© *The Journal of Risk and Insurance*, 2003, Vol. 70, No. 3, 509-526

ASSET ALLOCATION AND THE LIQUIDITY PREMIUM FOR ILLIQUID ANNUITIES

S. Browne
M. A. Milevsky
T. S. Salisbury

ABSTRACT

Academics and practitioners alike have developed numerous techniques for benchmarking investment returns to properly adjust seemingly high numbers for excessive levels of risk. The same, however, cannot be said for liquidity, or the lack thereof. This article develops a model for analyzing the *ex ante* liquidity premium demanded by the holder of an illiquid annuity. The annuity is an insurance product that is akin to a pension savings account with both an accumulation and decumulation phase. We compute the yield (spread) needed to compensate for the utility welfare loss, which is induced by the inability to rebalance and maintain an optimal portfolio when holding an annuity. Our analysis goes beyond the current literature, by focusing on the interaction between time horizon (both deterministic and stochastic), risk aversion, and preexisting portfolio holdings. More specifically, we derive a negative relationship between a greater level of individual risk aversion and the demanded liquidity premium. We also confirm that, *ceteris paribus*, the required liquidity premium is an increasing function of the holding period restriction, the subjective return from the market, and is quite sensitive to the individual's endowed (preexisting) portfolio.

> "*. . . If the insurance company has greater and longer surrender charges, then it can pay more, on fixed annuities, knowing the funds aren't going to leave, so the liability structure will be more stable . . .*" (*Best's Review*, October 2001, p. 43)

INTRODUCTION AND MOTIVATION

In the United States, the term annuity covers a wide spectrum of financial and insurance products. A *savings* (payin) annuity is akin to a bank account or savings bond

---

S. Browne is at the Columbia School of Business and M. A. Milevsky and T. S. Salisbury are at York University. This research is partially supported by a grant from the Mathematics of Information Technology and Complex Systems (MITACS) as well as The IFID Centre at the Fields Institute. Salisbury acknowledges support from NSERC. The authors would like to acknowledge helpful comments from M. Cao, P. Chen, G. Daily, J. Francis, A. Gottesman, G. Jacoby, S. Posner, K. SigRist and two anonymous *JRI* reviewers. A previous version of this article circulated under the title of: *A Note on the Solid Value of Liquid Utility*.

where money is accumulated over a period of time at a variable or fixed rate of interest. In contrast, a *consumption* (payout) annuity is similar to a pension that pays a periodic fixed or variable amount, which might also contain longevity insurance. The former is usually used prior to retirement, whereas the latter is used during the retirement years. According to the above-quoted A. M. Best survey, more than \$1 trillion are currently invested in various types of annuity products in the United States.

The common denominator of *fixed* (in contrast to variable) *savings* and *consumption* annuities is that they are quite illiquid. Namely, in stark contrast to a money market fund or savings bond that can be redeemed on a daily basis without any penalty, it is difficult or very costly to surrender (or cash-in) a fixed annuity. Whereas the reasons for this illiquidity differ depending on whether the product is in the accumulation or decumulation phase, the fact remains that continuous asset reallocation is virtually impossible with these products. Our article therefore asks a simple question: *What is the liquidity premium that a rational investor will demand to compensate for the inability to maintain a properly rebalanced and diversified portfolio?*

Thus, for example, the holder of a fixed (savings) annuity might be told that he or she cannot withdraw from (or cash out of) the product for the first 7 years of the contract. Or, in the event of a permissible early withdrawal during first 7 years, one might be "hit" with an x percent penalty, commonly referred to as market value adjustment. In the decumulation phase, illiquidity is often accepted as the cost of obtaining longevity insurance, whereas in the accumulation phase, *policyholders are told* that their return will exceed the yield of a comparably liquid instrument. Indeed, it is quite common to see a monotonic relationship between the magnitude of the early surrender charges on a fixed annuity—controlling for commissions—and the guaranteed yield *if* the product is held to maturity. Implicitly, investors (or more precisely, policyholders) are being promised compensation for the liquidity restrictions.

Recent academic literature has documented the empirical welfare gains from annuity products and annuitization, as well as the value of longevity insurance. For example, Mitchell et al. (1999) argued that consumers would be willing to "give up" 30 percent of their wealth to obtain a fairly priced annuity. Similarly, Brown and Poterba (2000) explained the extremely low levels of annuitization by arguing that married couples function as a miniannuity market. Blake and Burrows (2001) focused on the undesirable longevity risk taken by insurance companies issuing payout annuities, and the need for governments to issue mortality-linked bonds.

However, most of the literature discussing the costs and benefits of annuitization has ignored some of the problems created by having a portfolio that cannot be liquidated or rebalanced for long periods of time.

There are two lines of reasoning on the topic of liquidity compensation. Some financial economists, such as Longstaff (1995, 2001), argue that one should be compensated in equilibrium for illiquidity restrictions. In other words, all else being equal, a fixed income instrument that cannot be sold—or, for that matter, subsequently repurchased—over the life of the product, should provide investors with a higher yield. An alternative line of reasoning is that in equilibrium, investors should not be compensated for illiquidity, because they can lengthen their trading horizon when faced with such securities. In other words, they can use illiquid instruments to fund long-term liabilities, without demanding any compensation for this inconvenience. We refer the interested reader to Vayanos and Vila (1999) for a model that pursues this particular approach. In this article we pursue the former approach, although we discuss conditions in which the latter might apply.

Note, of course, that from an insurance company's perspective, liquidity restrictions are absolutely necessary to manage the duration mismatch (or risks) that otherwise would arise if incoming funds are invested in long-term projects, but yet instantaneously available to policyholders. Insurance companies must protect themselves by imposing a disintermediation (or market value adjustment) surrender charge. Therefore, from the perspective of the vendor of such products, our model should help determine the appropriate level of restrictions vis-á-vis the promised yield. Indeed, one of our results is that individuals that have very little "outside" wealth invested in the financial markets will demand a higher liquidity premium on the restricted annuities (e.g., within a pension plan) since they cannot offset the rebalancing risk with other investments.

This so-called liquidity premium cannot, of course, be determined in isolation; its value will depend on the alternative investments available, and the investor's willingness to make use of them. Thus, we will work in a framework in which there is both a fixed and a variable annuity, and we will impute the investor's level of risk aversion from the allocation chosen between the two annuities.

There is a nascent body of research on the general topic of liquidity, marketability, and the bid-ask spread. Various empirical and theoretical studies, such as Silber (1991), Amihud and Mendelson (1991), and more recently, Jacoby, Gottesman, and Fowler (2000), Garvey (2001), Brenner, Eldor, and Hauser (2001), Dimson and Hanke (2001), and Loderer and Lukas (2001), have argued and documented that the yield to maturity, or investment returns, on less liquid financial instruments might be higher compared to their identical liquid counterparts. In related articles, Faig and Shum (2002) investigated the relationship between illiquidity and portfolio choice, whereas Cao and Wei (2002) examined the valuation of restricted and hence highly illiquid stock options.

However, it appears that limited research has been done on developing a subjective metric for computing the demanded *ex ante* compensation for illiquidity. The exception is a series of articles by Longstaff (1995, 2001). We will provide a more detailed comparison to Longstaff's model, later in our analysis.

The remainder of this article is organized as follows. In the next section we demonstrate the simple economic intuition that underlies our model using a basic numerical example. The second section develops a formal utility-based model for the liquidity premium in the case of a *savings* (payin) annuity, where the time horizon is deterministic and the product is akin to a zero-coupon bond or a Certificate of Deposit. The third section solves the model using numerical techniques, with comparative statics and a comparison to Longstaff's approach discussed in two subsections. The fourth section provides a parallel analysis for a *consumption* (payout) annuity, where payments are received by the annuitant with embedded longevity insurance. Finally, the last section concludes the article.

Numerical Example

To understand the welfare loss from a lack of liquidity we offer the following stylized example. Consider a hypothetical investor (or policyholder) with \$100,000 to invest, and assume this sum is the bulk of their financial wealth. The investor decides to allocate 50 percent to a fixed annuity (risk-free asset) with liquidity restrictions and 50 percent to a risky equity (variable) annuity.[1] Further, we make the critical assumption that *the investor has picked this allocation because it maximizes his or her expected utility of wealth.*

In the language of Merton (1969), we let $\alpha_t^* = 1/2, \forall\, t \leq T$, denote the optimal allocation to the risky asset, and we let U_T^* denote the maximal expected utility, at the terminal horizon T. Merton (1969) demonstrated that an investor with constant relative risk aversion (CRRA) preferences for uncertain wealth at the terminal time T, modeled by $u(w) = w^{(1-\gamma)}/(1-\gamma)$, and faced with geometric Brownian motion asset dynamics, will select a time-invariant (myopic) investment policy. This well-known Merton result has been generalized to alternative asset processes and consumer preferences. See Kim and Omberg (1996) for more details on the necessary and sufficient conditions for myopic investment policies.

We caution the reader that an $\alpha_t^* = 1/2$ allocation, also known as constant proportional strategy, does *not* imply the portfolio is invested half in equities and half in cash, and then held *as is* until maturity. A buy-and-hold strategy is suboptimal in a classical Merton framework. Indeed, our 50/50 balance must be maintained by reacting to market movements and rebalancing the portfolio. In other words, rational utility-maximizing behavior requires frequent trading and rebalancing regardless of one's investment horizon or risk preferences. See Browne (1998) for more information on constant proportional strategies.

Suppose, for example, that the general stock market drops 30 percent within a short period of time. And, as a result, the value of the equity account (variable annuity) drops from \$50,000 to \$35,000 (=\$50,000 × 70 percent). The investor now has only \$85,000 in total, of which, by construction, 41 percent (=\$35,000/\$85,000) is in the equity account, and 59 percent (= \$50,000/\$85,000) is in the fixed annuity. The investor is holding a nonoptimal portfolio, which, in theory, should be rebalanced.

A rational investor will want to sell a portion of (or transfer from) the fixed annuity into the equity account to reestablish the optimal 50/50 mix between fixed and variable investments. Specifically, the investor will want to transfer \$7,500 from the fixed annuity to the variable account so that \$42,500 is invested in fixed assets, and \$42,500 is invested in variable assets; thus maintaining the delicate $\alpha_t^* = 1/2$ mix. Of course, if this investor has a large portfolio of other (e.g., nonpension) assets, they might use those to achieve the required asset allocation.

Our main point is that liquidity restrictions in the fixed annuity will impede the optimal process of reallocation. This is the forgone opportunity cost. Even the prudent buy-and-hold investor will want to rebalance assets after a substantial market movement.

---

[1] The 50/50 allocation is chosen for purely symmetrical reasons simply to illustrate the example. In the next section we provide results for general asset allocations and discuss the impact of the preexisting mix on the liquidity premium.

TABLE 1
Required Liquidity Premium Spread, in Basis Points, per Annum Assuming a Desired 50/50 Allocation to (Risky) Equities and (Safe) Cash Earning 5% p.a.

Coefficient of Relative Risk Aversion, $U(w) = w^{(1-\gamma)}/(1-\gamma)$			
	$\gamma = 1$	$\gamma = 2$	$\gamma = 3$
	$T = 15$ Years		
$\mu = 12.5\%$	62.29	34.59	23.92
$\mu = 15.0\%$	98.66	55.73	38.76
$\mu = 17.5\%$	139.34	79.80	55.75
$\mu = 20.0\%$	183.36	106.24	74.48
	$T = 10$ Years		
$\mu = 12.5\%$	47.56	25.84	17.74
$\mu = 15.0\%$	77.04	42.50	29.33
$\mu = 17.5\%$	110.78	61.90	42.90
$\mu = 20.0\%$	147.98	83.59	58.14
	$T = 5$ Years		
$\mu = 12.5\%$	28.10	14.80	10.05
$\mu = 15.0\%$	47.04	25.07	17.09
$\mu = 17.5\%$	69.54	37.44	25.62
$\mu = 20.0\%$	95.12	51.68	35.47
	$T = 1$ Year		
$\mu = 12.5\%$	6.67	3.38	2.27
$\mu = 15.0\%$	11.67	5.94	3.98
$\mu = 17.5\%$	17.95	9.17	6.16
$\mu = 20.0\%$	25.45	13.05	8.78

Notes: For example, an investor with a 50/50 allocation to equities and cash, with CRRA = 1 (log utility) preferences, would require a yield enhancement (lambda) of 77 basis points on the cash account (i.e., 5.77%) to compensate for the inability to rebalance for 10 years; this is assuming they expected the equity account to earn 15% p.a. during this period. In contrast, if the investor expects to earn 17.5% from the equity account, they would require a 111 basis point liquidity spread. Similarly, for a fixed 50/50 allocation and the same (subjective) equity return, a higher CRRA (i.e., a more risk-averse investor), requires less compensation for the inability to rebalance.

We argue that the only way to make up for the inability to adapt to market movements is to offer an enhanced yield on the fixed annuity. Stated differently, a rational investor will be willing to waive his or her ability to instantaneously rebalance the portfolio in exchange for an enhanced yield on the fixed annuity. Our definition of liquidity yield is meant to provide the same level of economic utility for the constrained investors, as the unenhanced risk-free asset provides to the unconstrained investor.

Table 1 applies our model—which we will fully develop in the next section—to a particular set of parameters and displays our main result. By static utility we mean the maximal utility that can be obtained from picking an asset mix and *holding it* for the entire horizon. By dynamic utility we mean the Merton (1969) values that arise from *rebalancing to maintain a 50/50 mix*. As one can see from the table, *ceteris paribus*, a longer time horizon, lower level of risk aversion, and higher subjective growth rate from the market all appear to imply a larger liquidity premium for the range of parameters we have investigated. The next section presents the formal model that was used to generate Table 1.

THE UTILITY MODEL

Our model draws heavily from the classical Merton (1969) framework, and we thus take the liberty of omitting some stages in the derivation. The unrestricted investor can rebalance and allocate assets in continuous time between two assets (subaccounts) under the annuity umbrella. The first is the market (risky, equity) asset that obeys a diffusion process:

$$dV_t = \mu V_t dt + \sigma V_t dB_t \qquad V_0 = 1, \quad 0 \leq t \leq T, \tag{1}$$

where B_t is a standard Brownian motion, μ is the subjective growth rate of the market, and σ is the subjective volatility. This leads to

$$V_T = e^{(\mu - \sigma^2/2)T + \sigma B_T}. \tag{2}$$

We stress the word *subjective* since the desire to rebalance, and the optimal allocation, will depend critically on the individual's assessment of future market returns and volatility.

The second asset is the fixed annuity, or the classically labeled risk-free asset, which obeys

$$dA_t = rA_t dt, \qquad A_0 = 1 \quad \Longleftrightarrow \quad A_T = e^{rT}. \tag{3}$$

In our (simplistic) model, the fixed annuity (bond) pays a constant yield-to-maturity regardless of the time horizon. In practice, of course, one might expect to see a nonflat yield curve, and, as a result, the return on the fixed annuity would be a function of the maturity of the product. However, our intention is to exclude, or control for, term-structure premium effects and focus exclusively on liquidity (marketability) issues. As such, we have decided to operate in a flat curve environment. Our main qualitative results are unaffected by the introduction of a stochastic term structure model, which would then force us to keep track of three assets, namely bonds, cash, and the variable account.

The end-of-period utility function is of the form

$$u(w) = \frac{w^{(1-\gamma)}}{1-\gamma}, \qquad \gamma \neq 1, \tag{4}$$

and $u(w) = \ln[w]$ when $\gamma = 1$. Furthermore, without any loss of generality, we assume the investor starts with one ($1) unit of account (wealth).

Following Merton (1969), the optimal control problem results in a partial differential equation (PDE), which leads to the maximal level of (dynamic) expected utility,

$$EU^{*}(r \mid \text{dynamic}) = \frac{1}{1-\gamma} e^{\xi(1-\gamma)T}, \tag{5}$$

where

$$\xi = r + \frac{(\mu - r)^2}{2\gamma\sigma^2}. \tag{6}$$

In this framework,

$$\alpha_t^{*} = \frac{\mu - r}{\gamma\sigma^2}, \tag{7}$$

which we label the *Merton Optimum.*

In contrast to the dynamic case, a static allocation—which is forced by the liquidity restrictions—will induce a maturity value of wealth that is the linear sum of the monies allocated to the two accounts. The expected utility from this static portfolio—with no liquidity enhancement—is defined as

$$EU(r \mid \text{static}) := EU[(1-\alpha)A_T + \alpha V_T] = EU\big[(1-\alpha)e^{rT} + \alpha e^{(\mu-\sigma^2/2)T+\sigma B_T}\big]. \tag{8}$$

And, by definition of the optimal allocation,

$$EU^{*}(r \mid \text{static}) \le EU^{*}(r \mid \text{dynamic}), \tag{9}$$

with equality occurring when $\alpha^{*} = 1$ or when $\alpha^{*} = 0$. So, we formally define the liquidity premium λ as the enhancement to r that will induce the same level of expected utility. In other words

$$EU^{*}(r + \lambda \mid \text{static}) = EU^{*}(r \mid \text{dynamic}). \tag{10}$$

In the static case, the maximal expected utility is obtained via

$$EU^{*}(r + \lambda \mid \text{static}) = \max_{\alpha} E\left[\frac{1}{1-\gamma}\big((1-\alpha)e^{(r+\lambda)T} + \alpha e^{(\mu-\sigma^2/2)T+\sigma B_T}\big)^{1-\gamma}\right]. \tag{11}$$

Now, since B_T is normally distributed with mean zero, and variance T, Equation (11) can be rewritten as

$$\begin{aligned} &EU^{*}(r + \lambda \mid \text{static}) \\ &\quad = \max_{\alpha} \int_{-\infty}^{\infty} \frac{1}{1-\gamma}\big((1-\alpha)e^{(r+\lambda)T} + \alpha e^{(\mu-\sigma^2/2)T+\sigma\sqrt{T}x}\big)^{1-\gamma} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx. \end{aligned} \tag{12}$$

In sum, the (maturity-dependent) parameter λ is the required yield to compensate for illiquidity. It is an implicit function of the time horizon T, the coefficient of relative risk aversion (CRRA) γ, and the return-generating process parameters r, μ, σ. Our objective is to solve for λ.

We note, again, that if the investor has a large amount of exogenous investible wealth that can be used for rebalancing his or her portfolio, the curvature of the utility function around the optimum will be much lower, and therefore the demanded liquidity premium will be lower as well.

Of course, any model that attempts to combine risk preferences, γ, and equity market parameters, μ, σ, comes face-to-face with the so-called equity risk premium anomaly. A large part of the economics literature is reasonably convinced that $\gamma < 3$. See, for example, Feldstein and Ranguelova (2001), or Friend and Blume (1975), for estimates in that range. Similarly, recent work by Mitchell et al. (1999) in the economic annuities literature, has employed values ranging from $\gamma = 1$ to $\gamma = 3$. Dramatically different evidence is provided by Mankiw and Zeldes (1991) where $\gamma = 35$ and Blake (1996) where $\gamma = 25$. And, whereas a value of $\gamma = 1$ corresponds with log utility that has appealing growth-optimal properties, our tables and numerical estimates provide a range of values for high and low risk aversion levels.

Another difficult issue is that if we use recent (Ibbotson Associates, 2001) capital market experience of $\mu - r = 6$ percent, and $\sigma = 20$ percent, then Equation (7) leads to an equity allocation of $\alpha_t^* = 246$ percent for a log-utility investor, and $\alpha_t^* = 123$ percent for a (more risk averse) $\gamma = 2$ investor. Clearly, these allocations are much higher than what is observed in practice.

Thus, to avoid this problem—while at the same time conditioning on a well-balanced portfolio—we decided to invert Equation (7) and locate market parameters that "fit" the Merton (1969) model. Specifically, we assume a preexisting asset allocation α, CRRA γ, and risk premium $\mu - r$, and solve for the (implied) subjective volatility assessment $\sigma = \sqrt{(m - r)/\gamma\alpha}$ that is consistent with Merton's optimum. The implied (subjective) volatility, which is higher than historical values, is motivated by a similar approach in the options market, and attempts to capture the possible model (jump) risks that are not reflected in the classical diffusion approach. The "adjust the volatility" approach to using asset allocation, and specifically option pricing models, originates with Rubinstein (1994).

SOLVING FOR THE REQUIRED λ

Due to the complexity of Equation (11), we are forced to use numerical methods to extract λ. We start by fixing a value for the risk-free rate r. Then, for any exogenously imposed value of the CRRA γ and a subjective rate of return μ, we impute the investor's subjective volatility σ from Equation (7). For simplicity, consider the case $\gamma \neq 1$ (the case of logarithmic utility $\gamma = 1$ can be treated similarly). Then we seek a value of λ such that the maximum of

$$F(\alpha, \lambda) = \int_{-\infty}^{\infty} \frac{1}{1-\gamma}\left((1-\alpha)e^{(r+\lambda)T} + \alpha e^{(\mu-\sigma^2/2)T+\sigma\sqrt{T}x}\right)^{1-\gamma} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \tag{13}$$

equals $U^{*}(r \mid \text{dynamic})$. In other words, we are seeking a solution to the pair of equations

$$F(\alpha, \lambda) = U^{*}(r \mid \text{dynamic}), \qquad \frac{\partial F}{\partial \alpha}(\alpha, \lambda) = 0. \tag{14}$$

This may be found numerically using Newton's method, where we alternate Newton steps in the λ and α variables. Basically, we solve two equations in two unknowns. We start with an initial approximation to the solution (α_0,λ_0). Then, we do a Newton step as a function of the first variable (holding the second variable fixed). This gives a better approximation, denoted by (α_1,λ_0). Then we hold the first variable fixed and look at it as a function of the second variable, and do a Newton step again. This gives an even better approximation (α_1, λ_1). Then we go back to the first variable and get a better approximation (α_2, λ_1), etc.

To carry this out we require expressions for the functions

$$F, \quad \frac{\partial F}{\partial \alpha}, \quad \frac{\partial^2 F}{\partial \alpha^2}, \quad \frac{\partial F}{\partial \lambda}. \tag{15}$$

But in fact, each of these is easily computed as an integral of simple functions against the standard normal density function. So to carry this out efficiently, all that is required is a method of rapid repeated calculation of such integrals. The method of choice is the Gauss-Hermite integration (see Press et al. 1997, Chapter 4), in which a single computation of nodes x_i and weights w_i allows one to write

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} f(x)\, dx \approx \sum_{i=1}^{N} w_i f(x_i) \tag{16}$$

for any regular function f. If the x_i are evenly spaced then the above formula would correspond to something like the trapezoidal rule or Simpson's rule, which is robust but not particularly accurate unless N is large. But if f is known to be highly smooth, then the nodes and weights can be tailored to achieve high accuracy with modest values of N. Gauss-Hermite quadrature makes the approximation of (16) exact for f any polynomial of degree less than $2N - 1$. Once again, Table 1 provides values for λ for various (subjective) levels of μ and time horizons T.

Comparative Statics

In Table 2, we display the required liquidity premium, λ, as a function of the underlying interest rate earned by the risk-free rate. As one can see, the greater the interest rate, the lower is the optimal liquidity premium. Though it might appear from Table 2 that the interest rate, *per se*, is what determines the required yield, it is the *actual spread* between the expected return from the market, μ, and the interest rate that drives this result. Indeed, with a $\mu = 12.5$ percent, the higher the level of interest rates, the lower the spread, and thus the lower is the opportunity cost of *not* being able to rebalance.

Once again, we caution the reader that underlying our result is an equity risk premium that also affects the opportunity loss. Thus, for example, when unrestricted cash earns

TABLE 2

Liquidity Premium as a Function of the Risk-Free Interest Rate, Conditional on a 50/50 Allocation to Equities versus Cash, and Assuming an Equity Return Expectation of 12.5%, over a 10-year Investment Horizon

Interest	CRRA = 1	CRRA = 2	CRRA = 3
4%	58.77	32.14	22.11
5%	47.56	25.84	17.73
6%	37.21	20.07	13.74
7%	27.81	14.89	10.16

See Notes to Table 1.

$r = 5$ percent, equity is expected to earn $\mu = 12.5$ percent, the investor has a CRRA $\gamma = 2$, *and* a preexisting portfolio of 50 percent cash and 50 percent equity, the implied subjective volatility assumption is $\sigma = 31.62$ percent.

Table 3 displays the required liquidity premium as a function of the preexisting asset allocation, different return expectations and a variety of high and low risk aversion levels. Thus, for example, a $\gamma = 3$ individual with a desired 20 percent allocation to risky equity, and an 80 percent allocation to unrestricted cash, will demand a liquidity premium of 35.86 basis points per annum as compensation for being unable to trade during a 10-year investment horizon.

As one can see from Table 3, the relationship between desired (or preexisting) equity holdings, and the demanded liquidity premium resembles an inverted parabola, and is zero at both ends. The intuition is as follows. A rational individual with a desired (or preexisting) allocation of either, 0 percent or 100 percent unrestricted cash, will not engage in any trading during the length of the investment horizon (with probability one) since there will never be a need to rebalance. However, as the portfolio moves toward a more balanced composition, the probability and magnitude of rebalancing increases, thus magnifying the required liquidity premium for not being able to trade.

The same inverted-parabolic relationship exists for higher levels of risk aversion, but in a decreasing manner since *the opportunity cost of not being able to trade is lower.* For example, one sometimes finds relatively high values of the risk aversion parameter γ in use. For example, Blake (1996) has estimated CRRA values as high as $\gamma = 25$ in the UK. Thus, if we were to take $\gamma = 25$, for example, then the only way a dynamic allocation of $\alpha = 50$ percent could be rational would be if the individual's estimate of future volatility was particularly low, in which case there is less difference between the risky and the risk-free asset, so the liquidity premium should also be low. This is indeed the case. To get a sense of the magnitudes, if $\mu = 15$ percent and $T = 10$, we compute $\lambda = 3.75$ b.p.

Finally, Table 4 provides a different perspective on the results. We start by fixing the level of μ, r, and σ, and then solve for the optimal static allocation α^* and liquidity premium λ, as a function of risk aversion γ. We contrast these numbers with the classical Merton (1969) allocation. We used historical values of $r = 5$ percent,

TABLE 3

Liquidity Premium as a Function of Asset Allocation, Equities (Risky Asset) versus Cash (Safe Asset), Assuming Cash Earns 5%, a 10-year Horizon and an Expected Return from Equity of 15% and 10%

Equity %	CRRA = 1	CRRA = 2	CRRA = 3	CRRA = 5	CRRA = 10	CRRA = 25
Expected return from equity of 15%						
0	0.00	0.00	0.00	0.00	0.00	0.00
10	47.92	39.00	31.44	22.17	12.58	5.44
20	74.12	48.74	35.86	23.34	12.43	5.17
30	82.38	49.45	35.18	22.26	11.60	4.76
40	82.11	46.87	32.74	20.41	10.51	4.28
50	77.04	42.50	29.33	18.10	9.25	3.75
60	68.60	36.83	25.16	15.40	7.82	3.16
70	57.29	29.97	20.28	12.32	6.22	2.50
80	42.96	21.84	14.63	8.80	4.41	1.77
90	24.79	12.11	8.00	4.76	2.36	0.94
100	0.00	0.00	0.00	0.00	0.00	0.00
Expected return from equity of 10%						
0	0.00	0.00	0.00	0.00	0.00	0.00
10	20.87	14.69	11.11	7.40	4.01	1.69
20	27.09	16.24	11.55	7.31	3.81	1.56
30	27.66	15.61	10.86	6.75	3.47	1.41
40	26.14	14.26	9.80	6.03	3.07	1.24
50	23.52	12.53	8.55	5.22	2.65	1.07
60	20.16	10.54	7.14	4.34	2.19	0.88
70	16.18	8.31	5.59	3.38	1.70	0.68
80	11.58	5.83	3.90	2.34	1.17	0.47
90	6.28	3.09	2.04	1.22	0.61	0.24
100	0.00	0.00	0.00	0.00	0.00	0.00

See Notes to Table 1.

$\mu = 11$ percent, and $\sigma = 20$ percent, which are consistent with the Ibbotson Associates (2001) numbers. As in Table 3, the liquidity premium is not monotonic. When CRRA is small, the Merton allocation to equities is near 100 percent, and when CRRA is large, the Merton allocation is near 0 percent. In either case the liquidity premium will be small, since there is little rebalancing in the liquid portfolio. The liquidity premium is largest when CRRA takes some intermediate value, in this example at about $\gamma = 2.7$. This, once again, is consistent with our main observation that individuals with a well-diversified (but small) investment portfolio and a relatively low level of risk aversion will demand the highest liquidity premium.

Comparison to Longstaff's Model

Although we took a similar approach to computing the welfare loss for liquidity restrictions, our article differs from Longstaff's 2001 work in a number of substantial ways. First, our model assumed a general constant relative risk aversion (CRRA) utility specification, in contrast to Longstaff's logarithmic utility model. This allowed us to

TABLE 4

Liquidity Premium and Asset Allocation as a Function of Relative Risk Aversion Equities (Risky Asset) versus Cash (Safe Asset) Assuming Cash Earns 5%, a 10-year Horizon, Historical Volatility of 20% and an Expected Return from Equity of 11%

	CRRA = 1	CRRA = 2	CRRA = 3	CRRA = 5	CRRA = 10	CRRA = 25
Merton allocation	150.00%	75.00%	50.00%	30.00%	15.00%	6.00%
Constrained allocation	NA	75.27%	48.93%	28.35%	13.72%	5.37%
Liquidity premium (b.p.)	NA	10.05	11.90	9.31	5.40	2.33

Notes: With CRRA = 1 the investor holds no fixed annuities so the constrained allocation and liquidity premium is NA. Using historical parameters, the liquidity premium peaks at a CRRA of approximately 2.73, with a value of 12.02 b.p. In addition, see Notes to Table 1 for explanations.

explore the critical impact of risk aversion on the value of liquidity, as well as the effect of holding period restrictions. Indeed, as Tables 1–3 indicated, risk aversion played an important role in determining the required liquidity premium. *Ceteris paribus*, the greater the aversion to risk, the lower the required liquidity premium, holding asset allocation constant. Also, whereas Longstaff modeled illiquidity in a stochastic volatility environment, and used trading strategies that were of bounded variation, we operated in a much simpler Merton (1969) environment, which allowed for closed-form solutions to the optimal portfolio holdings. (We traded off stochastic volatility for general utility.) Our liquidity premium—which was formulated as a yield, as opposed to Longstaff's discount—was obtained by solving a one-dimensional integral equation.

However, the most important distinction was that we focused on the individual's preexisting portfolio and asset allocation as a determinant of the liquidity premium. As one can see from Table 3, an individual with a very low, or very high, level of holdings in the illiquid bond (annuity) would not require as much compensation as the individual with a relatively well-balanced portfolio. The liquidity premium is directly related to the probability (and magnitude) of having to trade and rebalance during the life of the restriction. If the preexisting optimal portfolio is well balanced—i.e., close to equal amounts of equity and cash—there is a higher chance of the portfolio falling out of balance, and thus requiring trading to move back to the optimum.

As such, our conclusions complement Longstaff (2001), in that we concur that "discounts for illiquidity can be substantial," but we also demonstrate that the magnitude depends on the individual's risk aversion and preexisting portfolio.

PAYOUT ANNUITIES AND LONGEVITY INSURANCE

Within the universe of annuities, the natural context for the foregoing section is the accumulation phase, during which contributions are held and invested prior to retirement. We now turn to the payout phase of the life annuity in order to illustrate the applicability of the technique in the context of a random maturity. We assume that this involves the purchase of an immediate life annuity at time $t = 0$, entitling the holder to a continuous stream of payments, terminating upon death, which is now a random time $t = T$.

The payout annuity can be some combination of a fixed immediate annuity (FIA), which provides a fixed payment per unit time, and a variable immediate annuity (VIA), which provides a payment per unit time that varies depending on the value of some market asset V_t. If w dollars of the FIA are purchased, the consumer is entitled to continuous payment stream of $C_t^F = w/a_x(r)$ dollars per unit time, where the unit price of the FIA is

$$a_x(r) = \int_0^\infty e^{-rt}({}_t p_x)dt. \tag{17}$$

Here r denotes the risk-free interest rate, and $({}_t p_x)$ is the probability that the individual will survive to time t, conditional on being alive at the annuity purchase age x. The normalization is that each unit of the FIA pays \$1 per unit time.

Similarly, if w dollars of the VIA are purchased, the consumer receives payments based on $w/a_x(h)$ units of the market asset per unit time, where h is the assumed interest rate (AIR). In other words, at time t, payments accumulate at the rate of $C_t^V = we^{-ht}V_t/a_x(h)$ dollars per unit time, where we have normalized the market asset so that $V_0 = 1$.

As before, we will compare liquid and illiquid annuities. In the liquid case, the consumer is free to exchange FIA units for an economically equivalent number of VIA units at any time, and vice versa. In the illiquid case, the number of FIA and VIA units is fixed at the time of purchase. Other things being equal, the liquid annuity would provide greater utility to the consumer, so to compensate for this the illiquid annuity must provide an enhanced rate of return. As in the preceding section, we assume that it is the FIA that is so enhanced. In this context, we take this to mean that an investment of w dollars in the FIA produces a payment stream of $C_t^F = w/a_x(r + \lambda)$ dollars per unit time, where λ is the demanded liquidity premium.

We will assume that the AIR is chosen to be the risk-free rate, so that $h = r$. Such a restriction is not uncommon in annuity products available for sale, and is in fact typical of the liquid ones.

In the preceding section the consumer's utility involved only end-of-period wealth, since there were no funds available for consumption prior to that time horizon. In the present case, it is exactly the utility of consumption that is of interest, as it is discounted to take account of time preferences. Thus, if C_t denotes the payment stream generated by the life annuity, and if the function $u(\cdot)$ denotes the consumer's personal utility of consumption, then the mix between the fixed and variable annuities will be selected so as to maximize

$$E\left[\int_0^T e^{-rt}u(C_t)dt\right] = \int_0^\infty e^{-rt}({}_t p_x)E[u(C_t)]dt, \tag{18}$$

where T is the random time of death, and assuming independence of asset returns and mortality.

As before, we will assume an optimal 50/50 mix between the fixed and variable annuities in the liquid case, and then impute model parameters. As we illustrated quite extensively in the entitled section solving for the required λ, the technique can be used for any preexisting asset allocation. We have selected one particular mix to contrast the deterministic horizon results with their stochastic analog in the payout phase. We continue to assume geometric Brownian dynamics for the risky asset V_t, so

$$dV_t = \mu V_t dt + \sigma V_t dB_t, \qquad V_0 = 1. \tag{19}$$

Consistent with our results in Table 4, we use historical capital market values to obtain the implied asset mix as a function of risk aversion. Charupat and Milevsky (2002) consider the asset allocation problem in the setting of liquid annuities. Assuming $h = r$, and an exponential or Gompertz mortality function, they show that the Merton optimum

$$\alpha^* = \frac{\mu - r}{\gamma\sigma^2} \tag{20}$$

remains optimal in this new setting. In fact, this can be proved more generally—and is actually alluded to in Chapter 18 of Merton (1994)—and does not depend on the parametric form of the survival probabilities (${}_t p_x$). Denote by ϕ_t and ψ_t the number of units of the FIA and VIA held at time t, and assume that $h = r$. Then the payment stream is $C_t = \phi_t + \psi e^{-rt} V_t$, and it can be shown that the optimal choice of ϕ_t and ψ_t obeys

$$\alpha^* = \frac{\psi_t e^{-rt} V_t}{\phi_t + \psi_t e^{-rt} V_t} = 1 - \frac{\phi_t}{\phi_t + \psi_t e^{-rt} V_t}, \tag{21}$$

for α^* as above. In particular, from the assumption that $\alpha^* = 50$ percent, and the given (Ibbotson Associates, 2001) values for μ, r, and σ, we may impute a CRRA value of $\gamma = 3$, regardless of the form of the conditional probability of survival (${}_t p_x$). In other words, if the individual has a coefficient of relative risk aversion of $\gamma = 3$, and is faced with a market in which the expected return from the risky asset is $\mu = 11$ percent, with a volatility of $\sigma = 20$ percent, when the risk-free rate is $r = 5$ percent, then he will allocate exactly $\alpha = 50$ percent to each of the two asset classes.

It can further be shown that with this choice of allocation,

$$E[u(C_t)] = \frac{1}{1-\gamma} e^{\beta t} \left(\frac{w}{a_x(r)} \right)^{1-\gamma}, \tag{22}$$

where

$$\beta = \frac{(1-\gamma)(\mu - r)^2}{2\gamma\sigma^2}. \tag{23}$$

Thus,

$$U^* = \int_0^\infty e^{-rt}({}_t p_x) E[u(C_t)]dt = \frac{w^{(1-\gamma)}}{1-\gamma} \frac{a_x(r-\beta)}{a_x(r)^{(1-\gamma)}}, \tag{24}$$

in the dynamic liquid case.

In the static (illiquid) case, an initial allocation of α to the risky asset will result in holding $\phi_t = (1-\alpha)w/a_x(r+\lambda)$ FIA units, and $\psi_t = \alpha\ w/a_x(r)$ VIA units, and in a utility

$$\begin{aligned} F(\alpha,\lambda) &= \int_0^\infty e^{-rt}({}_t p_x) E[u(C_t)]dt \\ &= \int_0^\infty e^{-rt}({}_t p_x) \int_{-\infty}^\infty \frac{w^{1-\gamma}}{1-\gamma} \left(\frac{1-\alpha}{a_x(r+\lambda)} + \frac{\alpha e^{(\mu-r-\sigma^2/2)t+\sigma\sqrt{t}z}}{a_x(r)} \right)^{1-\gamma} \\ &\quad \times \frac{1}{\sqrt{2\pi}} e^{-z^2/2}\, dz\, dt. \end{aligned} \tag{25}$$

Our goal is, as before, to find the liquidity premium λ such that maximizing $F(\alpha,\lambda)$ over α reproduces the dynamic utility $U^*(r \mid \text{dynamic})$. We will do so by assuming Gompertz mortality, corresponding to an exponentially increasing hazard rate (force of mortality) of the form

$$h_{x+t} = \frac{1}{b} e^{(x+t-m)/b}. \tag{26}$$

(m, b) are the Gompertz parameters and x is the individual's age at the time of purchase. In this case the survival probability takes the form

$$({}_t p_x) = \exp\big(b h_x (1 - e^{t/b})\big). \tag{27}$$

As before, we use Newton's method to carry out the maximization and root finding, and we use Gauss-Hermite quadrature to rapidly evaluate the Gaussian integral in the expression for $F(\alpha,\lambda)$. We carry out the time integral using a related method, namely Gauss-Laguerre quadrature. As in Equation (16), the Gauss-Laguerre nodes and weights x_i and w_i are optimized for computing integrals of the form

$$\int_0^\infty e^{-at} t^c f(t)dt \approx \sum_{i=1}^N w_i f(x_i), \tag{28}$$

where f is well approximated by a polynomial function. We use $c = 0$ and must be careful to choose a in a narrow range of values for which the method is stable when applied to our integrands. But having done so, this gives a rapid and accurate algorithm.

We consider two cases, both corresponding to an age of 62 years at the time of annuitization. The first uses Gompertz parameters fit to the U.S. Society of Actuaries female

(IAM1996) mortality data (namely, $b = 8.78$ and $m = 92.63$) and yields an optimal allocation of $\alpha = 48.40$ percent and a liquidity premium of $\lambda = 13.07$ basis points. In the second case we use male mortality parameters $b = 10.5$ and $m = 88.18$, and compute $\alpha = 48.50$ and $\lambda = 12.52$ b.p.

To understand the factors influencing these results, one can calculate the conditional life expectancy, resulting in figures of $e_{62} = 26.62$ years (female) and $e_{62} = 22.78$ years (male). Using those time horizons in the fixed maturity problem of the previous section gives $\alpha = 47.68$ percent, $\lambda = 24.22$ b.p., and $\alpha = 47.93$ percent, $\lambda = 21.87$ b.p., respectively. These premiums are substantially higher than the Gompertz figures just computed, and a moment's reflection will spot the reason why. We saw that the liquidity premium increases rapidly with the time horizon, and in the annuity context most of the payments occur significantly earlier than the lifetime itself. Thus the effect of spreading payments out over the residual lifetime should be to reduce the liquidity premium. Indeed, even if the residual lifetime were to take on a deterministic value T, the mean time for which a payment is received is

$$\frac{1}{T}\int_0^T t\,dt = \frac{T}{2}. \tag{29}$$

Thus to appreciate the sensitivity of the results to the randomness of the life horizon T, we should not compare with the results of the preceding section, but rather with other lifetime distributions having the same means. As an extreme case, we compare the Gompertz results with deterministic lifetime distributions, that is, with survival functions

$$({}_t p_x) = \begin{cases} 1, & t < t_0 \\ 0, & t \geq t_0 \end{cases} \tag{30}$$

where t is set to the mean residual Gompertz lifetimes. Because of the discontinuity in $({}_t p_x)$ we use yet another quadrature method, namely Gauss-Legendre, which is optimized for integrals of the form $\int_0^{t_0} f(t)dt$. This gives optimal allocations of $\alpha = 48.48$ percent (female) and $\alpha = 48.63$ percent (male), and liquidity premiums of $\lambda = 13.02$ b.p. (female) and $\lambda = 12.20$ (male). These premiums are extremely close to those obtained under Gompertz mortality, which suggests that the premiums are not highly sensitive to the precise form of the hazard rate. Note, however, that both Gompertz figures are slightly higher, and it is tempting to describe the difference as a small additional premium for mortality risk.

CONCLUSION

This article has argued that the value of liquidity, or the lack thereof, can be assessed by returning to basic "Mertonian" (1969) principles. We define liquidity premium as the demanded enhancement to the risk-free rate that compensates for the inability to continuously rebalance an investment portfolio due to institutionalized restrictions. In other words, a liquidity premium added to an illiquid product should produce the same level of utility as the unrestricted product without the liquidity premium. Although we formulated our model within the context of a fixed savings or payout annuity—which traditionally is associated with substantial liquidity restrictions—our model can be applied more broadly. Our economic model led to an interesting mathematical problem, which was to locate the yield λ that equated maximal utility in a static portfolio to the (greater) utility from a portfolio that could be dynamically rebalanced in a Merton framework.

Our main results are as follows: We find that a log-utility ($\gamma = 1$) investor, with a preexisting 50/50 asset mix between fixed and variable *savings* annuities, would demand a liquidity premium of between 25 and 155 basis points per annum—depending on his or her expected return from the equity market—as compensation for the inability to rebalance a portfolio during a 10-year period. However, if the same investor has a preexisting asset mix that consists of 90 percent variable and only 10 percent fixed annuities, the required premium drops to between 5 and 40 basis points, depending on future market expectations. Furthermore, for investors that are more risk averse ($\gamma > 1$), and/or who are faced with shorter liquidity restrictions, the compensating liquidity premium can be much lower. Indeed, for a 1 year period, and coefficient of relative risk aversion ($\gamma = 3$) the premium ranges from *only* 2 to 8 basis points per annum above the risk-free yield. As such, we are careful to conclude that the question of liquidity is personal in nature, since it depends on attitudes toward financial risk, current portfolio holdings, and subjective expectations about future investment returns.

However, regardless of the magnitude of this effect, our article does find that impeding a consumer's ability to periodically rebalance his or her investment portfolio is detrimental to their economic well-being (utility) even if they have no expectation of doing so. This is regardless of their preexisting asset allocation, investment time horizon, or subjective market expectations.

Finally, research currently underway by the authors will attempt to develop a model in which only one-sided trading restrictions are imposed so that additional assets can be purchased, but not sold. We anticipate the liquidity premium will be lower in this case, but the amount by which it is reduced remains an open question.

REFERENCES

Amihud, Y., and H. Mendelson, 1991, Liquidity, Maturity and the Yields on U.S. Treasury Securities, *Journal of Finance*, 46: 1411-1425.

Blake, D., 1996, Efficiency, Risk Aversion & Portfolio Insurance: An Analysis of Financial Asset Portfolios Held by Investors in the UK, *Economic Journal*, 106: 1175-1192.

Blake, D., and W. Burrows, 2001, Survivor Bonds: Helping to Hedge Mortality Risk, *Journal of Risk and Insurance*, 68(2): 339-348.

Brenner, M., R. Eldor, and S. Hauser, 2001, The Price of Options Illiquidity, *Journal of Finance*, 56(2): 789-805.

Brown, J. R., and J. M. Poterba, 2000, Joint Life Annuities and Annuity Demand by Married Couples, *Journal of Risk and Insurance*, 67(4): 527-554.

Browne, S., 1998, The Return on Investment from Proportional Portfolio Strategies, *Advances in Applied Probability*, 30(1): 216-238.

Cao, M., and J. Wei, 2002, Restricted Stock Options, Schulich School of Business Working Paper.

Charupat, C., and M. A. Milevsky, 2002, Asset Allocation in Variable Annuities: A Note, *Insurance: Mathematics and Economics*, 30(2): 199-210.

Dimson, E., and B. Hanke, 2001, The Expected Illiquidity Premium: Evidence from Equity Index-Linked Bonds, Working paper, London Business School.

Faig, M., and P. Shum, 2002, Portfolio Choice in the Presence on Personal Illiquid Projects, *Journal of Finance*, 57(1): 303-327.

Feldstein, M., and E. Ranguelova, 2001, Individual Risk in an Investment-Based Social Security System, *American Economic Review*, 91(4): 1116-1125.

Friend, I., and M. E. Blume, 1975, The Demand for Risky Assets, *American Economic Review*, 65: 900-922.

Garvey, G. T., 2001, What is an Acceptable Rate of Return for an Undiversified Investor? Working paper, Claremont Graduate School.

Ibbotson Associates, 2001, *Stocks, Bonds, Bills and Inflation: 2000 Yearbook*, Chicago, IL: Ibbotson Associates.

Jacoby, G., A. Gottesman, and D. Fowler, 2000, The Capital Asset Pricing Model and the Liquidity Effect: A Theoretical Approach, *Journal of Financial Markets*, 3(1): 69-81.

Kim, T. S., and E. Omberg, 1996, Dynamic Nonmyopic Portfolio Behavior, *Review of Financial Studies*, 9(1): 141-161.

Loderer, C., and R. Lukas, 2001, The Pricing Discount for Limited Liquidity: Evidence from Switzerland, Working paper, University of Berne, Switzerland.

Longstaff, F. A., 1995, How Much Can Marketability Affect Security Values? *Journal of Finance*, 50(5): 1767-1774.

Longstaff, F. A., 2001, Optimal Portfolio Choice and the Valuation of Illiquid Securities, *Review of Financial Studies*, 14(2): 407-431.

Mankiw, N., and S. Zeldes, 1991, The Consumption of Stockholders & Nonstockholders, *Journal of Financial Economics*, 29: 97-112

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

Merton, R. C., 1994, *Continuous-Time Finance*, rev. ed., Oxford, UK: Blackwell.

Mitchel O., J. Poterba, M. Warshawsky, and J. Brown, 1999, New Evidence on the Money's Worth of Individual Annuities, *American Economic Review*, 89(5): 1299-1318.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 1997, *Numerical Recipes in C*, 2nd ed., Cambridge, MA: Cambridge University Press.

Rubinstein, M., 1994, Presidential Address: Implied Binomial Trees, *Journal of Finance*, 49(3): 771-818.

Silber, W., 1991, Discounts on Restricted Stock: The Impact of Illiquidity on Stock Prices, *Financial Analysts Journal*, 47: 60-64.

Vayanos, D., and J. Vila, 1999, Equilibrium Interest Rate and Liquidity Premium with Transaction Costs, *Economic Theory*, 13: 509-539.