because it cannot be easily sold. These discounts are often applied in the valuations of securities such as restricted stock, and in the valuations of businesses or business opportunities, generally privately held businesses. The discounts can vary from zero, in cases where the analyst feels that the restrictions on marketability are minimal, to average discounts around 35%, to even larger ranges advocated by some authors.<sup>2</sup> Given this large variance, one would hope that such discounts could be objectively quantified. In the litigation arena specifically, one would like a theory that could be replicated by opposing experts and that has standards controlling its operation.<sup>3</sup>

Unfortunately, only the crudest versions of such a theory currently exist. As discussed below, while there is some (probably biased) evidence on the average size of the marketability discount in the past, this evidence does not allow an analyst to determine the discount for a particular case without making subjective decisions about how that case compares to the average. There are also a limited number of theories, several of which are discussed below, that do provide some means of estimating the discount applicable to a particular firm or security. While these theories clearly represent potential advancements over simply using a historic average, they rely on old data that are not easily updated or are simply a more sophisticated form of subjective decision about the importance of various impediments to marketability.

This paper instead presents a theoretical model based on the CAPM, or capital asset pricing model, that allows for a quantification of the illiquidity discount based as much as possible on objective criteria. I believe it to be the first paper to apply the

---

<sup>1</sup> Shannon Pratt, Valuing a Business: The Analysis and Appraisal of Closely Held Companies, 1989.

<sup>2</sup> For example, one study argues that “[a]ppraisers can now cite a number of known arm’s length transactions in which the discount ranged up to 90 percent.” Moroney, Robert E. “Most Courts Overvalue Closely Held Stocks,” *Taxes – The Tax Magazine*, March 1973.

CAPM, or any theoretical model, to the determination of these discounts. Ultimately, there will be cases where the theory can be considered a complete success, because it will provide a methodology that will allow for such a calculation. Inevitably, in other cases, the restrictions on marketability will be vague, and thus the magnitude of their effects will still be subject to some interpretation. Nevertheless, even in those cases, the extent of estimation necessary will be greatly reduced by the theory described in this paper.

Finally, I note that it is often overlooked that there are two potential components to a discount for lack of marketability: the cost of bearing risk during period in which the asset cannot be sold, and the costs of not having immediate access to capital. In this paper, I consider principally the costs that the owner of an illiquid asset bears and propose a method for valuing those costs based on the economic theory and empirical observations of actual costs of risk-bearing in the securities markets. A brief discussion of the costs of lacking immediate access to capital is included at the end of this paper.

---

<sup>3</sup> More specifically, one would like a theory that satisfied the suggested criteria provided by the Supreme Court in *William Daubert v. Merrell Dow Pharmaceuticals, Inc.* 43 F. 3d 1311 (9<sup>th</sup> Cir. 1995)

## II. Literature Review

### A. Empirical Studies

There have been a number of studies on the discounts appropriate for restricted stock. Many of these are listed in Pratt or Mercer.<sup>4</sup> These studies generally show discounts of approximately 35% in examining the sample of sales prices of restricted stock with the contemporaneous price of unrestricted but otherwise identical stock.

One problem common to all of these studies is that they only look at cases in which restricted stock was actually sold. Yet, there are other ways in which the holder of a restricted stock can avoid the risk from that security. These may include selling short a similar amount of stock, engaging in an equity swap transaction, or purchasing a zero-cost collar.

By selling short a similar amount of stock, the owner of the restricted stock is completely hedged: any gains (losses) on the restricted stock are matched one-for-one with offsetting losses (gains) on the short position. The main obstacle to this type of transaction may come from an inability to short the security. Often the reason that an investor may own restricted stock is because she is an insider in a company, and therefore may be ineligible to sell the company's stock short.

A second way to hedge the risk of the restricted security is through an equity swap. An equity swap is simply an agreement to trade all of the cash flows from one equity for those of another. That is, all the dividend payments and the final capital gain or loss from each security is traded between two parties. In this manner, the holder of a restricted security can effectively receive the cash flows from another security, including a stock index or a bond portfolio.

A slightly more complicated method is to purchase a zero-cost collar for the security.<sup>5</sup> When an investor purchases a zero-cost collar, she is really purchasing a put,

---

<sup>4</sup> Z. Christopher Mercer, Quantifying Marketability Discounts: Developing and Supporting Marketability Discounts in the Appraisal of Closely Held Business Interests, Peabody Publishing, LP, 1997.

<sup>5</sup> This technique has come into prominence more recently. See, for example, "Executive Relief," *The Economist*, April 3, 1999. For an example of popular press on hedging, see "Few Options on Cheney Option Problem," *The New York Times*, August 26, 2000.

which protects her against any decline in the stock's price, and simultaneously writing a call in order to obtain additional funds. As the name implies, a zero-cost collar is structured so that the proceeds from writing the call exactly offset the costs of purchasing the put. If the strike prices of the call and the put are identical, the investor will completely hedge away any price risk from the underlying security.<sup>6</sup>

Whenever these methods are cheaper than selling the restricted security, we would expect that holders of restricted securities would take advantage of their existence. As a result, observed sale discounts would provide a downwardly biased measure of the inherent costs of a restriction if hedging were unavailable.<sup>7</sup>

Most of the marketability discount studies generally present only medians or average results, with perhaps some qualitative discussion of the types of restrictions that can affect the size of the discount. At least two studies do present statistical analyses of the size of the marketability discount, however. Silber<sup>8</sup> presents a regression, using data from transactions between 1981 and 1988, that quantifies some of the determinants of the difference between the prices of restricted and freely-traded stock. Specifically, Silber explains the log discount as a function of the log of the underlying firm's revenues, the log of the size of the restricted block relative to the total common stock, a dummy for whether the firm has positive earnings, and a dummy for whether "there is a customer relationship between the investor and the firm issuing the restricted stock." Silber reports that all the coefficients on the variables listed above are significant at the 10% level or better<sup>9</sup>, and the R<sup>2</sup> of his regression is 0.29.

---

<sup>6</sup> One of the interesting features of the collar, however, is that the strike price of the call can be set above that of the put, leaving a limited range in which the owner is still exposed to gains or losses due to movements in the price of the underlying security. The owner can then claim to still be affected by price risk and thus still have an ownership stake in the underlying security. To the extent that this is true, the collar would then not qualify as a constructive sale of the underlying security, and the investor would not have to pay any capital gains taxes at the point when the collar is established.

<sup>7</sup> For example, suppose that the cost of hedging a security was always 30% of its undiscounted value. Then in cases where the discount for bearing the risk was more than 30%, a hedgeable security could still be sold for 70% of its undiscounted value, as the purchaser would not need to actually bear the risk. In this scenario, only sale discounts of 30% or less would be observed, thereby understating the average underlying discount that would be observed in the absence of hedging.

<sup>8</sup> William L. Silber, "Discounts on Restrict Stock: The Impact of Illiquidity on Stock Prices," *Financial Analysts Journal*, July-August 1991.

<sup>9</sup> Silber states that all of the variables are significant at the 5% level using a one-tailed test, however the implied t-statistic on the earnings dummy is 1.61, which is not statistically significant at that level.

Trout<sup>10</sup> analyzes the discounts over the period 1968 to 1972 as a function of an exchange listing indicator, number of shares outstanding, percentage control involved in the transaction, an indicator for transactions of under 1% of the shares outstanding, and the value of the purchase (calculated at the undiscounted stock price).<sup>11</sup> All of the coefficients on these variables were significant at the 5% level or better, and the adjusted  $R^2$  of the regression was 0.24.

While these papers represent a step forward, it is not clear how to apply these older studies to analyses today or for other points in time. This is especially true because some of the variables such as revenue or the dollar value of a purchase will surely have different coefficients because of the changes in nominal price levels and the average sizes of many real economic variables.

## **B. Theoretical Solutions**

To my knowledge, there are three published theoretical methods that try to model the illiquidity discount.

### **1. Chaffee Put Analysis, or Accounting for the Possibility of Loss**

David B. H. Chaffee III argues that a European put is a good measure of the “risk caused by the inability to quickly and efficiently return the investment to a cash position.”<sup>12,13</sup> The basic thrust of Chaffee’s analysis is that by purchasing a put that allows her to sell at the current (or forward) price, an investor can be guaranteed that the illiquid investment will not result in a loss. Thus, the downside risk of illiquidity is eliminated. The put’s premium is then the value Chaffee assigns to the cost of illiquidity.

---

<sup>10</sup> Robert R. Trout, “Estimation of the Discount Associated with the Transfer of Restricted Securities”, *Taxes*, June 1977, pp. 381-385.

<sup>11</sup> Note also that both the Silber and Trout studies include variables that may pick up the effects of a minority discount or control premium. Therefore, these studies can be valuable in accounting for two separate adjustments at the same time.

<sup>12</sup> David B. H. Chaffee III, “Option Pricing as a Proxy for Discount for Lack of Marketability in Private Company Valuations,” *Business Valuation Review*, December 1993.

<sup>13</sup> It is not clear why one would use European puts, except for computational ease. An American put would more accurately measure the lack of ability to sell the asset at *any* point before a restriction expired.

One concern with Chaffee's theory is that by eliminating the downside risk but still leaving the investor with the potential for upside gains, it overstates the cost of risk that the owner of the restricted asset faces. For example, if an investor simply held restricted stock, she would benefit from any appreciation of the stock's value and suffer a loss when the stock declined in price. On the other hand, the owner of the combination of a restricted stock and a European put on that stock will, at the time the put expires, have benefited from an appreciation of the stock's value, but not have suffered a loss if the stock's price declined. Thus, the owner of the restricted stock and a put can be said to be in a better financial position than the owner of a similar unrestricted stock.<sup>14</sup> Consequently, the cost of the put is a potential upper limit on the actual costs of the restriction. As such, it is a good advance over not having any theory at all, but is likely to lead to upwardly biased estimates of the cost of illiquidity.

## **2. Longstaff Upper Bound Analysis, Or Accounting for the Possibility of Missing Out on Upward Gains**

Francis A. Longstaff has provided an alternative theoretical analysis of the maximum cost of non-marketability that is in some ways the converse of the Chaffee analysis.<sup>15</sup> Longstaff's technique focuses on the dynamics of a security's price during the period in which an investor holds a non-marketable share of that security. Longstaff essentially measures the difference between being able to sell the security at the expected optimal point during the period of illiquidity (accounting for reinvestment of the proceeds at the risk-free rate) and being forced to hold the security to the end of the period of illiquidity. This difference represents the maximum expected gain that an investor with perfect foresight would have if she were able to perfectly time her sale of the security. Longstaff defines that difference as the maximum cost of a lack of marketability.<sup>16</sup> It is an upper bound, or maximum value, because in the real world, even when an investor has

---

<sup>14</sup> This also does not account for the lack of access to capital while the restriction is in effect. In fact, because the holder of the stock would require additional capital to purchase a put at the moment she wished to hedge her risk, this theory compounds any error involved in ignoring this aspect of the problem.

<sup>15</sup> Francis A. Longstaff, "How Much Can Marketability Affect Security Values?" *Journal of Finance*, December 1995.

<sup>16</sup> Note that Longstaff also focuses on the price risk and not on the lack of access to capital.



a liquid asset, she cannot be sure that she will sell it at the optimal point.<sup>17</sup> Thus, while Chaffee places a bound on the cost of avoiding a loss, Longstaff measures the maximum expected gain that would occur if the illiquidity were removed.

By valuing this potential gain as a look-back option from the point when the asset again becomes liquid, Longstaff is able to quantify the maximum discount. Longstaff notes that for stocks with a standard deviation of returns in the range of 25% to 35%, his theoretical estimates of the upper bound are close to the 35% to 40% discounts cited by Pratt and the 34% discount cited by Silber. He therefore concludes that his estimate of the upper bound may be a reasonable approximation of the actual illiquidity discount.

### **3. Mercer Required Rate of Return Analysis**

Z. Christopher Mercer has developed what he terms the “Quantitative Marketability Discount Model” for valuing closely-held business interests. This model can apply, however, to either a business or a security. The basic outline of the Mercer model is as follows. First, an analyst estimates the current value of a firm using conventional means without reference to lack of marketability. The analyst then estimates the length of time that the firm will remain in a non-marketable condition. Let’s say that this will be for ten years. The analyst then grows the present value of the firm at an assumed “expected growth rate” for ten years. He then discounts that value back at the investor’s assumed “required holding period return.” This results in an assessment of the value of the firm with a marketability discount.

This formulation provides a sound theoretical means for explaining the current value of an asset or business. However, the analyst must still estimate some of the inputs to the model. First, it is not clear what the expected growth rate of a firm will be. Mercer suggests a careful analysis of the firm under study. Yet, this replaces the difficulty of determining the marketability discount with the challenge of determining an expected growth rate. This is of course an improvement in that if the firm has been operating for some time, there will be evidence of past rates of growth, whereas there may not be evidence of past discounts for selling partial interests in the firm.

---

<sup>17</sup> In fact, if markets are efficient, the stock is as likely to rise or fall after the investor sells it, and the lack of marketability should have no effect whatsoever on her risk-adjusted expected rate of return.

Second, the analyst must also determine the purchaser's expected holding period. As discussed below, this is likely to always be a difficult point in the determination of any marketability discount, and Mercer provides a useful list of factors that may affect this period.<sup>18</sup>

The final factor that the analyst needs to determine is the required holding period return. While Mercer presents an interesting theory purportedly based on the CAPM, as the last step, he adds in "specific risk premiums" that, according to Mercer, come from factors such as "appraiser judgment, appraiser experience, common sense, reasonableness, [and] comparisons with other relevant rates of return or market indicators." The first four of these factors are simply different ways of saying that the so-called specific risk premiums are subjective. The last, the comparison to other relevant rates of return or market indicators, is an interesting choice, because if the appraiser had a lot of information about sales of similar non-marketable assets, then he probably would not need to go through the Mercer analysis.

As discussed in the following section, the primary advancement of this paper is to combine the expected growth rate of the asset and the investor's required rate of return into an excess required rate of return for the investor and to provide a method of quantifying that excess required rate of return in an objective manner. The basis for this will be an extension of the CAPM model to measure the required return for illiquid assets. As with the Mercer study, the length of time that this excess rate of return will apply will still be difficult to determine; however, rather than having to estimate two different rates of return (the expected growth rate and the required holding period return), the methodology discussed below allows an analyst to calculate the combined effect of these two, and to do so without the use of subjective estimates.

---

<sup>18</sup> See Mercer, pp. 242-244.

### **III. A Theory Based on the CAPM and the Equity Risk Premium**

In this section, I propose a theory to measure the costs of illiquidity. As discussed above, the theory focuses on the extra risks imposed on the owner of a security or interest in a business enterprise, and not on the lack of access to capital. In brief, the theory uses market data on the additional return that investors require in order to hold a risky asset, measured by the equity risk premium, to extrapolate the extra return that the holder of an illiquid asset would require. The theory measures the extra required rate of return as if the illiquid asset were the sole component in the typical investor's portfolio. This may result in an overestimate of the illiquidity discount because it ignores the effects of possible diversification. This topic is discussed briefly in a later section.

#### **A. What Does the Equity Risk Premium Measure?**

The equity risk premium measures the additional expected return that an investor gets for holding a risky asset. In theory, it measures the expected return on equity less the expected return to a risk-free asset. In practice, it is usually measured as the historic return on a broad market index less the return to Treasury bonds over the same period.

According to the CAPM model, an investor is only rewarded for bearing systematic risk, while unsystematic risk does not result in any extra expected return above the risk-free rate. To measure the expected return to equity, one measures the covariance (or the beta, " $\beta$ ") between a stock and a market index and then adds beta times the equity risk premium to the risk-free rate. The lack of an extra return for nonsystematic, or idiosyncratic, risk comes from the fact that the typical investor can choose to hold only a small share of her wealth in any one asset, and thereby create a diversified portfolio. As a result, the nonsystematic risk in any individual asset contributes minimally to the variance of the overall portfolio. If the investor could earn an extra return for holding assets with nonsystematic risks, then she could hold a diversified portfolio of such assets in which those nonsystematic risks would mostly cancel each other out, and she would earn a rate of return in excess of the risk-free rate.

One advantage of using a theory based on the equity risk premium is that it can be applied across different time periods with different risk premiums. It can also be trivially shown to work in a theoretical world where all participants are risk-neutral. In that case, there would be no equity risk premium and thus no illiquidity discount under this model. This contrasts with the Chaffee or Longstaff analyses that both would still correctly provide a positive upper bound to the illiquidity discount, but would fail to recognize that that upper bound could be lowered to zero in a risk-neutral world.

## **B. A Theory of Applying the Equity Risk Premium to An Illiquid Security**

In order to apply the equity risk premium to the case of a holder of an illiquid security, I first define a number of terms as follows:

Let  $r$  represent the expected return to an asset. Specifically, let  $r_M$  represent the expected return to the market portfolio, let  $r_i$  represent the expected return to asset  $i$ , and let  $r_{RF}$  represent the (expected) return to the risk free asset.

Next, define  $\theta$  to be the equity risk premium, or  $r_M - r_{RF}$ . This represents the excess return of the market as a whole over the risk-free rate of return.

Define  $\beta$  to be the covariance of the security with respect to the market portfolio.

Let  $\sigma$  represent the volatility (standard deviation) of some series of returns. Specifically, let  $\sigma_M$  represent the standard deviation of the returns of the market portfolio,  $\sigma_i$  be standard deviation of the returns of asset  $i$ , and  $\sigma_e$  as the volatility of the undiversifiable risk of the asset in question, even if it were freely tradable. I.e.,  $\sigma_e$  is the volatility of  $r_i - (\alpha + \beta r_M)$ <sup>19</sup>, or that part of  $r_i$  that is not explained by market movements.

(Note that  $\beta$  equals  $\sigma_{iM}/\sigma_M^2$ , where  $\sigma_{iM}$  is the covariance between asset  $i$  and the market.)

Finally, define  $T$  to be the time that the asset is restricted.

---

<sup>19</sup>  $\alpha$  and  $\beta$  are generally determined by a regression of the returns of the stock on a market index.

To begin, assume that these quantities are all measurable, except for  $\sigma_\epsilon$ , which is calculated from other variables. As discussed below, this theory will still require a somewhat subjective analysis if one or more of these quantities, typically  $T$ , the time of the restriction, must be estimated based on qualitative data.

The theory is developed as follows. First, we account for the fact, discussed above, that some assets can be hedged. To do so, we first apply the following rule: *If the asset can be hedged, the illiquidity discount is the minimum of the costs of hedging and the direct costs of liquidity as measured below, while if the asset cannot be hedged, the illiquidity discount can be measured as discussed below.* This rule accounts for the fact that an investor would not receive an illiquidity discount that would exceed the costs of removing the risks associated with the illiquidity.

Assuming that the asset cannot be hedged, we begin by noting that, following the CAPM, the normal market risk premium for the asset in question is  $\beta\theta$ , or  $\theta\sigma_{iM}/\sigma_M^2$ , representing the covariance of the asset's returns multiplied by the equity risk premium. This represents the contribution that the asset makes to the variance of the average, or market, portfolio. This term is normally added to the risk-free rate in order to determine the required return to equity for an asset. It represents the extra expected return to holding the market portfolio as a result of the extra (undiversifiable) risk that holding the market portfolio entails.

As discussed above, to find the extra expected return to a restricted asset, we begin with the assumption that the restricted asset cannot be broken up and will represent the entire portfolio of the purchaser. Under this assumption, the restricted asset's contribution to the variance of the investor's portfolio equals the variance of that asset itself. Because the required rate of return depends on the variance of an investor's portfolio, the necessary risk premium is then proportional to the variance of the restricted asset and not just to its covariance with the market portfolio (representing its contribution to overall market risk in cases where all assets can theoretically be broken up into infinitesimally small pieces and shared equally by all investors).

Therefore, rather than bearing the undiversifiable risk,  $\sigma_{iM}$ , or the covariance of the asset's returns with that of the market, the holder of a restricted asset bears the entire

risk of that asset or the asset's covariance with itself,  $\sigma_{ii}$ , or  $\sigma_i^2$ . Consequently, her required return above the risk-free rate is no longer  $\theta\sigma_{iM}/\sigma_M^2$ , but instead is  $\theta\sigma_i^2/\sigma_M^2$ .

Thus, the expected/required rate of return above the risk-free rate is  $\theta$  the equity risk premium, scaled up by  $\sigma_i^2/\sigma_M^2$ , or the ratio between the risk born by holding the individual security and the risk borne by holding the market portfolio. This overall risk,  $(\sigma_i^2/\sigma_M^2)\theta$ , can be rewritten to break up  $\sigma_i$  into the systematic and unsystematic components of the asset's volatility. If we do so, we note that  $(\sigma_i^2/\sigma_M^2)\theta$  can be written as  $\theta(\beta\sigma_M^2 + \sigma_\epsilon^2)/\sigma_M^2$ , which in turn can be written as  $\theta(\beta + \sigma_\epsilon^2/\sigma_M^2)$ .<sup>20</sup> Because  $\sigma_\epsilon^2/\sigma_M^2$  is necessarily positive, the entire return is necessarily greater than or equal to  $\beta\theta$ .

For ease of notation, we now define  $\phi$  to be  $\sigma_\epsilon^2/\sigma_M^2$ , the excess expected return over the risk free rate per percentage point of equity risk premium for a restricted stock.<sup>21</sup> Note that  $\phi$  equals  $(\sigma_i^2/\sigma_M^2) - \beta$ .  $\phi$  then equals the ratio of the excess variance of the asset being measured over the market portfolio to the variance inherent in the market portfolio, less that portion of the excess variance that is correlated with the market portfolio and therefore already incorporated in the value of the unrestricted asset. As per the previous paragraph,  $\phi$  will equal zero when the asset is identical to the market portfolio and will otherwise be positive.  $\phi$  will then represent the extra discount for risk *above the discount present simply for correlation with the market portfolio*.

For valuation purposes, we can then define the illiquidity discount per unit of time as  $\phi\theta$ . Consider now an asset that will have a value of 1 after time T. Because of the illiquidity discount, the value of that asset today is only  $e^{-\phi\theta T}$  times the value it would otherwise have. Therefore, the illiquidity discount is  $1 - e^{-\phi\theta T}$ .<sup>22,23</sup>

<sup>20</sup> Because variances are additive  $\sigma_i^2 = \beta\sigma_M^2 + \sigma_\epsilon^2 + \text{cov}(\sigma_M, \sigma_\epsilon)$ , where the covariance term is zero by construction.

<sup>21</sup> This results in a similar analysis to that of Meulbroek, Lisa K., "The Efficiency of Equity-Linked Compensation: Understanding the Full Cost of Awarding Executive Stock Options," Harvard Business School working paper, 2000. Meulbroek uses the Sharpe ratio to determine the upper bound of the ratio of required rate of return for increased volatility.

<sup>22</sup> Note that a restricted investment in the market portfolio would receive no discount for lack of liquidity. This is because the investor could simply short a market index at any point that she wished to sell the restricted asset. This would remove all of the price risk of holding the restricted investment until the restriction expired.

<sup>23</sup> This may be derived as follows: Under the CAPM, the asset price equals its expected value at time T times the exponential of  $-(r_{RF} + \beta\theta)T$ , while under the analysis of a restricted security, its value is the expected value at time T times the exponential of  $-(r_{RF} + \sigma_i^2/\sigma_M^2\theta)T$ . The percentage discount is therefore

To determine the reasonableness of this formula, we compare the results obtained by calculating the discount to a broad array of stocks. Specifically, we examine the five hundred companies that were members of the S&P 500 index as of December 1999. For each of these companies, we calculate its historical volatility over the five calendar years 1989 through 1993 and 1994 through 1998, dropping any years in which the company was not publicly traded for the full year. Similarly, we calculate the volatility of the S&P 500 index over the same time periods. The equity risk premium is calculated either beginning in 1926, the first year for which Ibbotson presents data, or from 1946, to show the results using post-war data only. The differences in these two measures are not significant. Finally, we account for the dividend yield of the S&P 500 Index. Because these dividends will be received before the stock will be marketable, the cash flows corresponding to the dividends will be illiquid for a lesser period of time, as discussed below. These inputs to the discount formula are shown in Table 1.

<b>Table 1</b>					
<b>Year</b>	<b>Average Volatility of S&amp;P 500 Member Company<sup>1</sup></b>	<b>Volatility of S&amp;P 500</b>	<b>S&amp;P 500 Dividend Yield</b>	<b>Intermediate-Term Equity Risk Premium<sup>2</sup></b>	
				<b>1926-</b>	<b>1946-</b>
1989-1993	30.4%	12.7%	3.2%	7.7%	7.2%
1994-1998	31.4%	14.5%	2.2%	7.9%	7.6%
<sup>1</sup> S&P 500 member company list as of December 15, 1999. These companies are treated as representative of previous S&P 500 members.					
<sup>2</sup> Average risk premium for the five years in each period, using data from Ibbotson Associates' <u>Stocks, Bonds, Bills and Inflation Yearbook</u> .					

To calculate the implied illiquidity discounts for the “typical” company, we assume that  $\beta$  equals one. Under this assumption, we find that  $\phi$  equals 4.75 for the 1989-1993 period and 3.72 for the 1994-1998 period. We then calculate the implied discount for a two-year restriction (corresponding to the required holding period under SEC Rule 144 for restricted stock) and an equity risk premium equal to the average

---

$\{\exp[-(r_{RF} + \beta\theta)T] - \exp[-(r_{RF} + \sigma_i^2/\sigma_M^2\theta)T]\} / \exp[-(r_{RF} + \beta\theta)T]$ . This simplifies to  $1 - \exp[-(\sigma_i^2/\sigma_M^2 - \beta)\theta T]$ , which in turn equals  $1 - e^{-\phi\theta T}$ .

excess return of stocks over intermediate-term bonds, as reported by Ibbotson Associates. In addition, we calculate the implied discount for a one-year restriction, corresponding to the average time until dividends are received. These two discounts are averaged so that two years' worth of dividends are effectively assumed to be received one year into the holding period, and the stock itself has its illiquidity period end after the two years. The results are presented in Table 2.

If we calculate the average volatility of an S&P 500 member stock, we see that the implied discounts are approximately 49% in the earlier time period and approximately 41% in the later period. Breaking up the data into percentiles by volatility of the S&P 500, from low to high, reveals some interesting patterns. First, we see that even though the average  $\phi$  in each period is higher than the  $\phi$  based on the average volatility, the average of the discounts is less than the discount of the “representative” stock. This is a result of Jensen's Inequality and the nonlinearity in the discount formula. It further suggests that were we to break up the percentiles more finely, or to replace the choice of a constant  $\beta$  of one across each stock with individual values for  $\beta$ , the average illiquidity discount would fall further.



Table 2					
Volatility	$\phi$	Implied Illiquidity Discount		Longstaff Theoretical Upper Bound	Chaffee Put Analysis
		Based on Different Equity Risk Premiums			
		1926-	1946-		
1989-1993					
Mean	4.75	50.4%	48.4%	38.3%	32.6%
10th percentile	1.54	20.4%	19.4%	24.4%	22.1%
20th percentile	2.13	27.1%	25.8%	27.3%	24.4%
30th percentile	2.71	33.0%	31.5%	30.0%	26.5%
40th percentile	3.23	38.0%	36.3%	32.3%	28.2%
50th percentile	3.92	44.0%	42.1%	35.1%	30.3%
60th percentile	4.59	49.2%	47.2%	37.7%	32.2%
70th percentile	5.54	55.8%	53.7%	41.2%	34.7%
80th percentile	7.72	67.9%	65.7%	48.6%	39.6%
90th percentile	12.09	82.9%	81.1%	61.8%	47.5%
Percentile Average	4.83	46.5%	44.8%	37.6%	31.7%
1994-1998					
Mean	3.72	42.3%	40.5%	39.8%	33.7%
10th percentile	1.20	16.3%	15.4%	26.0%	23.4%
20th percentile	1.67	22.0%	20.9%	29.0%	25.7%
30th percentile	1.98	25.5%	24.2%	30.8%	27.1%
40th percentile	2.48	30.8%	29.3%	33.6%	29.2%
50th percentile	2.91	35.0%	33.4%	35.8%	30.8%
60th percentile	3.54	40.7%	38.9%	38.9%	33.1%
70th percentile	4.32	47.2%	45.2%	42.6%	35.6%
80th percentile	5.84	57.7%	55.6%	49.3%	40.0%
90th percentile	10.69	79.0%	77.1%	67.6%	50.6%
Percentile Average	3.85	39.4%	37.8%	39.3%	32.8%

Next, we note that for low volatility stocks, the illiquidity discount based on the equity risk premiums tends to be below the Longstaff or Chaffee bounds, while for more volatile stocks, the reverse is true.<sup>24</sup> This would indicate the need for caution in applying this analysis to highly volatile stocks, because the results could exceed the theoretical maximums suggested by other papers. Perhaps fortunately, in the case where a stock is

<sup>24</sup> The Longstaff and Chaffee analyses are also adjusted to account for the interim dividend yield in the same manner as described above.

highly volatile, it is also more likely that the direct costs of bearing risk will outweigh the brokerage costs of hedging risk, and therefore it is less likely that the risk would have to be borne.

Finally, we note that the discounts shown in the table are at the upper end of the generally presented empirical estimates. As discussed above, observed discounts are probably a downwardly biased estimate of underlying risks because of the possibility of hedging, while the model presented in this paper may be upwardly biased since it assumes that the restricted asset comprises the entirety of the investor's portfolio.

### **C. A Theory of Applying the Equity Risk Premium to a Discounted Cash Flow Valuation of a Firm**

If a project or company is not traded, it is often valued using a discounted cash flow model. This is generally done by projecting cash flows and discounting those cash flows back at the calculated weighted average cost of capital or WACC. Often an illiquidity discount is then applied to the final valuation.

This final haircut for illiquidity is generally quite subjective. It may be based on an analyst's intuition or else on some average discount from the literature. It may also be based on the methodologies discussed above. As discussed above, none of these methodologies explicitly accounts for the volatility of the company being valued.

Based on the model presented in this paper, another proposed solution is to adjust the WACC by adding in an illiquidity premium as discussed above. Specifically, one would use the same cash flows as before and, for each year in which the company is expected to be illiquid, to adjust the WACC by adding the extra return for illiquidity. That is, the new WACC should equal the old WACC plus  $(\sigma_i^2/\sigma_M^2 - \beta)\theta$ , or  $\phi\theta$ .

Table 3, below, presents an example of how this analysis would be applied to the valuation of a firm. For this example, we assume that a firm has cash flows of 100, coming due one year from the present, and that its cash flows are then expected to grow by 10% per year. The firm is financed purely by equity. To determine the cost of capital, we assume that the risk-free rate is 7%, the firm's  $\beta$  is 1, the equity risk premium is 8%, and the ratio  $\sigma_i^2/\sigma_M^2$  equals 3. Under these assumptions, the firm's standard cost of capital, without the inclusion of any illiquidity effects, is 15%. If one explicitly accounts

for illiquidity in the determination of the required rate of return, one comes up with a value of 31%. Table 3 then shows how the firm is valued under these assumptions, both with no discount for illiquidity, and under various illiquidity assumptions.

<b>Table 3</b>			
Year	Cash Flow	Present Value	
		WACC	Full Discount
1	100.0	87.0	76.3
2	110.0	83.2	64.1
3	121.0	79.6	53.8
4	133.1	76.1	45.2
5	146.4	72.8	38.0
Terminal Value	3,221.0	1392.5	151.7
Sum		1791.1	429.1

  

	Discount	Implied Value
	20%	1432.9
Explicit	25%	1343.3
Illiquidity	30%	1253.8
Discount	35%	1164.2
	40%	1074.7

  

Years Illiquid	Value	Implied Discount
1	1572.4	12.2%
2	1389.6	22.4%
3	1237.1	30.9%
4	1109.7	38.0%
5	1003.4	44.0%
forever	429.1	76.0%

In the table, the first set of values uses different potential explicit illiquidity discounts applied directly to the figure of 1,791.1 calculated without any consideration of marketability. The discounts range from 20 to 40 percent, which is generally typical of the range observed in business valuations. The second set of valuations is based on different estimates of when the investment will become liquid. As can be seen, there is a fairly large spread between the lowest and highest values. This highlights the importance of an accurate determination of the likely period of illiquidity, either through the review of business plans or a study of comparable situations.

One important question is how to estimate  $\sigma_i$  for the company. If the company's shares are publicly traded, this is a straightforward exercise. Even if not, this is not really a new question, as analysts have learned to estimate  $\beta$  based on a selection of peer groups. Similarly, one can use the peer group methodology to estimate the company's unobserved volatility.<sup>25</sup> If an analyst wanted to go beyond simple peer group averages, he could use regression analysis to see how stock price volatility for the peer group members depends on financial characteristics such as the volatility of earnings or revenues. This relationship could then be used to estimate what the stock price volatility of the company under study would be if it were publicly traded. Once this is completed, the valuation would proceed as a standard discounted cash flow analysis, with the illiquidity-adjusted WACC used as the discount rate.

#### **D. Determining the Likely Length of the Period of Illiquidity**

While all of the previous analyses could be done in a completely objective manner, there still remains one area where such objectivity may or may not be possible. This is in the determination of the length of time for which the asset will remain illiquid.

In some cases, this determination may be completely straightforward. For example, an investor may receive stock that she cannot sell for a certain period, after which she can sell the stock freely. Clearly, in that case the length of the restriction is the amount of time before the stock can be sold.

A more complicated case would be if the stock could only be sold in small quantities, for example under SEC Rule 144. In that case, one could approximate the length of the restriction by considering the amount of time until half the stock could be sold, or could undertake a more complicated analysis of valuing different batches of the stock as if they had different restrictive time periods associated with them. But, even that type of analysis may require one to make some estimates of future market volume in order to estimate how quickly the stock could be sold under the Rule 144 limitations.

---

<sup>25</sup> An analyst should be careful to note that the volatility from the equity of peer group members is akin to a levered beta, in that it applies only to the equity component of the capital structure. The excess volatility has to be adjusted to reflect the volatility of the company as a whole. This can be done most easily by making the same assumption as with levered and unlevered betas: that debt prices do not covary with the market.

An even more difficult case would occur if the length of the restrictions depends on future events. For example, a private company could go public or be bought out in a transaction in which the restricted stock is exchanged for (possibly restricted) stock of an acquirer. In these cases, the analyst will have to carefully consider the potential future events in determining the length of the restriction.

In that regard, it may seem that at some level the entire analysis proposed above again depends on a subjective view of the strength of some restriction on marketability. To some extent that is correct. However, there will be cases when the length of time is clearly known, and the subjectivity disappears. In other cases, there may be evidence from peer groups, for example, on how long it takes, on average, until a private company in a certain industry or with certain financial characteristics goes public or is acquired. At the very least, the subjectivity is moved from the amorphous question of whether certain liquidity restrictions are important to the question of how those restrictions affect the likely length of time until the asset can be sold. Such information is more likely to be found in a business plan than is the cost of illiquidity. Interestingly, this change of question may be particularly important in cases where one is undertaking a valuation with the benefit of hindsight. That is, the time until a firm goes public or is acquired may be known in retrospect; to the extent that this is a good proxy for the beliefs at the time of the valuation, it may help answer the question of the length of time that an interest in the firm would be non-marketable.

## **E. Additional Considerations for the Model**

For those cases where the lack of access to capital is a consideration, one could add in the additional costs of borrowing as an additional difference between restricted and unrestricted assets. Note that this assumes that the party would be willing to sell the unrestricted asset. I do not address the means of assessing such costs in this paper.

A second consideration is whether there is a different discount if the illiquid asset is to be part of a larger portfolio or was going to be held by an investor who was not planning to sell the asset even if it were not restricted. In these cases, the illiquidity of the asset would seem to matter less and the discount discussed above would appear to be

an overestimate of the actual discount. In some circumstances, this would no doubt be the case. The analysis presented above was based on the assumption that the illiquid asset formed the investor's entire portfolio and therefore its variance equaled its contribution to the variance of the portfolio. In other cases, the generalized approach would be to determine the variance of an investor's portfolio with the asset under consideration as restricted, compared to what that variance would be if the asset were not restricted and the investor were free to diversify. To the extent that the restriction causes the investor to hold a riskier portfolio, she would need to be compensated for this risk with an extra expected rate of return, which would translate into a discount for the purchase of the restricted asset. The calculation of the amount that the restricted asset adds to a portfolio's variance would be determined by comparing the portfolio's estimated variance with the restricted asset to the estimated variance of the portfolio the investor would have maintained had the asset not been restricted.

## **F. Conclusion**

This equity-risk premium based approach for calculating illiquidity discounts has a number of benefits over using simple average discounts or over other methodologies discussed above. These are summarized as follows:

- Because illiquidity discounts are a rate over time, there will be different discounts depending on the period of time for which the asset is illiquid so that a short-term restriction would create less of a discount than would a long-term discount, holding everything else constant. This is also true in the Mercer, Chaffee, and Longstaff methodologies, but not when using average discounts or the Silber or Trout equations.
- By using different equity risk premia for different time periods, the illiquidity discount can likewise change over time, which is realistic and allows one to estimate these discounts for periods for which there has been no empirical study.
- The discount is related to the volatility or riskiness of the asset whose price risk must be borne.
- The discount ties into the literature on other market risks and returns.
- The calculation of an illiquidity discount is objective (or at least relatively so) because it can be calculated based on volatility (actual for a security, from a peer group for a company) and the equity risk premium.

In sum, the theoretical model of valuing restrictions on liquidity described in this paper relates the economic literature on equity risk premiums in general to the extra risk that the holder of the restricted security must bear. The result is a framework for measuring illiquidity discounts that vary over time, with the length of the restriction, and with the riskiness of the illiquid asset. Perhaps most importantly, the analysis is less subjective than the type of analysis often used in practice today.