

# NEW MARKOV CHAIN MODELS TO ESTIMATE THE PREMIUM FOR EXTENDED HEDGE FUND LOCKUPS

by

Kun Soo Park,  
Graduate School of Management,  
KAIST Business School, Seoul, Korea 130-722  
kunsoo@business.kaist.ac.kr, +82-2-958-3329

and

Ward Whitt,  
Department of Industrial Engineering and Operations Research,  
Columbia University, New York, NY 10027-6699  
ww2040@columbia.edu, 212-854-7255

## *Abstract*

To estimate the premium an investor should expect from extended hedge fund lockups, Derman et al. (2009) proposed a three-state discrete-time Markov Chain to model the state of a hedge fund, allowing the state to change randomly among the states “good,” “sick” and “dead” every year. The lockup premium measures the consequence of being stuck with a sick fund. To be more realistic, we propose an alternative three-state continuous-time Markov Chain model, which allows state changes continuously in time. We develop new techniques for parameter fitting, exploiting nonlinear programming. We fit the parameters indirectly to readily available hedge fund performance measures: the persistence factor, the death rate and the variance of annual returns, estimated from TASS hedge fund data.

# NEW MARKOV CHAIN MODELS TO ESTIMATE THE PREMIUM FOR EXTENDED HEDGE FUND LOCKUPS

Kun Soo Park,  
Graduate School of Management, KAIST Business School, kunsoo@business.kaist.ac.kr  
Ward Whitt,  
Department of IE and OR, Columbia University, ww2040@columbia.edu

## *Abstract*

A *lockup period* for investment in a hedge-fund is a time period after making the investment during which an investor cannot freely redeem his investment. Since longer lockup periods have recently been imposed, it is important to estimate the premium an investor should expect from extended lockups. In Derman et al. (2009) we proposed a parsimonious three-state discrete-time Markov Chain (DTMC) to model the state of a hedge fund, allowing the state to change randomly among the states “good,” “sick” and “dead” every year. According to the lockup condition, the investment can be immediately redeemed from a dead fund (when it ceases to operate), to the extent possible. The lockup premium measures the consequence of being stuck with a sick fund. We fit the DTMC model parameters indirectly to readily available hedge fund performance measures, in particular to the persistence factor, the death rate and the variance of annual returns, estimated from TASS hedge fund data. We then showed how the DTMC model can be applied to estimate the lockup premium. In this paper, we propose an alternative three-state absorbing continuous-time Markov Chain (CTMC) model, which allows state changes continuously in time instead of yearly. The CTMC model is more realistic and flexible, but requires new techniques for parameter fitting. We employ nonlinear programming to solve the new calibration equations. We show that the more realistic CTMC model is a viable alternative to the previous DTMC model.

## **1. Introduction**

A *lockup period* for investment in a hedge-fund is a time period after making the investment during which an investor cannot freely redeem his investment. Hedge funds often require a lockup period in order to invest in illiquid assets (Aragon, 2007). The importance of investment liquidity was highlighted by the recent financial crisis of 2007-2009 (Golts and Kritzman, 2010).

In recent years, the required lockup period has increased. Boyle et al. (2010) found that in the Hedge Fund Research database, the average lockup period for funds with the lockup

condition is one year and the range of the lockup period is from one to four years. Ang and Bollen (2009) also pointed out that typical lockup period is one to three years. These suggest that the extended lockups have been more popular recently. The emergence of extended hedge fund lockups continues during the recent financial crisis; Kazemi (2010) reported that during the recent financial crisis, hedge funds with long lockup period could avoid selling their assets at distressed prices. Also, Ben-David et al. (2010) found that hedge funds with short lockup period is more likely to face selloffs caused by redemptions.

We thus wish to estimate the premium from extended hedge fund lockup. In doing so, we take the point of view of a manager of a fund of funds, who has to choose between two investments in similar funds in the same strategy category. Thus, we *define* the premium for extended hedge fund lockup as the annual fixed rate of return that compensates the difference of expected returns in two hedge-fund investments, with and without the extended lockup condition; see §3. This definition accounts for the lost gains due to the inability to rebalance the investment portfolio in hedge funds, but not for other lost investment opportunities, so this is a conservative estimate of the lockup premium. Investors can separately consider the cost of other lost investment opportunities. We believe that the lockup premium calculated here may be helpful for investors to see whether the shareholder restrictions caused by the extended lockup condition in hedge funds can be offset by the additional returns from the illiquid investment.

Since hedge funds are not required to report their returns by regulatory authorities, data on returns of hedge funds are relatively limited, compared to data for other securities. Thus, despite of the importance of extended hedge fund lockups, there has been limited research on the lockup premium for hedge fund investment; see §2.

### 1.1. The Initial Discrete-Time Model

In Derman et al. (2009) (*hereafter*, DPW09) we proposed a parsimonious discrete-time Markov chain (DTMC) model that can be calibrated to observable performance measures from hedge fund returns: the persistence factor, the death rate, and volatility (as measured by the variance or standard deviation of annual returns). We fit the model to hedge-fund return data from the Tremont Advisory Shareholder Services (TASS) database, using data from 2001-2005. We then applied the fitted DTMC model to calculate the lockup premium.

The DTMC model in DPW09 is similar in spirit to previous DTMC models used in finance, such as the DTMC model of yearly bond credit rating migration given on pp. 626-627 of Hull

(2003). In bond credit rating, there are easily identified states, namely, the different credit ratings, ranging from *AAA* to *CCC* and default. With these well-specified states, the transition probabilities are easy to estimate from the observed proportions of changes in historical data.

The DTMC model for hedge funds in DPW09 is less straightforward. As with bond ratings, the DTMC was used to model the “state” of the hedge fund, but the state is not so easy to define. However, the rules for hedge fund lockup suggest a simple framework: In DPW09, three states were postulated: good, sick, and dead. In a good state, the fund has above-average performance, so an investor wants to keep his investment in the fund. In a sick state, the fund shows below-average performance, so an investor in this fund would want to redeem his investment and reinvest in another hedge fund in a good state, if allowed. In a dead state, the investor suffers a low return due to poor performance, and the fund becomes extinct. However, in a dead state, the lockup condition becomes invalid and the investor receives the remaining balance from the dead fund. In a dead state, the investor can immediately invest in another fund, which we take to be in the good state. (The model can incorporate partial redemption of the investment from a dead fund by assigning an appropriate return value for a dead state.)

From the perspective of the lockup premium, the critical state is the sick state. There is no extra lockup penalty associated with a good fund or a dead fund, but there is with a sick fund. With the nominal one-year lockup, we assume that an investor will reinvest in a good fund every year, if the current fund is not judged to be in a good state. In contrast with an extended lockup period, the investor will not be able to reinvest when the fund is judged to be in a sick state. The investor must keep his investment in the sick fund. Meanwhile, the state of the sick fund will evolve in an uncertain manner. It may continue to produce mediocre returns and be judged sick, it may get worse and “die,” or it may recover and become a good fund. The DTMC model was used to capture the likelihood of the different alternatives.

Of course, the state (“health”) of a hedge fund cannot be directly observed, but symptoms are observed, notably the stream of returns. In DPW the state was estimated by the level of relative returns. Two thresholds were postulated:  $U$  and  $L$ , with  $U > L$ . A fund is judged to be in a good state if the relative returns are above  $U$ , in a sick state if the relative returns are in the interval  $(L, U)$ , and in a dead state if the relative returns fall below  $L$ . Assuming that the three-state DTMC only can have one-step (one-year) transitions to a neighboring state, the DTMC has three parameters; the three transition probabilities  $P_{G,S}$ ,  $P_{S,G}$  and  $P_{S,D}$ . Since the true state is not observable, an indirect procedure was used to estimate the

transition probabilities. To fit the parameters in the DTMC model based on the TASS hedge fund data, three important hedge-fund performance measures were used: the performance persistence factor, the death rate and the standard deviation of annual returns. The transition probabilities were determined by solving a system of equations.

In DPW09 we relied heavily on the persistence of hedge fund returns. A persistence level  $\gamma$  means that *“for every 1 percentage point earned above the average in the current year, we expect to earn  $\gamma$  percentage points above the average in the next year.”* We acknowledge that the existence of persistence in hedge funds is controversial, but we found strong evidence based on the TASS hedge fund data. In DPW09 we estimated the persistence by doing a regression analysis on the hedge fund return data from the TASS database. Zero persistence is contained in the 95% confidence intervals for only three of the eleven hedge fund strategy categories; see Table 1 of DPW09. (The quality of the TASS data remains a concern, however.)

Given the fitted DTMC model, the lockup premium was then calculated as the compensation for the restricted rebalancing opportunities from a sick state fund to a good state during the lockup period. Specifically, the calculation is done by comparing the expected return of the same hedge fund with and without the extended lockup condition. The lockup premium is the annual fixed rate of return that compensates the difference of expected returns in the two hedge-fund investments, with and without the extended lockup condition.

## 1.2. The Proposed Continuous-Time Model

Unfortunately, the DTMC model with yearly transitions has the limitation that the state of a hedge fund can only change once per year. Since transition was restricted to neighboring states, it requires at least two years for a fund to transition from good to sick and then sick to dead. In reality, the fortune of hedge funds can change much more rapidly. Thus we are motivated to consider alternative models that allow the hedge fund state to change continuously in time.

For that purpose, in this paper we propose to model the state of the hedge fund as an absorbing continuous-time Markov chain (CTMC) or, more specifically, an absorbing birth-and-death process. With the CTMC model, there are again three parameters, but now the parameters are the instantaneous transition rates to neighboring states: There are death rates  $\mu_G$  and  $\mu_S$  for transitioning from good to sick and from sick to dead, respectively. And there is a birth rate  $\lambda_S$  for transitioning from sick to good. The fund ceases to exist in a dead state, so the dead state is an absorbing state.

With the CTMC model, we still assume that the fund generates returns in discrete time. We

use discrete time because the returns are reported infrequently. We assume that the investor will be able to redeem his investment from a dead fund at these reporting times, whenever the fund becomes dead. Thus, from the investor's perspective, there are only two relevant states at these reporting times: good and sick. Thus, the successive states at the return times becomes a two-state ergodic DTMC, whose transition probabilities are determined by the transient probabilities of the CTMC. We thus need to use the transient transition probabilities of the absorbing CTMC in order to determine the one-step transition probabilities of the ergodic DTMC. Given the ergodic DTMC, we can calculate the lockup premium, much as before; see §3.

Allowing state changes continuously in time is more realistic and allows greater flexibility in the model fitting. However, new calibration methods are needed. We use the same measures of hedge fund performance. Just as in DPW09 and Derman et al. (2010), we rely heavily on the persistence of hedge fund returns. It turns out that we can again fit the three model parameters (now the three transition rates of the CTMC) to the persistence, the death probability and the annual return volatility. However, the fitting becomes more complicated. We are no longer able to obtain closed-form expressions for the parameters (the transition rates of the CTMC, as opposed to the transition probabilities of the DTMC in DPW09). Instead, we exploit nonlinear programming and an iterative algorithm to carry out the fitting. However, once this algorithm has been developed, we are able to fit the parameters as easily and rapidly as before. Moreover, the flexibility provided by the CTMC allows us to fit to a wider range of parameters. For example, with the previous DTMC model, the death rate had to be less than 0.06; with the CTMC model, it can be as high as 0.13.

Overall, we show that the more realistic and flexible CTMC model is also a viable alternative to estimate the premium for extended lockup of hedge funds. The estimated lockup premium is thus presumably more accurate. However, in our numerical experiments we found that the computed lockup premium as a function of hedge fund performance measures does not differ greatly from the premium estimated by the DTMC model. We thus regard this paper as confirming that the previous more rudimentary DTMC model yields a reasonable approximation. Both highly stylized approaches seem appropriate given the quality of the data. The methods here may also be useful to guide the development of more sophisticated methods in the future with more refined data.

### 1.3. The Rest of the Paper

The rest of the paper is organized as follows: We start in §2 by giving a literature review. Then in §3 we carefully define the lockup premium. The new Markov chain models are then defined in §4. In §5 we develop our new algorithms for fitting the parameters. We indicate how to calculate the lockup premium in §6. In §7 we perform sensitivity analysis, showing how the model parameters and the lockup premium depend on basic hedge fund performance measures. Finally, in §8 we draw conclusions.

## 2. Literature Review

We will be brief in our literature review, because an extensive review was given in DPW09. We will emphasize recent contributions since DPW09.

First, it is good to put the definition of lockup premium in perspective with the broader literature on liquidity. Clearly, an investor under the lockup condition naturally should require compensation or additional return from his investment in hedge funds for limited rebalancing activities during the lockup period, which we define as the lockup premium. Our definition of the lockup premium corresponds to the liquidity premium in the asset pricing literature; see Longstaff (1995), Longstaff (2001), and Browne et al. (2003). As stated in §1 of this paper and the above papers, the calculation of the premium can be done by comparing the expected returns in the two investments, with and without the liquidity condition. This approach of the liquidity premium calculation is also relevant for accounting regulation (FAS No. 157, “Fair Value Measurements”) to discount illiquidity in a portfolio (Ang and Bollen, 2009).

Similar to DPW09, Ang and Bollen (2009) propose a variant of a binomial lattice model to calculate the lockup premium as a function of the lockup period and notice period. They estimate a two-year lockup with a three-month notice period is approximately 1%, which is also similar to our result. However, while their model allows a time-dependent default probability, their model assumes independence of increment returns. We think that independent increments may not hold, especially for an illiquid investment like a hedge fund. Our view is supported by Jagannathan et al. (2006), Koh et al. (2003), Agarwal and Naik (2000), Edwards and Caglayan (2001), and DPW09, who all provide support the existence of persistence in hedge-fund returns in various ways. See, the discussion and the references in DPW09 and Derman et al. (2010) for the details. Recent studies like Boyson et al. (2010) which study downside correlations in hedge funds also support our view.

There also are a few purely empirical studies, without employing stochastic models. First, Aragon (2007) empirically compares the performance of hedge funds with and without the one-year lockup condition. The average difference between returns of the hedge funds with and without the lockup condition is estimated by 4 to 7% in Aragon (2007). However, his study does not differentiate various lockup periods in the data and regards all different lockup periods as one year. Thus the estimated lockup premium is not a function of the lockup period. De Roon et al. (2009) also estimate the lockup premium empirically where an investor has a portfolio of stocks, bonds, and hedge funds. They estimate the three-month lockup costs the investor 4.11% annually. However, just as Aragon (2007), they did not calculate the lockup premium for other lockup periods than three months.

### 3. Definition of the Lockup Premium

In this section, we carefully define the premium for extended hedge-fund lockups that we will use with our new Markov chain models. Since the hedge fund returns in the database are reported monthly at most, it is reasonable to consider a discrete process. We thus consider a discrete-time return stochastic process  $B \equiv \{B_i : i = 1, 2, 3, \dots\}$ . Each time period  $i$  in the process represents an *updating time* for the returns. We can study the model with different updating time period  $T_u$  for updates, if we wish. For example,  $T_u = 0.25$  implies quarterly updating of the process  $B$ , while  $T_u = 0.5$  represents semi-annual updating of returns. We assume that the number of updates in a year,  $k \equiv 1/T_u$ , is always an integer, which usually is the case in practice. While  $T_u = 1$  in the DTMC model,  $1/T_u$  can be any integer in the CTMC model.

We let  $B_i$  represent a continuously compounded (random) rate of return for the  $i^{th}$  updating period, by which we mean that  $e^{B_i}$  is the (random) value at the end of  $i^{th}$  updating period ( $i \cdot T_u$  year) of one dollar invested in this investment at the beginning of  $i - 1^{th}$  updating period ( $(i - 1) \cdot T_u$  year). (Notice that  $B_i$  depends on  $T_u$  in this definition. For example,  $B_i$  is an annually-updated return process when  $T_u = 1$ , whereas  $B_i$  is a semi-annually updated return process when  $T_u = 0.5$ .)

Consequently, the (random) total value at the end of  $n$  years of one dollar invested in this investment at the beginning of the first year,  $V_{n, T_u}$ , is the  $(n \cdot k)$ -fold product

$$V_{n, T_u} = e^{\sum_{i=1}^{n \cdot k} B_i} , \quad (3.1)$$

where  $k$  is the number of updating (reporting) periods per year and  $n$  is the number of years.



We let  $r_n$  be the deterministic value for which

$$e^{nr_n} = \mathbb{E}[V_{n,T_u}] \quad (3.2)$$

for  $V_{n,T_u}$  in (3.1); i.e., we let  $r_n$  be the constant rate of return, with continuous compounding, that yields the same expected value  $\mathbb{E}[V_n]$  over  $n$  years. We call  $r_n$  simply the *rate of return* of this investment. What we have done follows common practice. We have “backed out” the rate of return  $r_n$  from the expected cash value  $\mathbb{E}[V_n]$ . By (3.1) and (3.2),  $r_n$  can be expressed directly as

$$r_n = \frac{\log \mathbb{E}[V_n]}{n} = \frac{1}{n} \log \left( \mathbb{E} \left[ e^{(\sum_{i=1}^{n-k} B_i)} \right] \right), \quad (3.3)$$

where we use the natural logarithm (base  $e$ ).

Now consider two different investment opportunities in hedge funds with the same strategy, but one with a conventional 1-year lockup and the other with an  $n$ -year lockup condition. Let  $B^1$  be the return stochastic processes with the 1-year lockup condition; let  $B^2$  be the return stochastic processes with the  $n$ -year lockup condition. Let the  $n$ -year *lockup premium*  $p_n$  be

$$p_n \equiv r_n^1 - r_n^2, \quad (3.4)$$

where  $r_n^i$  is the rate of return of  $B^i$ , defined as in (3.3), over an  $n$ -year horizon.

Instead of working with  $B_i$  directly, we focus on *relative return rates* for each fund strategy. To do so, we let  $\alpha_i \equiv \mathbb{E}[B_i]$ , the mean return rate for a particular hedge fund strategy within updating period of return  $i$ . This quantity is estimated by the mean of  $B_i$  over all funds within that strategy. Then the (random) relative return rate is

$$R_i \equiv B_i - \mathbb{E}[B_i] \equiv B_i - \alpha_i. \quad (3.5)$$

Combining equations (3.1) and (3.5), we see that the (random) total value at the end of year  $n$  from investor  $j$  is

$$V_{n,T_u}^j = \prod_{i=1}^{n-k} e^{(\alpha_i + R_i)} = e^{(\sum_{i=1}^{n-k} \alpha_i)} e^{(\sum_{i=1}^{n-k} R_i^j)}. \quad (3.6)$$

and the difference between the expected total returns is

$$\mathbb{E}[V_{n,T_u}^1] - \mathbb{E}[V_{n,T_u}^2] = e^{(\sum_{i=1}^{n-k} \alpha_i)} \left( \mathbb{E} \left[ e^{(\sum_{i=1}^{n-k} R_i^1)} \right] - \mathbb{E} \left[ e^{(\sum_{i=1}^{n-k} R_i^2)} \right] \right). \quad (3.7)$$

Hence, the premium in (3.4) becomes

$$p_n \equiv r_{n,T_u}^1 - r_{n,T_u}^2 = \frac{1}{n} \left[ \log \left( \mathbb{E} \left[ e^{(\sum_{i=1}^{n-k} R_i^1)} \right] \right) - \log \left( \mathbb{E} \left[ e^{(\sum_{i=1}^{n-k} R_i^2)} \right] \right) \right], \quad (3.8)$$

which is independent of the average rates  $\alpha_i$ . We need a model that describe the behavior of  $R_i^j$  in (3.8).

## 4. The New Markov Chain Models

In this section, we propose a model that uses an absorbing CTMC to model the evolution of the fund state in time and an ergodic DTMC to model the state of the fund at the updating times (when returns are reported). (Since there is a positive probability of transition from good to sick in finite time and a positive probability of transition from sick to dead in finite time, the absorbing CTMC would eventually end up in the dead state with probability 1 (over an infinitely long horizon), whereas the ergodic DTMC has a proper limiting steady-state distribution.) In the CTMC model, a fund changes its state in continuous time, but the investment updates take place in discrete time. This leads us to an ergodic DTMC for investment update that is based on the CTMC; see Ross (2003) for background on both kinds of Markov chains.

### 4.1. The Transition Matrix of the Absorbing CTMC

Just as in DPW09, our CTMC has three states for a hedge fund: good, sick and dead. Now we assume that transitions among these states occur according to a CTMC, specifically a birth-and-death process.

In the DTMC model in DPW09, a hedge fund in a good state cannot reach a dead state until two years. In contrast, with the CTMC model, a fund can be in a dead within one year.

There is a cost, however: for the CTMC model, we are unable to fit the model parameters simply by explicitly solving three equations in three unknowns. Instead, we develop an algorithm to carry out the model fitting numerically. With our algorithm, the parameter fitting for the CTMC model is not substantially harder than for the DTMC model.

In our proposed CTMC model we replace the three-state absorbing DTMC in DPW09 by a two-state absorbing CTMC. The states now are  $G$  (Good) and  $S$  (Sick); we do not directly use the state  $D$  (Dead) here, but we will be able to account for it. As usual, we specify the CTMC by specifying its infinitesimal transition rate matrix  $Q$ . That means we specify the birth and death rates. Let  $\mu_G$  be the death rate in  $G$ , the rate of transition down to state  $S$  from state  $G$ . Let  $\lambda_S$  be the birth rate in state  $S$ , the rate of transition up to state  $G$  from state  $S$ . Let  $\mu_S$  be the death rate in state  $S$ , implicitly the rate of transition down to state  $D$  from state  $S$ . (The fund may leave state  $S$  to go to state  $D$ , but gets absorbed in  $D$ . We do not need to include the state  $D$  in our transition rate matrix.) Here is the infinitesimal transition matrix,

with the parameters above:

$$Q = \begin{matrix} G \\ S \end{matrix} \begin{pmatrix} -\mu_G & \mu_G \\ \lambda_S & -(\lambda_S + \mu_S) \end{pmatrix}. \quad (4.9)$$

We now want to derive the time-dependent transition probability matrix  $P(t)$  for this CTMC. It is well-known that  $P(t)$  is the solution to the matrix ordinary differential equation

$$P(t)' = P(t)Q, \quad P(0) = I, \quad (4.10)$$

where  $I$  is the identity matrix, so that  $P(t)$  is the matrix exponential  $P(t) = e^{tQ}$ . If we diagonalize  $Q$  so that  $Q = UDU^{-1}$ , where  $D$  is a diagonal matrix and  $UU^{-1} = I$ , then we can write  $P(t) = Ue^{tD}U^{-1}$ ; see §4.8 and the appendix of Karlin and Taylor (1975). Since  $D$  is a diagonal matrix, the  $i^{\text{th}}$  diagonal element of  $e^{tD}$  is related to the corresponding diagonal element of  $D$ , i.e.,  $(e^{tD})_{i,i} = e^{D_{i,i}t}$  for  $t > 0$ . Let  $\Lambda(t)$  be a diagonal matrix of the form

$$\Lambda(t) = \begin{matrix} G \\ S \end{matrix} \begin{pmatrix} e^{\eta_G t} & 0 \\ 0 & e^{\eta_S t} \end{pmatrix}, \quad (4.11)$$

with the two parameters  $\eta_G$  and  $\eta_S$  being the eigenvalues of the matrix  $Q$ , while the columns of  $U$  are the associated right eigenvectors. The resulting formula for  $P(t)$  is

$$P(t) = U\Lambda(t)U^{-1}. \quad (4.12)$$

The characterization (4.12) implies that  $P_{i,j}(t) = A_{i,j}e^{\eta_1 t} + B_{i,j}e^{\eta_2 t}$  for  $t \geq 0$  and all state pairs  $(i, j)$ , where  $\eta_1$  and  $\eta_2$  are the eigenvalues of  $Q$  and  $A_{i,j}$  and  $B_{i,j}$  are appropriate constants. Since  $P(0) = I$ , we necessarily have  $A_{i,i} + B_{i,i} = 1$  for  $i = 1, 2$  and  $A_{i,j} + B_{i,j} = 0$  for  $i \neq j$ . If  $0 > \eta_1 > \eta_2$ , then asymptotically  $P_{i,j}(t) \sim A_{i,j}e^{-\eta_1 t}$  as  $t \rightarrow \infty$ , which means that the ratio approaches 1. As a consequence, necessarily  $A_{i,j} > 0$  for all state pairs  $(i, j)$ ;  $B_{i,j} = -A_{i,j}$  for  $i \neq j$ .

As usual, we find the eigenvalues of  $Q$  by finding the determinant of  $\eta I - Q$ . The characteristic polynomial as a function of the variable  $\eta$  is the quadratic equation

$$(\eta + \lambda_S + \mu_S)(\eta + \mu_G) - \lambda_S \mu_G = 0, \quad (4.13)$$

which has two strictly negative roots, as required for the formula in (4.11) to yield bonafide probabilities. In particular, solving the quadratic equation, we obtain

$$\eta = \frac{-(\lambda_S + \mu_S + \mu_G) \pm \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S \mu_G}}{2}. \quad (4.14)$$

Since the term inside the square root can be rewritten as  $(\mu_G - \mu_S)^2 + \lambda_S^2 + 2\mu_G \lambda_S + 2\lambda_S \mu_S$ , it is nonnegative. The first term clearly dominates the square root in absolute value. So we indeed have two negative roots.

Now we find eigenvectors corresponding to the eigenvalues in (4.14). Given eigenvalues, the eigenvectors form the null space of  $(Q - \eta I)$ , i.e., a matrix  $U$  such that  $(Q - \eta I)U = 0$ . We arrange eigenvalues  $\eta_G, \eta_S$  as  $\eta$  matrix:

$$\eta = \begin{pmatrix} \eta_G \\ \eta_S \end{pmatrix} = \begin{pmatrix} \frac{-(\lambda_S + \mu_S + \mu_G) - \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}}{2} \\ \frac{-(\lambda_S + \mu_S + \mu_G) + \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}}{2} \end{pmatrix}. \quad (4.15)$$

Such an eigenvector matrix  $U$ , where the columns of  $U$  are eigenvectors of  $Q$ , can be easily found by algebraic manipulation or by a symbolic calculation package such as Mathematica. One such eigenvalue matrix is

$$U = \begin{pmatrix} \frac{(\lambda_S + \mu_S - \mu_G) - \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}}{2\lambda_S} & \frac{(\lambda_S + \mu_S - \mu_G) + \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}}{2\lambda_S} \\ 1 & 1 \end{pmatrix}. \quad (4.16)$$

Its inverse matrix is then

$$U^{-1} = \begin{pmatrix} -\frac{\lambda_S}{\sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}} & \frac{\lambda_S + \mu_S - \mu_G + \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}}{2\sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}} \\ \frac{\lambda_S}{\sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}} & \frac{-\lambda_S - \mu_S + \mu_G + \sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}}{2\sqrt{(\lambda_S + \mu_S + \mu_G)^2 - 4\mu_S\mu_G}} \end{pmatrix}. \quad (4.17)$$

Thus, we now have derived the components of  $P(t)$  in (4.12). We have derived  $P(t)$  as a nonlinear function of  $\mu_G$ ,  $\lambda_S$  and  $\mu_S$  from (4.15)-(4.17).

## 4.2. The Associated Ergodic DTMC

We are now ready to specify the evolution of the hedge fund at successive updating times. We will characterize these transitions by an ergodic two-state DTMC, whose transition probabilities are obtained from the time-dependent transition probabilities of the absorbing CTMC developed above.

Since a dead fund is replaced immediately by a good fund at the updating time, we make an ergodic two-state DTMC with transition matrix

$$P = \begin{matrix} G \\ S \end{matrix} \begin{pmatrix} 1 - P_{G,S}(T_u) & P_{G,S}(T_u) \\ 1 - P_{S,S}(T_u) & P_{S,S}(T_u) \end{pmatrix}. \quad (4.18)$$

We construct  $P$  in (4.18) by letting  $P_{G,S} = P_{G,S}(T_u)$  and  $P_{S,S} = P_{S,S}(T_u)$  and then making the DTMC ergodic by letting the row sums be 1. This procedure is tantamount to inserting an instantaneous transition from state  $D$  to  $G$  at time  $T_u$ , which is the time of a single transition in the DTMC. This transition probabilities are used to model the process  $R_i^j$  in (3.8).

As usual, the steady-state vector for this DTMC is obtained by solving the equation  $\pi = \pi P$  for a probability vector  $\pi \equiv (\pi_G, \pi_S)$ . In this simple  $2 \times 2$  case, we can give the steady-state probability vector  $\pi$  explicitly as

$$\pi \equiv (\pi_G, \pi_S) = \left( \frac{1 - P_{S,S}(T_u)}{1 - P_{S,S}(T_u) + P_{G,S}(T_u)}, \frac{P_{G,S}(T_u)}{1 - P_{S,S}(T_u) + P_{G,S}(T_u)} \right). \quad (4.19)$$

## 5. Parameter Fitting in the CTMC Model

Just as we fit the three parameters  $p \equiv P_{G,G}$ ,  $q \equiv P_{S,G}$  and  $r \equiv P_{S,S}$  in the DTMC transition matrix to performance measures estimated from the TASS data in DPW09, here instead we fit the three parameters  $\lambda_S$ ,  $\mu_S$ , and  $\mu_G$  in the CTMC transition rate matrix (4.9) to the three hedge-fund performance measures.

### 5.1. The Persistence and Death Fitting Equations

The first performance measure for model fitting is the persistence factor of relative returns. As indicated in the introduction, a persistence level  $\gamma$  means that for every 1 percentage point earned above the average in the current year, we expect to earn  $\gamma$  percentage points above the average in the next year. We define a time period  $T_p$  to measure the persistence factor. For example,  $T_p = 1$  implies that the persistence factor is measured with annual relative returns whereas  $T_p = 0.5$  means that it is measured with semi-annual relative returns. Papers on persistence of hedge-fund returns use different measurement times ranging from one quarter to three years (See, §4 of DPW09 for the details). Thus, allowing variable  $T_p$  could be useful if one has different view on measurement time for persistence.

The time periods  $T_u$  and  $T_p$  can be different in our model, while they were both 1 in the DTMC model.

We let  $Y_G$ ,  $Y_S$ , and  $Y_D$  represent the annual relative returns that an investor makes at each updating time from a good, sick, and dead fund, respectively. As before, we will take these as given parameters, based on the data, but we will also discuss how to get reasonable estimates of these values below.

We now proceed toward parameter fitting for this new model. We first define an equation that fits observed persistence within a specific strategy. As in DPW09, we allow different persistence factors in the states  $G$  and  $S$ . Paralleling (6.5)-(6.6) in DPW09, we obtain the new equations

$$\gamma_G \cdot Y_G = P_{G,G}(T_p) \cdot Y_G + P_{G,S}(T_p) \cdot Y_S + \{1 - P_{G,G}(T_p) - P_{G,S}(T_p)\}Y_D \quad (5.20)$$

and

$$\gamma_S \cdot Y_S = P_{S,G}(T_p) \cdot Y_G + P_{S,S}(T_p) \cdot Y_S + \{1 - P_{S,G}(T_p) - P_{S,S}(T_p)\}Y_D . \quad (5.21)$$

Both sides of the equations above represent expected relative returns, starting with a good or sick state. Both exploit the definition of persistence.

Next, paralleling equation (6.4) in DPW09, we also have an equation based on the probability of the fund ceasing to exist (dying) within a specified time. Specifically, we introduce an equation giving the probability of death within the time period  $T_d$ , where  $T_d$  may differ from  $T_u$  and  $T_p$ ; i.e.,

$$\delta \equiv \delta(T_d) = \pi_G\{1 - P_{G,G}(T_d) - P_{G,S}(T_d)\} + \pi_S\{1 - P_{S,G}(T_d) - P_{S,S}(T_d)\} . \quad (5.22)$$

Notice that the new model is more flexible because the three times  $T_u, T_p$ , and  $T_d$  can be different.

With equation (5.22), we can derive the survival probability from the CTMC model, which is closely related to the death probability. At time  $t$ , the *survival probability* of a fund is defined as

$$S(t) = P_{G,G}(t) + P_{G,S}(t) \quad \text{for } t \geq 0. \quad (5.23)$$

Figure 1 displays the survival probabilities for the CTMC model with  $\delta = 0.03, 0.06$  and  $0.09$  (and other parameter values from Table 1 below). The survival probability for  $\delta = 0.09$  is possible only in the CTMC model, since for the DTMC model,  $r$  becomes negative when  $\delta \approx 0.07$ . As we see from Figure 1, median fund life is less than 10 years for  $\delta = 0.09$ . Since this median hedge fund life is within the range of Rouah (2006) and Park (2007), it should be worth considering  $\delta = 0.09$ . We could not do this before in DPW09.

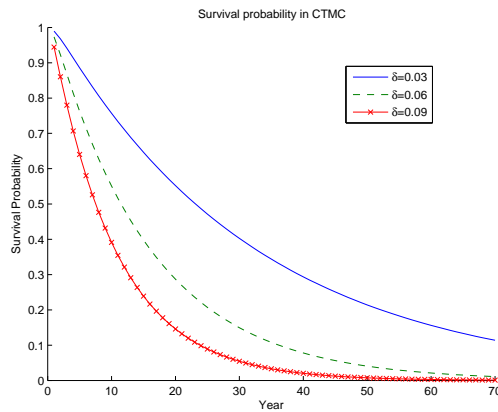


Figure 1: Survival probability curves for the CTMC model with  $\delta = 0.03, 0.06$ , and  $0.09$  with  $T_u = T_p = T_d = 1$ . The parameters values  $\mu_G, \lambda_S, \mu_S, Y_G, Y_S, Y_D$ , and  $\sigma$  are from Table 1.

## 5.2. Solving the Three Equations in Three Unknowns

We now show how to solve the three equations (5.20), (5.21) and (5.22) for the three unknowns  $\mu_G, \mu_S$  and  $\lambda_S$ . Unfortunately, we have been unable to obtain explicit solutions for the desired parameters as we did in DPW09. Hence, we develop an effective numerical algorithm.

Suppose that we start with a candidate initial parameter triple  $(\mu_G, \lambda_S, \mu_S)$ . Given that parameter triple and the specified time  $t$ , we calculate the transition probabilities  $P_{G,G}(t)$ ,  $P_{G,S}(t)$ ,  $P_{S,G}(t)$ , and  $P_{S,S}(t)$  in (4.12)–(4.14) by calculating the eigenvalues and eigenvectors of the infinitesimal matrix  $Q$  in (4.9). Afterwards we calculate the steady-state probability vector  $\pi \equiv (\pi_G, \pi_S)$  in (4.19) of the two-state DTMC in (4.18). We then calculate the right-hand sides of the three equations (5.20)–(5.22). Our goal is to have three bonafide equations, where the two sides of the equations are equal, but in the iteration we will not achieve that. Based on the errors we see, we update the parameter triple  $(\mu_G, \lambda_S, \mu_S)$  and repeat until the errors in the three equations (5.20)–(5.22) are negligible. Notice that  $T_u, T_d$ , and  $T_p$  are all constant so does not add any complexity to the system of equations.

Since we are confronted with a three-dimensional iteration, we do not want to proceed in a haphazard way. Hence, we apply nonlinear programming to do this iteration. The idea is to find parameter triple  $(\lambda_S, \mu_S, \mu_G)$  minimizing errors between the right-hand and left-hand sides of equations (5.20), (5.21) and (5.22). To formulate a minimization problem, we define three error functions  $\epsilon_1$ ,  $\epsilon_2$  and  $\epsilon_3$  as a function of parameter triple  $(\lambda_S, \mu_S, \mu_G)$  as follows:

$$\begin{aligned}\epsilon_1 &\equiv \epsilon_1(\lambda_S, \mu_S, \mu_G) = \delta(T_u) - \pi_G\{1 - P_{G,G}(T_d) - P_{G,S}(T_d)\} - \pi_S\{1 - P_{S,G}(T_d) - P_{S,S}(T_d)\}, \\ \epsilon_2 &\equiv \epsilon_2(\lambda_S, \mu_S, \mu_G) = \gamma_G \cdot Y_G - P_{G,G}(T_p) \cdot Y_G - P_{G,S}(T_p) \cdot Y_S - \{1 - P_{G,G}(T_p) - P_{G,S}(T_p)\}Y_D, \\ \epsilon_3 &\equiv \epsilon_3(\lambda_S, \mu_S, \mu_G) = \gamma_S \cdot Y_S - P_{S,G}(T_p) \cdot Y_G - P_{S,S}(T_p) \cdot Y_S - \{1 - P_{S,G}(T_p) - P_{S,S}(T_p)\}Y_D.\end{aligned}\tag{5.24}$$

Our objective, then, is to find  $\lambda_S$ ,  $\mu_S$  and  $\mu_G$  such that  $\epsilon_1(\lambda_S, \mu_S, \mu_G) = \epsilon_2(\lambda_S, \mu_S, \mu_G) = \epsilon_3(\lambda_S, \mu_S, \mu_G) = 0$ . To obtain values of  $\epsilon_1, \epsilon_2$ , and  $\epsilon_3$  for a given parameter triple of  $\lambda_S, \mu_S$  and  $\mu_G$ , we have to calculate  $P_{G,G}(T), P_{G,S}(T), P_{S,G}(T)$ , which are elements of  $P(t)$  matrix in (4.12). As indicated above, this involves finding eigenvalues and eigenvectors of  $Q$  matrix in (4.9). From (4.14), we derived eigenvalues as a function of  $\lambda_S, \mu_S$  and  $\mu_G$ . Given the eigenvalues, the eigenvectors can be calculated as in (4.16), but also in other ways. Since  $Q$  is only a  $2 \times 2$  matrix, calculation of the eigenvectors for given eigenvalues can be done easily. One way is to use the Schur decomposition algorithm, as in Anderson et al. (1999), which is implemented in MATLAB as the *eig* function. Then  $\Lambda(t)$  can be calculated easily from (4.11), so we can easily compute the  $U$  and  $\Lambda$  matrices numerically. The final step is to compute  $P_{G,G}(t), P_{G,S}(t)$  and  $P_{S,G}(t)$  from  $P(t) = U\Lambda(t)U^{-1}$ .

We can obtain the desired parameter triple  $(\lambda_S, \mu_S, \mu_G)$  by solving the following constrained

minimization problem:

$$\begin{aligned}
& \min_{\lambda_S, \mu_S, \mu_G} \max\{|\epsilon_1|, |\epsilon_2|, |\epsilon_3|\} \\
& s. t. \\
& \epsilon_1 = \delta(T_u) - \pi_G(1 - P_{G,G}(T_u) - P_{G,S}(T_u)) - \pi_S\{(1 - P_{S,G}(T_u) - P_{S,S}(T_u))\}, \\
& \epsilon_2 = \gamma_G \cdot Y_G - P_{G,G}(T_p) \cdot Y_G - P_{G,S}(T_p) \cdot Y_S - \{1 - P_{G,G}(T_p) - P_{G,S}(T_p)\}Y_D, \\
& \epsilon_3 = \gamma_S \cdot Y_S - P_{S,G}(T_p) \cdot Y_G - P_{S,S}(T_p) \cdot Y_S - \{1 - P_{S,G}(T_p) - P_{S,S}(T_p)\}Y_D \\
& \lambda_S, \mu_S, \mu_G \geq 0
\end{aligned} \tag{5.25}$$

With (5.25), we regard  $\lambda_S$ ,  $\mu_S$  and  $\mu_G$  as the variables. Since the transition probabilities  $P_{G,G}(t)$ ,  $P_{G,S}(t)$  and  $P_{S,G}(t)$  are functions of the BD rates  $\lambda_S$ ,  $\mu_S$  and  $\mu_G$  through the eigenvalue and eigenvector calculation, we must regard (5.25) as a nonlinear programming (NLP) problem, for which it is natural to apply an iterative procedure. However, since we only have three variables, we are able to solve the NLP (5.25) easily. One effective way is to use Sequential Quadratic Programming (SQP), as in Schittkowski (1986). With SQP, at each iteration, an approximation is made of the Hessian of the Lagrangian function using a quasi-Newton updating method. That is then used to generate a QP subproblem whose solution is used to form a search direction for a line search procedure. This algorithm is implemented in MATLAB via the functions *fminsearch* and *fmincon*. Both functions solve (5.25) within seconds.

### 5.3. The Variance and the Return of a Good Fund

So far, we have assumed that the three returns  $Y_G$ ,  $Y_S$  and  $Y_D$  in the three states can be specified in advance. We can specify one of these, which we take to be  $Y_G$ , by exploiting an equation for the variance of the relative annual returns in steady state. Since we are focusing on the relative returns, the variance  $\sigma^2$  coincides with the second moment.

To express the steady state variance  $\sigma^2$ , we need the steady-state probabilities. Given that  $\pi_G$  and  $\pi_S$  are the stationary probabilities for the transition matrix in (4.18), the stationary probability that the fund dies at the end of  $T_u$  years when it starts alive is

$$\pi'_D = \pi_G [-P_{G,G}(T_u) - P_{G,S}(T_u)] + \pi_S [1 - P_{S,G}(T_u) - P_{S,S}(T_u)] \tag{5.26}$$

We then also define  $\pi'_G$  and  $\pi'_S$  accordingly, using (4.19):

$$\begin{aligned}
\pi'_G &= \pi_G \cdot P_{G,G}(T_u) + \pi_S \cdot P_{S,G}(T_u) \\
\pi'_S &= \pi_G \cdot P_{G,S}(T_u) + \pi_S \cdot P_{S,S}(T_u)
\end{aligned} \tag{5.27}$$

where  $(\pi_G, \pi_S)$  is defined in (4.19).



We obtain an expression for the steady-state variance of returns by supposing that the returns in good, sick and dead states are the deterministic values  $Y_G$ ,  $Y_S$  and  $Y_D$ , respectively. Thus, the variance satisfies the equation

$$\sigma^2 = \pi'_G \cdot Y_G^2 + \pi'_S \cdot Y_S^2 + \pi'_D \cdot Y_D^2, \quad (5.28)$$

where  $\pi'$  is defined in (5.26) and (5.27).

We applied equation (5.28) to determine  $Y_G$  for given  $Y_S \equiv -1.5\sigma$  and  $Y_D \equiv -2.0\sigma$ , just as DPW09. To do so, we simply iterate on the two parameters  $Y_G$  and  $\sigma^2$ , using a simple bisection search. As in Figure 7 of DPW09, we found that there is a nearly linear positive relation between  $Y_G$  and  $\sigma^2$  for given  $Y_S$  and  $Y_D$ . Thus, we are able to easily achieve any desired  $\sigma$ , such as  $\sigma \approx 0.1$ , by iterating  $Y_G$ .

While we can use judgement to specify the values  $Y_S$  and  $Y_D$ , to obtain the returns  $Y_S$  and  $Y_D$  in the sick and dead states for simplicity, as we have more reliable and accurate hedge-fund returns, we note that we can also classify relative returns by introducing two thresholds  $U$  and  $L$ . We can classify the fund as good, sick or dead if the relative return falls in the interval  $[U, \infty)$ ,  $(L, U)$  or  $(-\infty, L]$ , respectively. With this classification, we can let  $Y_S$  and  $Y_D$  be the average returns in the sick and dead intervals. The fitting algorithm provides an estimate of the parameters  $\mu_G$ ,  $\mu_S$ ,  $\lambda_S$  and  $Y_G$  as a function of the time parameters  $T_p$ ,  $T_u$  and  $T_d$ , the persistence parameters  $\gamma_G$  and  $\gamma_S$ , the death parameter  $\delta$  and the relative return parameters  $Y_S$  and  $Y_D$ . As we will illustrate, this provides an opportunity to do sensitivity analysis.

#### 5.4. Application of the Algorithm

Below are parameter values obtained using the NLP in (5.25) and iterating  $Y_G$  values. In the following table,  $\epsilon$  records the maximum absolute value of errors in equations (5.20), (5.21) and (5.22), obtained where we elected to terminate the algorithm. We let  $T_p = T_u = 1$  for simplicity but these may have any values, depending on the time measure from the data.

Table 1: Parameter values for the CTMC model as a function of  $\delta$  with  $Y_S = -0.15$ ,  $Y_D = -0.20$ ,  $\sigma = 0.1$ ,  $\gamma_G = \gamma_S = 0.5$ ,  $T_p = T_u = T_d = 1$

$\delta$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S$	$Y_D$	$\sigma$	$\epsilon$
0.00	0.2133	0.4798	0.0000	0.0667	-0.15	-0.20	0.1000	$2.2626 \times 10^{-7}$
0.03	0.2191	0.5533	0.1250	0.0684	-0.15	-0.20	0.1000	$2.8191 \times 10^{-7}$
0.06	0.2262	0.7025	0.3495	0.0699	-0.15	-0.20	0.1000	$3.1488 \times 10^{-7}$
0.09	0.2386	1.0980	0.8726	0.0713	-0.15	-0.20	0.1000	$5.3025 \times 10^{-6}$

Unlike the DTMC model, where the parameter  $r$  becomes negative if  $\delta$  exceeds 0.07 for the base-case parameter values, for the CTMC we can fit the model to  $\delta$  up to around 0.13. When  $\delta \approx 0.13$ , we observe that the CTMC lockup premium becomes nearly 0.

## 6. The Lockup Premium Calculation

Once we have fit all the parameters, we can calculate the lockup premium. Denote the initial state of a fund by  $S_0$ . For a benchmark case (no extended lockup), at the end of each updating time ( $T_u$ ), investor can redeem his money from sick fund and reinvest the money in the good fund. Thus all funds starts with a good state each time. In this case, the fund's annual return up to  $n$  year is

$$\mathbb{E} \left[ e^{\sum_{i=1}^{n \cdot k} R_i^1} | S_0 = G \right] = [P(T_u)_{G,G} \cdot e^{Y_G} + P(T_u)_{G,S} \cdot e^{Y_S} + P(T_u)_{G,D} \cdot e^{Y_D}]^{n \cdot k}. \quad (6.29)$$

Now we consider a fund in  $n$ -year lockup case. The fund's expected return up to  $n$  year if the fund is under  $n$ -year lockup is

$$\mathbb{E} \left[ e^{\sum_{i=1}^{n \cdot k} R_i^2} | S_0 = G \right]. \quad (6.30)$$

Calculation of (6.30) is not immediate as (6.29) but this can be done by a recursion. Recall that the transition probability in the CTMC model,  $P(t)$ , is obtained in (4.18). Define

$$m(t, s) \equiv \mathbb{E} \left[ e^{\sum_{i=1}^t R_i^2} | S_0 = G, S_t = s \right] \cdot P(t)_{G,s}, \quad (6.31)$$

where  $P(t)$  defined in (4.12) represents the probability of reaching  $S_t$  from  $G$  at time  $t$ .

Conditioning on the state of a fund at time  $t$ , (6.31) is obtained from the following recursion formulas:

$$\begin{aligned} m(t, G) &= P(T_u)_{G,G} \cdot e^{Y_G} \cdot m(t-1, G) + P(T_u)_{S,G} \cdot e^{Y_G} \cdot m(t-1, S) \\ m(t, S) &= P(T_u)_{G,S} \cdot e^{Y_S} \cdot m(t-1, G) + P(T_u)_{S,S} \cdot e^{Y_S} \cdot m(t-1, S) \quad \text{and} \\ m(t, D) &= (1 - P(T_u)_{G,G} - P(T_u)_{G,S}) e^{Y_D} \cdot m(t-1, G) \\ &\quad + (1 - P(T_u)_{S,G} - P(T_u)_{S,S}) e^{Y_D} \cdot m(t-1, S), \end{aligned} \quad (6.32)$$

where  $m(1, G) = P(T_u)_{G,G} \cdot e^{Y_G}$ ,  $m(1, S) = P(T_u)_{G,S} \cdot e^{Y_S}$  and  $m(1, D) = (1 - P(T_u)_{G,G} - P(T_u)_{G,S}) e^{Y_D}$ . Notice that if a fund becomes dead before year  $n$ , it starts with a good state. Furthermore, the new good fund is now under 1-year lockup instead of  $n$ -year. Because of this,

care must be taken for a sample path once a fund becomes dead. We finally have

$$\mathbb{E} \left[ e^{\sum_{i=1}^{n \cdot k} R_i^2} | S_0 = G \right] = m(n \cdot k, G) + m(n \cdot k, S) + \sum_{t=1}^{n \cdot k} m(t, D) (P(T_u)_{G,G} e^{Y_G} + P(T_u)_{G,S} e^{Y_S})^{n \cdot k - t}. \quad (6.33)$$

Finally, the premium is calculated from (3.8) given that  $S_0 = G$ . The definition of the benchmark case is flexible.

## 7. Sensitivity Analysis for the New Markov Chain Model

In this section we do sensitivity analysis on the three performance measures,  $\gamma$  ( $\gamma_G, \gamma_S$ ),  $\delta$  and  $\sigma$ , respectively. We also analyze the impact of time periods used to estimate the performance measures, i.e., the updating time ( $T_u$ ), the time used to compute the persistence factor ( $T_p$ ), and the time period for computing the probability of death ( $T_d$ ) on the lockup premium.

### 7.1. How the Lockup Premium Depends on $\gamma$ ( $\gamma_G, \gamma_S$ ), $\delta$ and $\sigma$

We first study how much the lockup premium depends on the death probability  $\delta$ . Then, we see how much the premium depends on the model parameters  $\gamma$  ( $\gamma_G, \gamma_S$ ) and  $\sigma$ . When changing one performance measure in the sensitivity analysis, we fix all other measures just as the benchmark case with  $\delta = 0.03$  in Table 1.

Table 1 shows the model parameters for  $\delta = 0, 0.03, 0.06$ , and  $0.09$  and Figure 2 shows the lockup premium for different  $\delta$  values. The figures suggest that as the death probability  $\delta$  increases, the  $n$ -year lockup premium decreases. Notice that the CTMC model can incorporate the probability of death being  $0.09$  which is not possible in the DTMC model.

Table 2 shows the model parameters for  $\gamma = 0.4, 0.5$ , and  $0.6$  and Figure 3 (a) shows the lockup premium for  $\gamma = 0.4, 0.5$ , and  $0.6$ . We observe that as the persistence factor  $\gamma$  increases, the  $n$ -year lockup premium increases.

Table 2: Parameter value sets for  $\gamma$  ranging from 0.3 to 0.6

$\gamma$	$\delta$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S$	$Y_D$
0.3	0.03	0.3750	0.9646	0.1184	0.0659	-0.15	-0.20
0.4	0.03	0.2868	0.7315	0.1210	0.0670	-0.15	-0.20
0.5	0.03	0.2204	0.5531	0.1240	0.0690	-0.15	-0.20
0.6	0.03	0.1642	0.4142	0.1317	0.0704	-0.15	-0.20

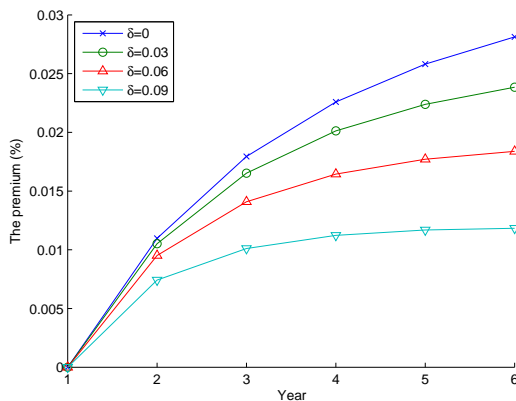


Figure 2: The lockup premium for the CTMC model for  $\delta = 0, 0.03, 0.06$ , and  $0.09$ . The parameter values come from Tables 2.

We next consider two separate persistence factors,  $\gamma_G$  and  $\gamma_S$ . The parameter values are given in Table 3 and the sensitivity of the lockup premium with respect to  $\gamma_G$  and  $\gamma_S$  is shown in Figure 3 (b). As  $\gamma_G$  increases and  $\gamma_S$  decreases, the fund in the state  $G$  tends to be in the same state while the fund in the  $S$  state tends to be in other states than the state  $S$  in the next time period. Since there is less chance of being in the state  $S$ , it is observed that the lockup premium decreases.

Table 3: Parameter value sets for  $\gamma_G$  and  $\gamma_S$

$\gamma_G$	$\gamma_S$	$\delta$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S$	$Y_D$
0.4	0.6	0.03	0.2238	0.4430	0.1041	0.0585	-0.15	-0.20
0.5	0.5	0.03	0.2204	0.5531	0.1240	0.0690	-0.15	-0.20
0.6	0.4	0.03	0.1962	0.6807	0.1612	0.0768	-0.15	-0.20

We lastly check the sensitivity of the lockup premium with respect to  $\sigma$ . Our TASS database analysis estimates  $\sigma$  of annual returns for each year is lower than 0.20 in most cases. We here highlight the sensitivity of the lockup premium for  $\sigma = 0.05, 0.10$ , and  $0.15$  with  $\gamma = 0.5$ . Table 4 provides the parameter value sets and Figure 4 shows the corresponding lockup premiums. We see that the premium increases in  $\sigma$ .

## 7.2. How the lockup premium depends on $T_u$ , $T_p$ , and $T_d$ .

The CTMC model is flexible in calibration such that the three key-performance measures can be estimated in different time periods. Since the three measures can be estimated independently, allowing different time period for each measure could be helpful to an investor who has different

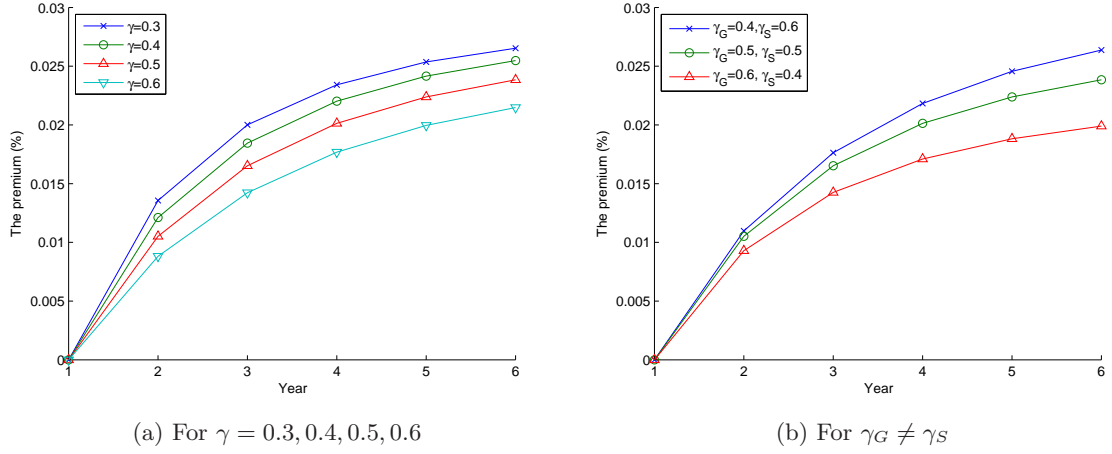


Figure 3: The lockup premium for the CTMC model (a) for  $\gamma = 0.3, 0.4, 0.5$  and  $0.6$  for other parameter values in Tables 2 and (b) for  $\gamma_G \neq \gamma_S$  for other parameter values in Table 3.

Table 4: Parameter value sets for  $\sigma = 0.05, 0.10, 0.15$

$\sigma$	$\delta$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S = -1.5\sigma$	$Y_D = -2.0\sigma$
0.05	0.03	0.8434	0.3721	0.5025	0.0342	-0.075	-0.10
0.10	0.03	0.8434	0.3721	0.5025	0.0684	-0.150	-0.20
0.15	0.03	0.8434	0.3721	0.5025	0.1026	-0.225	-0.30

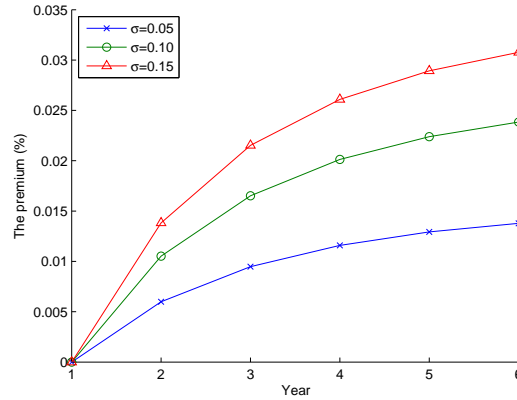


Figure 4: The lockup premium for  $\sigma = 0.05, 0.10,$  and  $0.15$ .

views in these estimation-time periods. In this section, we investigate how the lockup premium in the CTMC model depends on the three measurement times  $T_u, T_p,$  and  $T_d$ .

We first study how much the lockup premium depends on the updating period  $T_u$ . As  $T_u$  decreases, an investor can rebalance his portfolio of hedge funds more frequently. For a fixed lockup period, as  $T_u$  decreases, the investor has more updating frequencies in a year, which is

restricted by the lockup condition. Thus if all the other parameter values are the same, the lockup premium may increase as  $T_u$  decreases. However, at the same time, the investor more frequently rebalances a fund in the state  $D$  with a fund in the state  $G$ , which may decrease the lockup premium. In Figure 5, we numerically calculate the lockup premium up to six years for  $T_u = 0.25, 0.5$ , and 1 year. Parameter values are shown in Table 5. In order to compare the lockup premium for different  $T_u$ , we adjust annual relative return rates under good, sick, and dead states from  $i^{th}$  to  $i + 1^{th}$  updating time by multiplying  $T_u$ , i.e.,  $T_u Y_G$ ,  $T_u Y_S$ , and  $T_u Y_D$ . We then fit such parameters to the standard deviation of return for  $T_u$ , i.e.,  $\sigma\sqrt{T_u}$ . The figure shows that as the rebalance period increases, the lockup premium decreases. Thus it can be said that the investor requires more compensation for the lockup condition when there are more rebalancing opportunities that are missed because of the lockup period.

Table 5: Parameter value sets for  $T_u = 0.25, 0.5$ , and 1

$T_u$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S \equiv -1.5T_u\sigma\sqrt{T_u}$	$Y_D \equiv -2.0T_u\sigma\sqrt{T_u}$	$\sigma\sqrt{T_u}$
0.25	0.5900	0.1084	0.0376	0.1068	-0.038	-0.05	0.0500
0.5	0.4358	0.2758	0.0531	0.0893	-0.075	-0.10	0.0708
1	0.2191	0.5533	0.1250	0.0684	-0.150	-0.20	0.1000

We also investigate the impact of  $T_p$  on the lockup premium. As  $T_p$  decreases, the persistence factor is observed in a shorter period. If the persistence factors are the same for different  $T_p$ , the impact of the persistence factor becomes higher with shorter  $T_p$  than with the longer one. Figure 6 shows that the lockup premium is increasing as  $T_p$  decreases, if  $\gamma$  remains the same for each  $T_p$ . Table 6 shows the parameter values.

Table 6: Parameter value sets for  $T_p = 0.25, 0.5$ , and 1

$T_p$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S \equiv -1.5T_u\sigma$	$Y_D \equiv -2.0T_u\sigma$
0.25	0.8163	2.0287	0.1103	0.0156	-0.15	-0.20
0.5	0.4182	1.0431	0.1148	0.0322	-0.15	-0.20
1	0.2191	0.5533	0.1250	0.0667	-0.15	-0.20

Lastly, we investigate how the lockup premium depends on  $T_d$ . If the death probability  $\delta$  is the same for different  $T_d$ , the annual death probability with smaller  $T_d$  tends to produce more funds in state  $D$  for given time period. Figure 7 suggests that the lockup premium is increasing as  $T_d$  increases, just as the premium increases as  $\delta$  increases for fixed  $T_d$ . Table 7 shows the parameter values.

Table 7: Parameter value sets for  $T_d = 0.25, 0.5,$  and  $1$

$T_d$	$\mu_G$	$\lambda_S$	$\mu_S$	$Y_G$	$Y_S \equiv -1.5T_u\sigma$	$Y_D \equiv -2.0T_u\sigma$
0.25	0.2744	3.5967	3.3273	0.0726	-0.15	-0.20
0.5	0.2262	0.7022	0.3491	0.0699	-0.15	-0.20
1	0.2191	0.5533	0.1250	0.0684	-0.15	-0.20

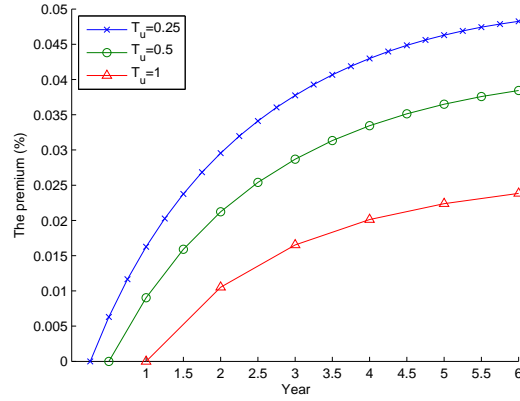


Figure 5: The lockup premium for  $T_u = 0.25, 0.5$  and  $1$  years.

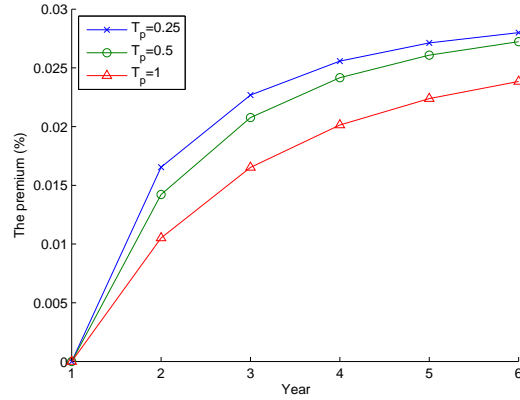


Figure 6: The lockup premium for  $T_p = 0.25, 0.5$  and  $1$  years.

### 7.3. Comparison of the CTMC and the DTMC Models

In this section, we compare the lockup premium in the CTMC model (Figure 2) to the one in the DTMC model (Figure 8 (a) of DPW09).

First, we observe that the lockup premiums in the DTMC model and the CTMC model have similar values. However, we also observe that the lockup premium in the CTMC model is slightly larger than for the DTMC model under the same performance measures and lockup period.

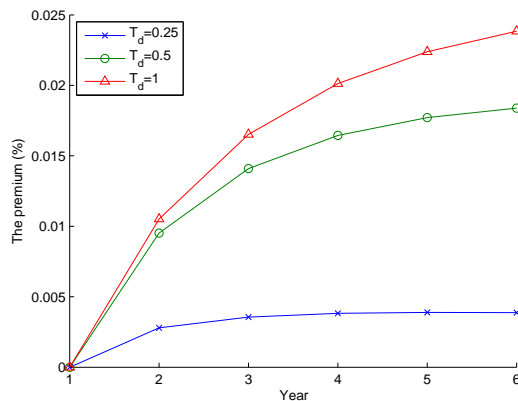


Figure 7: The lockup premium for  $T_d = 0.25, 0.5$  and 1 years.

We compare the transition probabilities in the DTMC and CTMC models under the same performance measures and measurement times (benchmark case) in order to understand this difference. We see that the transition probabilities from the state  $G$  to  $G$  in both models are almost the same (0.850 in the DTMC model, 0.846 in the CTMC model) and so are the transition probabilities from  $G$  to  $S$ . The probabilities from  $S$  to  $G$  are also similar (0.358 in the DTMC model and 0.365 in the CTMC model). However, the difference in transition probabilities from  $S$  to  $S$  (0.493 in the DTMC model and 0.543 in the CTMC model) is relatively bigger than the differences in other probabilities. Thus, a fund in the state  $S$  is more likely to stay in  $S$  in the next time period in the CTMC model than in the DTMC model, which makes the lockup premium slightly bigger in the CTMC model.

The transition probability from the state  $S$  to the state  $D$  in the DTMC model (0.132) is bigger than that (0.092) in the CTMC model. Recall that once a fund becomes dead in the state  $D$ , the lockup condition is not valid any more and the fund is replaced with a fund in the state  $G$ , decreasing the lockup premium. This can also increase the lockup premium in the CTMC model.

Comparing the three equations (5.20)-(5.22) that are used to parameter fitting in the CTMC model to (6.4)-(6.6) in DPW09 in the DTMC model, we see that the difference is additional terms in (5.20) and (5.22) that incorporate the transition from  $G$  to  $D$  in one time period. As a result of these additional terms, the transition probability from  $S$  to  $S$  is larger and the transition probability from  $S$  to  $D$  is smaller in the CTMC model than in the DTMC model. We thus conclude that while the lockup premium in the CTMC model increases only slightly by allowing the transition from  $G$  to  $D$  in one year, which is caused by allowing the transition from the state  $G$  to the state  $D$  within one time period.



## 8. Conclusion

In this paper, we have developed new Markov chain models to calculate the premium from extended hedge-fund lockups. As in our previous paper DPW09, we use a highly stylized three-state model to represent the state of a hedge fund, using the states “good,” “sick” and “dead.” Since investors can redeem their investment (or part of it) and reinvest if the fund dies, the important state for extended hedge fund lockup is the sick state.

We define the lockup premium as a compensation to the investor for the opportunity cost of restricted re-balancing activities; see §3. In order to calculate the lockup premium, we compare two identical hedge funds with and without extended lockups. The premium is the fixed annual rate of return that makes the expected returns of the two funds the same.

The new CTMC model for the state of a hedge fund allows the fund state to change continuously in time. The previous DTMC model in DPW09 required transitions from good to dead to take at least two years. With the new CTMC model, that restriction is removed. The CTMC model also provides increased flexibility in modeling and performance-measure estimation. First, the CTMC model allows flexible performance measure estimation, e.g., semi-annually or quarterly from the data. Second, the CTMC model allows wider ranges of acceptable performance measures than the DTMC model does.

In §4 we gave an explicit representation of the ergodic two-state DTMC model based on the three CTMC model parameters  $\mu_G$ ,  $\mu_S$  and  $\lambda_S$ . In order to calibrate these three parameters, in §5 we developed a nonlinear programming algorithm, which can be easily solved numerically with MATLAB. In order to calibrate the standard deviation of returns,  $\sigma$ , we used an efficient iterative search algorithm. Thus we succeeded in developing an efficient algorithm to carry out the model fitting. The model fitting yields the parameters  $\mu_G$ ,  $\mu_S$ ,  $\lambda_S$  and  $Y_G$  plus the premium for extended lockup as a function of the time parameters  $T_p$ ,  $T_u$  and  $T_d$ , the persistence parameters  $\gamma_G$  and  $\gamma_S$ , the death parameter  $\delta$  and the relative return parameters  $Y_S$  and  $Y_D$ , where the parameters  $\gamma_G$ ,  $\gamma_S$ ,  $\delta$ ,  $Y_S$  and  $Y_D$  can be estimated from historical data, as we illustrated by applying the TASS data. We then can compute the premium for extended lockup as a function of the number  $n$  of years of extended lockup and again the time parameters  $T_p$ ,  $T_u$  and  $T_d$ , the persistence parameters  $\gamma_G$  and  $\gamma_S$ , the death parameter  $\delta$  and the relative return parameters  $Y_S$  and  $Y_D$ .

We then conducted sensitivity analysis to show how the premium for extended lockup depends on the variables. The sensitivity analysis quantifies how the lockup premium increases

as a function of the persistence factor  $\gamma$  and the standard deviation  $\sigma$ , but decreases as a function of the death probability  $\delta$ . We also have examined how the lockup premium depends on different values of measurement times,  $T_u$ ,  $T_p$  and  $T_d$ . We conclude that the impact of three performance measures -  $\delta$ ,  $\gamma$  and  $\sigma$  - have a significant impact on the lockup premium just as we saw for the DTMC models in DPW09. Also, we conclude that while the impact of  $T_p$  are relatively smaller on the lockup premium, the impact of  $T_u$  and  $T_d$  can be significant.

The numerical values of the lockup premium with the new Markov chain models are very similar to those for the previous DTMC model. We thus conclude that the more rudimentary DTMC model does not suffer greatly from its restriction to yearly updating. A slight increase in the lockup premium in the new model can be caused by allowing the transition from good to sick within one year.

Our approach here may also be useful to study extended lockup premiums for other investment opportunities that have the good-sick-dead structure treated here; e.g., private equity funds, where limited partners are expected to maintain their interest in the fund until it is unwound, typically over a ten-year life.

**Acknowledgments.** The authors thank Emanuel Derman for suggesting this research problem and for his helpful comments on this paper. The second author acknowledges support from NSF grants DMI-0457095 and CMMI-0948190.

## References

- Agarwal, V., N. Y. Naik. 2000. Multi-period performance persistence analysis of hedge funds. *Journal of Financial and Quantitative Analysis* **35**(3) 327–342.
- Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney, D. Sorensen. 1999. *LAPACK User's Guide*. Third Edition, SIAM, Philadelphia.
- Ang, A., N.P.B. Bollen. 2009. Lockup up by a lockup: valuing liquidity as a real option. *Financial Management* **39**(3) 1069–1095.
- Aragon, G. 2007. Share restrictions and asset pricing: Evidence from the hedge fund industry. *Journal of Financial Economics* **83**(1) 33–58.
- Ben-David, I., F. Franzoni, R. Moussawi. 2010. Hedge fund stock trading in the financial crisis of 2007-2008. *Working paper* .
- Boyle, P., S. Li, Y. Zhu. 2010. Hedge fund redemption restrictions, financial crisis, and fund performance. *Working paper, Wilfrid Laurier University* .

- Boyson, N., R. Stulz, C. Stahel. 2010. Hedge fund contagion and liquidity shocks. *Journal of Finance* **65**(5) 1789–1816.
- Browne, S. J., M. A. Milevsky, T. S. Salisbury. 2003. Asset allocation and the liquidity premium for illiquid annuities. *The Journal of Risk and Insurance* **70**(3) 509–526.
- De Roon, F. A., J. Guo, J. R. ter Horst. 2009. Being locked up hurts. *SSRN Working paper* .
- Derman, E., K. S. Park, W. Whitt. 2009. Markov chain models to estimate the premium for the extended hedge fund lockup. *Wilmott Journal* **1**(5-6) 263–293, Available at <http://www.columbia.edu/~ww2040/allpapers.html>.
- Derman, E., K. S. Park, W. Whitt. 2010. A stochastic-difference-equation model for hedge-fund relative returns. *Quantitative Finance* **10**(7) 701–733.
- Edwards, F. R., M. O. Caglayan. 2001. Hedge fund performance and manager skill. *Journal of Futures Markets* **21**(11) 1003–1028.
- Golts, M., M. Kritzman. 2010. Liquidity options. *Reverse Stret Working Paper Series* .
- Hull, J. C. 2003. *Options, Futures, and Other Derivatives*. 5th ed. Prentice-Hall.
- Jagannathan, R., A. Malakhov, D. Nonikov. 2006. Do hot hands persist among hedge fund managers. NBER Working paper, 12015.
- Karlin, S., H. M. Taylor. 1975. *A First Course in Stochastic Processes*. Academy Press.
- Kazemi, H. 2010. Asset allocation and illiquidity risk. *CAIA Association* .
- Koh, F., W. T. H. Koh, M. Teo. 2003. Asian hedge funds: Return persistence style and fund characteristics. *Working Paper, Singapore Management University* .
- 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.
- Park, H. 2007. Risk Measures for Hedge Funds and a Survival Analysis. *Ph.D thesis, University of Massachusetts* .
- Ross, S. M. 2003. *Introduction to Probability Models*. Eighth ed. Academy Press.
- Rouah, F. 2006. Competing Risks in Hedge Fund Survival. *Ph.D thesis, McGill University* .
- Schittkowski, K. 1986. Nlqpl: A fortran-subroutine solving constrained nonlinear programming problems. *Annals of Operations Research* **5**(2) 485–500.