© *The Journal of Risk and Insurance*, 2003, Vol. 70, No. 3, 489-508

# An Empirical Study on the Lapse Rate: The Cointegration Approach

Weiyu Kuo
Chenghsien Tsai
Wei-Kuang Chen

**Abstract**

We use the cointegration technique to reexamine the contending lapse rate hypotheses: the emergency fund hypothesis and the interest rate hypothesis. We find that the unemployment rate affects the lapse rate in both the long and short run, whereas the interest rate causes variations in the lapse rate mainly in the long run. This evidence seems to be in favor of the emergency fund hypothesis. However, according to the impulse response analysis of the estimated error-correction model, the interest rate overwhelms the unemployment rate on the overall impact on the dynamics of lapse rate. In other words, the interest rate hypothesis is favored against the emergency fund hypothesis in the sense that the interest rate is more economically significant than the unemployment rate in explaining the lapse rate dynamics.

## Introduction

Understanding the dynamics of the lapse rate is important to insurance companies. First, policy lapse might make the insurer unable to fully recover his initial expenses—the cost of procuring, underwriting, and issuing new business. The insurer pays these expenses at or before the time of issue but earns profits over the life of the contract. The insurer thus might incur losses from lapsed policies. Second, policy lapse might involve mortality or morbidity adverse selection. Policyholders who have adverse health or other insurability problems tend not to lapse their policies, causing the insurer to experience more claims than expected if the lapse rate is high. Third, policy lapse, or more specifically, early surrender, poses a liquidity threat to the

Weiyu Kuo is associate professor, Department of International Trade; Chenghsien Tsai is associate professor, Department of Risk Management and Insurance; and Wei-Kuang Chen is professor, Department of Money and Banking, all at the National Chengchi University. The authors would like to formally express their gratitude to Khari Cook and Michele Tibbs at the American Council of Life Insurance who provided the lapse rate data for the period from 1951 to 1965 and to Larry Y. Tzeng at National Taiwan University for his insightful comments. Helpful comments from two anonymous referees and Richard MacMinn are also gratefully acknowledged.

insurer.[1] Most insurers include in their contracts a provision that grants the policyholder who elects to terminate the policy a right to a cash surrender value. This policyholder's option to demand the policy's cash surrender value at any time could cause a liquidity problem to insurers. Life insurers usually have an excess of income over expenditures that is more than sufficient to meet the demand for cash surrender values. During difficult economic times, however, demands by policyholders might be so great that it becomes necessary for insurers to liquidate assets at depressed prices. The disintermediation that happened to U.S. life insurers during the 1980s demonstrated the adverse impact of excessive early surrenders. Many life insurers experienced negative cash flows for the first time since the Great Depression of the 1930s (Black and Skipper, 2000, p. 111). Furthermore, to alleviate the liquidity threat, the insurer has to adopt a more liquid investment policy than would otherwise be required, which implies lower investment returns. Understanding lapse behavior is therefore important for an insurer's liquidity and profitability.

Understanding the relationship between the lapse rate and interest rate is even more important. If the lapse rate is related to the interest rate, an insurance company's cash flows will depend on the interest rate. Babbel (1995) and Briys and Varenne (1997) found that the dependence of a policy's cash flow on the interest rate was critical to the duration and convexity of the policy. They showed that misspecification of lapse rate sensitivity to the interest rate could cause large errors in the effective duration estimates and even greater errors in the convexity estimates for insurance policies. Furthermore, Albizzati and Geman (1994) and Grosen and Jorgensen (2000) demonstrated that the option to surrender the policy for its cash value could account for a substantial portion of the present value of all future premiums. If the surrender option is exercised rationally with changes in the interest rate, it could account for more than 50 percent of the contract value. However, these valuation models lack a robust lapse rate model. Understanding the relationship between the lapse and interest rates is therefore crucial to insurance companies.

What causes lapses has attracted certain academic interest for some time. Two hypotheses have been proposed to explain lapse behavior. The emergency fund hypothesis contends that policyholders utilize cash surrender values as emergency funds when facing personal financial distress. Moreover, policyholders may not be able to maintain premium payments for insurance coverage and are therefore inclined to terminate their policies during personal financial distress. A testable implication of this hypothesis is that the lapse rate would increase during economic recessions. On the other hand, the interest rate hypothesis conjectures that the lapse rate rises when the market interest rate increases because the latter acts as the opportunity cost

[1] There are subtle differences between lapse and surrender. Policyholders could actively terminate their policies through surrendering or let their policies lapse by not paying premiums. Early surrender usually involves cash value payments, but policy lapse might not because some policies deduct premiums due from cash surrender values when policyholders stop paying premiums to keep policies valid until the cash values are used up. Since the traditional measure of the lapse rate includes both lapsed and surrendered policies, we broaden the meaning of lapse in this study to take account of surrender as well as lapse.

for owning insurance contracts. Furthermore, when interest rates rise equilibrium premiums fall. There is a greater likelihood that a newly acquired policy will provide the same coverage at a lower premium. Policyholders thus tend to surrender their policies to exploit higher yields and/or lower premiums available in the markets. Since no theoretical models have been established to explain the causes of policy lapse, the relative significance of the emergency fund hypothesis and the interest rate hypothesis remains an empirical issue.[2]

Few empirical studies looked into the plausibility of these two hypotheses in explaining lapse behavior. Outreville (1990) found consistent support for the emergency fund hypothesis based on two lapse rate measurements in both the United States and Canada.[3] The support for the interest rate hypothesis is weak because interest rate variables are significant in only one of four cases. The interest rate hypothesis found its ground from more recent actuarial studies in the *Transactions of Society of Actuaries Reports* (Cox et al., 1992) and the Annuity Persistency Study in the 1995–1996 reports (pp. 559-638). These studies found that the lapse rate increased with the spread between the policy's credited interest rate and the market interest rate. These somewhat inconsistent results are probably due to the differences in sample periods and methods. Whereas the sample periods for the actuarial studies were the late 1980s and 1990s, the sample in Outreville (1990) covered up to 1979 only and missed the wide swing in interest rates during the 1980s and 1990s. Outreville performed ordinary least squares (OLS) analysis with the Cochrane-Orcutt adjustment for the first-order serial correction of the residuals. However, the analyses in the actuarial reports were performed using univariate analysis without control variables. Since evidence is relatively scarce and inconclusive, the underlying causes of policy lapse are still open to debate. This study aims to empirically construct a model that captures the dynamics among the lapse rate, interest rate, and unemployment rate.

Our empirical modeling has two advantages over the models developed in previous studies. First, we use the cointegrated vector autoregression model developed by Engle and Granger (1987) to construct our empirical model. Cointegration modeling is able to separate the potential long-term relationship among the lapse rate, interest rate, and unemployment rate from their short-term adjustment mechanisms. This is very different from the OLS method used by Outreville (1990) that focuses mainly on the short-term dynamics and ignores the potential long-term relations. A long-term relationship among the lapse rate, interest rate, and unemployment rate could result from the existence of an equilibrium relationship coupled with a partial adjustment to the equilibrium. It could also be the product of the aggregate behavior of policyholders based on their rational expectation of the future lapse rate. The interest rate and unemployment rate might therefore influence the lapse rate through a long-term channel that normally cannot be identified using OLS techniques. The

[2] Note that the emergency fund hypothesis and the interest rate hypothesis are not mutually exclusive and can both be important in explaining lapse behavior.

[3] One of the lapse rate measurements is reported by the Life Insurance Marketing and Research Association and the other by the American Council of Life Insurance.

cointegrated vector autoregression model could shed light on the distinction between the emergency fund hypothesis and the interest rate hypothesis in a more comprehensive framework.

Second, our data cover a longer sample period than previous studies. We use 48 observations of annual lapse rates, interest rates, and unemployment rates dating from 1951 to 1998, whereas previous articles used less than 28 observations with sampling periods of less than 25 years. A larger sample size enhances the power of our estimation. In addition, according to Otero and Smith (2000), the total sample length is more important than the number of observations for the Johansen cointegration tests to identify cointegration. Thus, our longer data also secure the robustness of our cointegration analysis so long as we can presume that the process is sufficiently stationary for our statistical methods.

We find that the unemployment rate is statistically significant in explaining the short-term lapse rate dynamics whereas the explanatory power of interest rate is relatively weak in the short run. This result is consistent with that of Outreville (1990). We also find that two significant cointegrated long-term relationships exist between the lapse rate and unemployment rate and between the interest rate and unemployment rate that are not identified in Outreville's model. In other words, both the interest rate and unemployment rate affect the lapse rate through long-term mechanisms. Overall, our results indicate that the unemployment rate affects the lapse rate in the long and short run. The interest rate causes variations in the lapse rate mainly in the long run, but its short-term impact is less obvious. Outreville's findings can be accommodated in our more flexible empirical model.

The statistically significant results involving the long-term and short-term dynamics tell us nothing about the relative economic importance of the interest rate and unemployment rate in causing variations in the lapse rate. To shed light on this issue, we perform impulse-response analysis based on the estimated error-correction model. We find that the lapse rate responds insignificantly to the shocks from the unemployment rate but significantly to the shocks from the interest rate. Combining these results, we conclude that the interest rate hypothesis is more consistent with the evidence than the emergency fund hypothesis.

The article is organized as follows. Section 2 briefly discusses the cointegration vector atuoregression model. Section 3 describes our data and the estimation of error-correction models. The implications of the estimated model are also discussed in this section. Section 4 concludes this article and suggests future research directions.

## COINTEGRATION ANALYSIS

Since the seminal paper by Engle and Granger (1987), cointegration analysis has become a popular method for economists to study the long-term relationships between various economic variables, such as consumption and income, short- and long-term interest rates, and stock prices and dividends. The basic intuition behind cointegration analysis is that even though a group of nonstationary variables might individually wander extensively, these variables can be expected to wander in such a way that they do not drift too far apart from one another. That is, although individually they are time series with unit roots, a particular linear combination of them is stationary. Such variables are said to be cointegrated. We delineate the definition and estimation procedure for cointegrated vectors as follows.

A $(n \times 1)$ vector time series $\mathbf{y_t}$ is defined as cointegrated in $(d, b)$ order if each of the series is individually an $I(d)$ process, namely, a nonstationary process with $d$ unit roots, whereas a certain linear combination of the series $\mathbf{a}'\mathbf{y_t}$ is an $I(d\text{-}b)$ process for some nonzero $(n \times 1)$ constant vector $\mathbf{a}$. The vector $\mathbf{y_t}$ considered in this study contains three variables $y_{1t}$, $y_{2t}$, and $y_{3t}$, where $y_{1t}$ is the lapse rate, $y_{2t}$ is the interest rate, and $y_{3t}$ is the unemployment rate. Suppose that $\mathbf{y_t}$ is cointegrated in the (1,1) order. Then, according to the Granger Representation Theorem of Engle and Granger (1987), $\mathbf{y_t}$ follows an error-correction model of the form

$$\mathbf{C(L)\Delta y_t} = \boldsymbol{\mu} + \gamma \mathbf{y_{t-1}} + \varepsilon_{\mathbf{t}}, \tag{1}$$

where $\mathbf{C(L)}$ is a $3 \times 3$ matrix polynomial in the lag operator $\mathbf{L}$ of order $p$, $\boldsymbol{\Delta}$ is the first-order difference operator, $\boldsymbol{\mu}$ is an intercept vector, $\gamma$ is a $3 \times 3$ constant matrix, and $\varepsilon_{\mathbf{t}}$ is a white noise error term vector.

Cointegration analysis generally involves four steps: (1) ensuring that the individual elements of $\mathbf{y_t}$ are $I(1)$ processes, (2) determining the order of the vector autoregression (VAR) model, (3) performing cointegration tests to determine the rank of the cointegrated system, and (4) estimating the error-correction model. We utilize the augmented Dickey-Fuller (ADF) unit root test to examine if the elements of $\mathbf{y_t}$ are $I(1)$ processes individually. Since the ADF test depends critically on the assumption about the underlying process and the estimated regression, we conduct the test based on three different assumptions that correspond to cases 1, 2, and 4 in Hamilton (1994, p. 502).[4] In particular, we estimate the following regressions:

$$y_{it} = \theta_i y_{it-1} + \rho_{i1}\Delta y_{it-1} + \rho_{i2}\Delta y_{it-2} + \cdots + \rho_{ip-1}\Delta y_{it-p+1} + \varepsilon_{it}, \quad i = 1, 2, 3, \tag{2}$$

$$y_{it} = \mu_i + \theta_i y_{it-1} + \rho_{i1}\Delta y_{it-1} + \rho_{i2}\Delta y_{it-2} + \cdots + \rho_{ip-1}\Delta y_{it-p+1} + \varepsilon_{it}, \quad i = 1, 2, 3, \tag{3}$$

[4] There are four cases for the ADF unit root test discussed in Hamilton (1994). These cases cover all potential combinations of the underlying unit root process under the null hypothesis and the estimated regression under the alternative hypothesis. In cases 1 and 2, the underlying process is assumed to be a random walk without drift that is estimated by a stationary autoregression without an intercept in the former and with an intercept in the latter. In contrast, the underlying process is assumed to follow a random walk with a drift term in both cases 3 and 4 that is estimated by a stationary autoregression with an intercept in the former and with an intercept and a deterministic time trend in the latter. The reason why we decide not to test case 3 is that it can be deemed as a special case of case 4.

and

$$y_{it} = \mu_i + \delta_i t + \theta_i y_{it-1} + \rho_{i1}\Delta y_{it-1} + \rho_{i2}\Delta y_{it-2} + \cdots + \rho_{ip-1}\Delta y_{it-p+1} + \varepsilon_{it}, \quad i = 1, 2, 3, \quad (4)$$

where $y_{it}$ is the $i$th element of $\mathbf{y_t}$. Obviously, Equations (2), (3), and (4) differ from one another in the assumption about whether an intercept or a deterministic time trend is included in the regression. We test the following three null hypotheses that correspond to the above three regressions:

$$H_{20} : \theta_i = 1, \quad H_{30} : \mu_i = 0 \text{ and } \theta_i = 1, \text{ and } H_{40} : \theta_i = 1 \text{ and } \delta_i = 0.$$

After the unit root test, the order of the VAR model must be determined. The Akaike information criterion (AIC) and the Bayesian information criterion (BIC) derived by Schwarz (1978) are normally used for this purpose. It is known that BIC is more parsimonious than AIC in the sense that it usually selects a model with a lower order as the optimal model than the one chosen by AIC. In view of our limited data, we decide to adopt BIC to determine the order of the VAR model.

The maximum likelihood multivariate cointegration tests developed by Johansen (1991) are used to conduct the cointegration test. The hypothesis of interest involves the rank of $\gamma$. If the rank of $\gamma$ is $q$ and $q \leq n - 1$, then one can decompose $\gamma$ into two $n \times q$ matrices $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ such that $\gamma = \boldsymbol{\alpha}\,\boldsymbol{\beta}'$. The matrix $\boldsymbol{\beta}$ contains $q$ linear cointegration parameter vectors whereas $\boldsymbol{\alpha}$ is a matrix consisting of $n$ error-correction parameter vectors. The maximum likelihood estimate of $\boldsymbol{\alpha}$ is obtained using the OLS regression of $\Delta\mathbf{y_t}$ on $\Delta\mathbf{y_{t-1}}, \ldots, \Delta\mathbf{y_{t-p+1}}$ , and 1 whose residual is $\hat{\varepsilon}_{\mathbf{0t}}$. Similarly, the maximum likelihood estimate of $\boldsymbol{\beta}$ can be obtained from the OLS regression of $\mathbf{y_{t-1}}$ on $\Delta\mathbf{y_{t-1}}, \ldots,$ $\Delta\mathbf{y_{t-p+1}}$, and 1 whose residual is $\hat{\varepsilon}_{\mathbf{1t}}$. Based on the residuals $\hat{\varepsilon}_{\mathbf{0t}}$ and $\hat{\varepsilon}_{\mathbf{1t}}$, we have the residual product matrices

$$\mathbf{S}_{ij} = T^{-1}\sum_{t=1}^{T} \hat{\varepsilon}_{it}\hat{\varepsilon}'_{jt}, \quad i, j = 0, 1. \quad (5)$$

We then solve the eigenvalue system

$$\left|\lambda\mathbf{S}_{11} - \mathbf{S}_{10}\mathbf{S}_{00}^{-1}\mathbf{S}_{01}\right| = 0$$

for eigenvalues $\hat{\lambda}_1 > \cdots > \hat{\lambda}_n$ and eigenvectors $\hat{\boldsymbol{\Psi}} = \left(\hat{\psi}_{\mathbf{1}}, \ldots, \hat{\psi}_{\mathbf{n}}\right)$. The estimates for $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ are given by $\hat{\boldsymbol{\alpha}} = \mathbf{S}_{01}\hat{\boldsymbol{\beta}}$ and $\hat{\boldsymbol{\beta}} = \left(\hat{\psi}_{\mathbf{1}}, \ldots, \hat{\psi}_{\mathbf{q}}\right)$, where $\hat{\psi}_{\mathbf{1}}, \ldots, \hat{\psi}_{\mathbf{q}}$ are the eigenvectors associated with the $q$ largest eigenvalues. Two Johansen's maximum likelihood tests, the maximal eigenvalue test and the trace test, can then be used to determine the number of cointegration vectors. The statistic from the maximal eigenvalue test for the null hypothesis of $q$ cointegration vectors against the alternative of $q + 1$ cointegration vector is

$$\hat{\lambda}_{\max} = -T\ln(1 - \hat{\lambda}_{q+1}).$$

The trace test statistic for the null hypothesis of at most $q$ cointegration vectors is

$$\hat{\lambda}_{\text{trace}} = -T \sum_{j=q+1}^{n} \ln(1 - \hat{\lambda}_j).$$

If the results are consistent with the hypothesis of at least one cointegration vector, one then uses the maximum likelihood method to test the hypotheses regarding the restriction on $\boldsymbol{\beta}$.[5]

After determining the number of cointegration vectors through the maximal eigenvalue test and trace test, we continue to estimate the following error-correction model of the lapse rate, interest rate, and unemployment rate.

$$\Delta \mathbf{y_t} = \boldsymbol{\mu} + \gamma \mathbf{y_{t-1}} + \xi_1 \Delta \mathbf{y_{t-1}} + \cdots + \xi_\mathbf{p} \Delta \mathbf{y_{t-p+1}} + \varepsilon_\mathbf{t},$$

where the vector $\mathbf{y_t}$ consists of the lapse rate, interest rate, and unemployment rate at time $t$, $\Delta$ is the first-order difference operator, $\boldsymbol{\mu}$ is a $3 \times 1$ intercept vector, $\gamma$ is a $3 \times 3$ constant matrix, $\varepsilon_\mathbf{t}$ is a $3 \times 1$ white noise error term vector, and the optimal lag $p$ is determined according to BIC.

## EMPIRICAL ESTIMATION AND RESULTS

### Data

We acquire the lapse rate data from the *Life Insurance Fact Book*, an annual statistical report of the American Council of Life Insurance (ACLI). The data in the *Fact Book* are derived from the annual statements filed by life insurance companies with the National Association of Insurance Commissioners, ACLI's surveys, and/or external sources such as government agencies and trade associations. Our sample contains the annual voluntary termination rates for all ordinary life insurance policies in force from 1951 to 1998. The voluntary termination rate equals the ratio of the number of lapsed or surrendered policies to the mean number of policies in force. Sampled policies cover permanent insurance (including universal life, variable life, variable-universal life, and traditional whole life that includes limited payment, continuous premium, joint whole life, single premium, adjustable life, monthly debit, and other fixed-premium products), term insurance (including decreasing, level and other term, term additions), and endowment insurance.[6] Compared with the data in Outreville (1990) covering the period 1955–1979, our sample spans a longer period and extends over the era of highly volatile interest rates in the 1980s and early 1990s. We obtain

[5] Those that are interested in greater detail about the derivation of the maximal eigenvalue test and the trace test may refer to Johansen (1991).

[6] We are aware that the changed mix of various types of policies over the last several decades might have some impact on the lapse rate and the relations among the lapse rate, interest rate, and unemployment rate. However, we have no adequate data to assess such an impact. This is a limitation of our study.

the 90-day Treasury rate[7] from the U.S. Financial Database and unemployment rate from the Taiwan Economic Journal Database.[8]

### Estimation of the Error-Correction Model

Following the cointegration analysis procedure, we first verify that the lapse rate, interest rate, and unemployment rate are $I(1)$ processes before testing for the cointegrating relations among them. We employ the ADF test to examine whether there are unit roots in these three variables. As mentioned before, the asymptotic distribution of the unit root test depends on whether the selected optimal regression includes an intercept, a deterministic time trend, or none of them. We thus need to decide which specification to use for the ADF test. Specifically, we follow the general principle suggested by Hamilton (1994) to fit a specification that is a plausible description of the data under both the null hypothesis and the alternative.

Figure 1 shows the time series for the lapse rate, interest rate, and unemployment rate from 1951 through 1998. It is apparent that there is a time trend in these three series from 1951 through 1982, although there is a serious setback during the later period from 1983 to 1998. Based on this observation, the general principle would guide us to perform the ADF test based on regression (4).[9]

We also conduct the ADF test on the first-order differences of the lapse rate, interest rate, and unemployment rate to confirm that all of these series are $I(1)$ processes rather than processes with higher-order integration, that is, $I(j)$, $j > 1$. If this is the case, the ADF test would reject the null hypothesis of a unit root in these differenced series at conventional significance levels. Since the first-order differences of these series, as shown in Figure 2, fluctuate almost randomly around zero, we should include neither an intercept nor a deterministic time trend in the regression of the ADF test. In other words, we choose the form of regression (2) for the first-order differences of the series.

Table 1 reports the results of the unit root test on the lapse rate, interest rate, unemployment rate, and their first-order differences. All of the ADF statistics for the

[7] Since the duration of life insurance policies is probably measured in decades, long-term interest rates might be more appropriate. We choose the 90-day T-bill rate to maximize our sample period because long-term Treasury rates have shorter sample periods in our data set. Nevertheless, our results are robust to Treasury rates with different maturities. The cointegration relations among the lapse rate, interest rate, and unemployment rate when using long-term rates such as the 1-year T-bill rate, 5-year T-note rate, and 10-year T-bond rate are qualitatively identical to those in this article. The results of cointegration analysis based on long-term interest rates are available from the authors upon request.

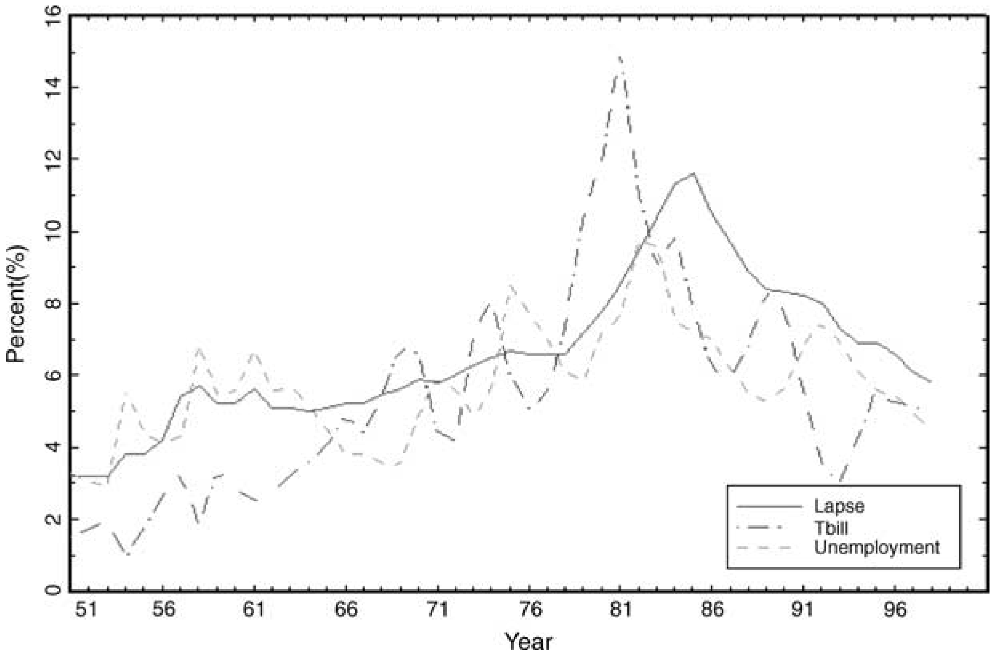
[8] Since the 90-day Treasury bill rates in the database are recorded monthly, we transform the monthly interest rates into the annual rates using the following compounding method:

$$\text{annual interest rate } r = \left(1 + \frac{m_1}{12}\right)\left(1 + \frac{m_2}{12}\right)\ldots\left(1 + \frac{m_{12}}{12}\right) - 1,$$

where $m_i$ denotes the interest rate in month $i$, $i = 1, 2, \ldots, 12$. The unemployment rates are also reported monthly in the database and we use the average of the monthly rates as the unemployment rate for that year.

[9] We also perform the tests based on regressions (2) and (3) to insure that the acceptance of the null hypothesis of a unit root is robust to different specifications for the underlying data-generating process.

**FIGURE 1**
Time Series of the Lapse Rate, Interest Rate, and Unemployment Rate: 1951–1998

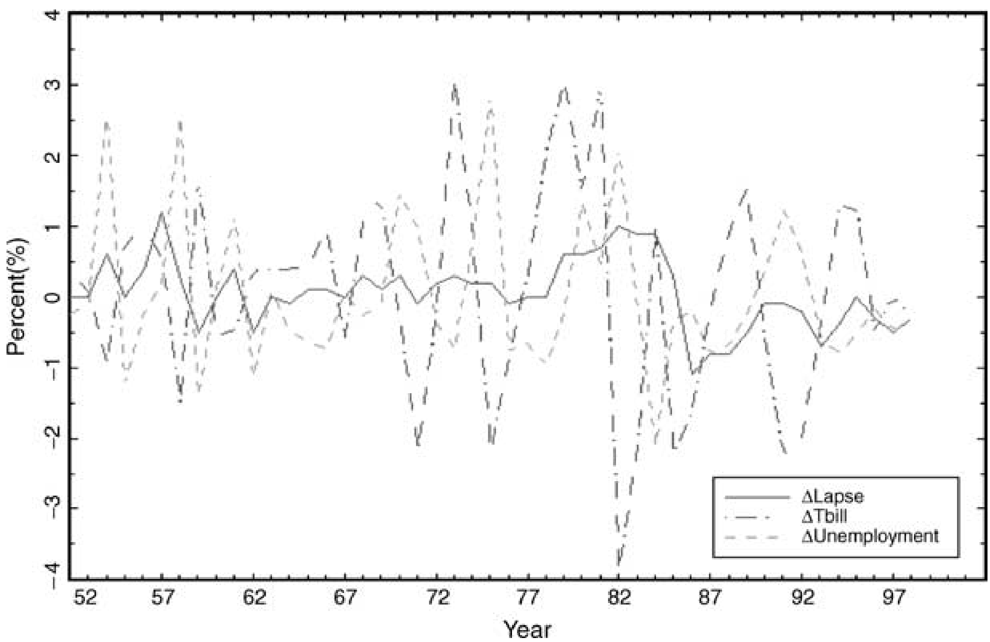



level of three series are not significant at the 5 percent level, implying that the null hypothesis of a unit root cannot be rejected for the lapse rate, interest rate, and unemployment rate. In addition, the corresponding statistics for their first-order differences are significant at the 1 percent significance level and thus suggest rejecting the null hypothesis. Based on these results, we conclude that the lapse rate, interest rate, and unemployment rate follow nonstationary $I(1)$ processes individually.

To conduct the multivariate cointegration tests of Johansen (1991), we first have to determine the order of the error-correction model that in turn depends on the order of the vector autoregression model consisting of the lapse rate, interest rate, and unemployment rate. The optimal model is VAR (2), according to BIC. The results of the maximal eigenvalue test and the trace test are reported in Table 2. Both tests reveal two cointegrating vectors among the lapse rate, interest rate, and unemployment rate.

A nonzero cointegrating vector represents the influence from a long-term force. The cointegrating vector specifies a long-term relation among the levels of lapse rate, interest rate, and unemployment rate. Any deviation from this relation will cause the lapse rate to change and the impact of this deviation will last for a long period of time because it is the levels of variables that cause the lapse rate to change. Since a nonzero cointegrating vector has enduring effects, it represents the influence in the long run. As mentioned above, separating such a long-term relation from its short-term counterpart is the main merit of our cointegration analysis.

**FIGURE 2**
Time Series of First-Order Differenced Lapse Rate, Interest Rate, and Unemployment Rate: 1952–1998




The economic literature offers two explanations for the existence of the cointegration relation. Engle and Granger (1987) described cointegrated variables (lapse rate, interest rate, and unemployment rate in our case) as being in equilibrium when the stationary linear combination of their levels is at its unconditional mean that is usually assumed to be zero. The system is out of equilibrium when this combination of levels (of the lapse rate, interest rate, and unemployment rate) is not zero, which occurs frequently in the real world.[10] However, since the combination is stationary, there is always a tendency for the system to return to equilibrium. Engle and Granger called the nonzero stationary cointegration vector the "equilibrium error." Error-correction models thus imply a world in which a long-term equilibrium relationship exists among the variables (the lapse rate, interest rate, and unemployment rate) in the economy and in which the equilibrium error induces changes in the dependent variable.

Campbell and Shiller (1988) proposed another explanation for the existence of the cointegration relation that emphasized the possibility that the change in the lapse rate results from policyholders' forecasts of the relevant variables in a rational

[10] Possible explanations for the economy being out-of-equilibrium are the imperfections in the market such as transaction cost (in surrendering and repurchasing life insurance), incomplete information (about the economy to individual policyholders), the (policyholders') time lag in obtaining or being aware of the information, and bounded rationality (of policyholders).

**TABLE 1**
Unit Root Tests for the Lapse Rate, Interest Rate, and Unemployment Rate

| Variables | Augmented Dickey-Fuller Test[a] | | |
|---|---|---|---|
| | $H_{20}$ | $H_{30}$ | $H_{40}$ |
| $L_t$ | −2.061 | −1.518 | −0.270 |
| $I_t$ | −2.392 | −2.428 | −0.522 |
| $U_t$ | −2.903 | −2.742 | −0.426 |
| $\Delta L_t$ | −3.166** | – | – |
| $\Delta I_t$ | −5.754** | – | – |
| $\Delta U_t$ | −4.958** | – | – |

[a] Lag length $p$ in the regressions (2), (3), and (4) are selected based on AIC. The purpose of using AIC rather than BIC is to include enough lags of $\Delta y_{it-1}$ to eliminate autocorrelations of the residuals. Lag length in every regression is chosen to ensure that the $Q$-statistic of Ljung and Box (1979) at 8 lags indicates absence of autocorrelation in the residuals. The critical values of the ADF tests are based on MacKinnon (1991).
** Significant at the 1 percent level.

expectation framework. The essential feature of this mechanism is that certain variables (the interest rate and unemployment rate in our case) contain information incorporated into the policyholders' rational expectations about the future variations in the variable of interest (the lapse rate in our case) that follows an integrated process. Policyholders then act on their expectations about future lapse rates according to the information on the historical lapse rate, interest rate, and unemployment rate and the expected lapse rates turn out to be the true ones in the end. In this instance, we do not need an economic theory to delineate the long-term equilibrium relationship among the lapse rate, interest rate, and unemployment rate. What we need, instead, is a rational expectation model that includes the necessary information to predict future lapse rates.

## Discussions about Estimation Results

Based on the results from the maximal eigenvalue test and the trace test, we estimate an error-correction model with two cointegrating vectors of the lapse rate, interest rate, and unemployment rate as shown in the following:

$$\begin{bmatrix}\Delta L_t\\ \Delta I_t\\ \Delta U_t\end{bmatrix} = \begin{bmatrix}\underset{(3.036)}{0.005^{***}}\\ \underset{(-1.783)}{-0.015^{*}}\\ \underset{(3.657)}{0.018^{***}}\end{bmatrix} + \begin{bmatrix}\underset{(-3.530)}{-0.275^{***}} & \underset{(5.337)}{0.153^{***}}\\ \underset{(-0.146)}{-0.036} & \underset{(-1.082)}{-0.160}\\ \underset{(-0.321)}{-0.044} & \underset{(2.489)}{0.207^{**}}\end{bmatrix} \begin{bmatrix}1 & 0 & \underset{(-4.109)}{-1.742^{***}}\\ 0 & 1 & \underset{(-3.491)}{-2.553^{***}}\end{bmatrix} \begin{bmatrix}L_{t-1}\\ I_{t-1}\\ U_{t-1}\end{bmatrix}$$
$$+ \begin{bmatrix}\underset{(3.818)}{0.391^{***}} & \underset{(-1.739)}{-0.087^{*}} & \underset{(-3.530)}{-0.208^{***}}\\ \underset{(-1.048)}{-0.553} & \underset{(1.758)}{0.454^{*}} & \underset{(-0.147)}{-0.044}\\ \underset{(0.712)}{0.212} & \underset{(0.207)}{0.030} & \underset{(0.984)}{0.169}\end{bmatrix} \begin{bmatrix}\Delta L_{t-1}\\ \Delta I_{t-1}\\ \Delta U_{t-1}\end{bmatrix} + \begin{bmatrix}\varepsilon_t^L\\ \varepsilon_t^I\\ \varepsilon_t^U\end{bmatrix}, \tag{6}$$

**TABLE 2**
The Maximal Eigenvalue Test and the Trace Test of Johansen (1991)

| | | The Maximal Eigenvalue Test | | |
|---|---|---|---|---|
| Null | Alternative | $\hat{\lambda}_{max}$ | 95% Critical Value | 90% Critical Value |
| $r = 0$ | $r = 1$ | 29.70* | 21.12 | 19.02 |
| $r \leq 1$ | $r = 2$ | 15.20* | 14.88 | 12.98 |
| $r \leq 2$ | $r = 3$ | 4.44 | 8.07 | 6.50 |
| | | The Trace Test | | |
| Null | Alternative | $\hat{\lambda}_{trace}$ | 95% Critical Value | 90% Critical Value |
| $r = 0$ | $r \geq 1$ | 49.35* | 31.54 | 28.78 |
| $r \leq 1$ | $r \geq 2$ | 19.65* | 17.86 | 15.75 |
| $r \leq 2$ | $r = 3$ | 4.44 | 8.07 | 6.50 |

*Note:* The tests are performed based on $\Delta \mathbf{y_t} = \mu + \gamma \mathbf{y_{t-1}} + C\Delta \mathbf{y_{t-1}} + \varepsilon_\mathbf{t}$.[12] We use Microfit 4.0 to obtain relevant statistics.

where $\mathbf{E} = [\varepsilon^L_\mathbf{t}\ \varepsilon^I_\mathbf{t}\ \varepsilon^U_\mathbf{t}]' \sim N(\mathbf{0}, \hat{\mathbf{\Sigma}})$ and

$$\hat{\mathbf{\Sigma}} = \begin{bmatrix} 8.28 \times 10^{-6} & 6.94 \times 10^{-6} & 2.85 \times 10^{-6} \\ 6.94 \times 10^{-6} & 2.20 \times 10^{-4} & -7.38 \times 10^{-5} \\ 2.85 \times 10^{-6} & -7.38 \times 10^{-5} & 7.03 \times 10^{-5} \end{bmatrix}.$$

The specifications for the individual variables in the error-correction VAR model and the results from their misspecification tests are shown in Table 3. Table 3 shows that the estimated models for the lapse rate, interest rate, and unemployment rate are generally well specified except for the fat tails of the lapse rate and unemployment rate residuals.[11]

The lapse rate equation in the vector error-correction system (6) reveals that changes in the lapse rate are due to two distinct forces: the change in the lagged variables $[\Delta L_{t-1}\ \Delta I_{t-1}\ \Delta U_{t-1}]$ and the levels of the lagged variables $[L_{t-1}\ I_{t-1}\ U_{t-1}]$. The impact on the level of lagged variables can be represented by two cointegration vectors $L_t - 1.742U_t$ and $I_t - 2.553U_t$ denoted as *ECM*1 and *ECM*2, respectively. These two cointegration vectors further imply the existence of a cointegration vector among the lapse rate, interest rate, and unemployment rate: $L_t - 0.556I_t - 0.321U_t$. Therefore, changes in

[11] The fat tail of the lapse rate residuals results from the two residuals being larger than two standard deviations, 0.96 percent in 1956 and −0.62 percent in 1986. Three large unemployment rate residuals occurred in 1954, 1958, and 1975 with sizes of 1.95 percent, 1.95 percent, and 2.0 percent, respectively. Looking closely at the lapse rate and unemployment rate data for these 5 years, we find that there are abrupt changes in their first-order differenced series that cannot be explained reasonably by our error-correction model. Nevertheless, we believe that our model fits the data well in general since Gonzalo (1994) showed that the estimation for cointegration vectors through various methods is consistent even when the errors are non-Gaussian.

**TABLE 3**
Error-Correction Models for the Lapse Rate, Interest Rate, and Unemployment Rate

| | $\Delta L_t$ | | | $\Delta I_t$ | | | $\Delta U_t$ | | |
|---|---|---|---|---|---|---|---|---|---|
| Variables | Coefficients | *t*-value | *p*-value | Coefficients | *t*-value | *p*-value | Coefficients | *t*-value | *p*-value |
| Intercept | 0.005 | 3.036 | 0.004 | −0.015 | −1.783 | 0.082 | 0.018 | 3.657 | 0.001 |
| $\Delta L_{t-1}$ | 0.391 | 3.818 | 0.000 | −0.553 | −1.048 | 0.301 | 0.212 | 0.712 | 0.481 |
| $\Delta I_{t-1}$ | −0.087 | −1.739 | 0.090 | 0.454 | 1.758 | 0.086 | 0.030 | 0.207 | 0.837 |
| $\Delta U_{t-1}$ | −0.208 | −3.530 | 0.001 | −0.045 | −0.147 | 0.884 | 0.169 | 0.984 | 0.331 |
| $ECM1_{t-1}$ | −0.275 | −5.808 | 0.000 | −0.036 | −0.146 | 0.885 | −0.044 | −0.321 | 0.750 |
| $ECM2_{t-1}$ | 0.153 | 5.337 | 0.000 | −0.160 | −1.082 | 0.286 | 0.207 | 2.489 | 0.017 |
| $R^2$ | 0.701 | | | 0.172 | | | 0.464 | | |
| $DW$ | 1.926 | | | 1.918 | | | 2.041 | | |
| $LB-Q$ | $\chi^2$ (6) = 1.765[0.940] | | | $\chi^2$ (6) = 2.033[0.917] | | | $\chi^2$ (6) = 2.862[0.826] | | |
| $ARCH$ | $\chi^2$ (6) = 2.223[0.898] | | | $\chi^2$ (6) = 2.977[0.278] | | | $\chi^2$ (6) = 2.761[0.838] | | |
| *Normal* | $\chi^2$ (2) = 12.286[0.002] | | | $\chi^2$ (2) = 1.890[0.389] | | | $\chi^2$ (2) = 8.273[0.016] | | |
| $FF$ | $\chi^2$ (1) = 0.638[0.425] | | | $\chi^2$ (1) = 0.654[0.419] | | | $\chi^2$ (1) = 0.663[0.415] | | |

$ECM1_{t-1} = L_{t-1} - 1.742U_{t-1}$; $ECM2_{t-1} = I_{t-1} - 2.553U_{t-1}$

the lapse rate could result from the changes in lagged variables and/or a nonzero cointegrating vector.

System (6) also indicates that both the interest rate and unemployment rate have influences on the lapse rate in the long run because the cointegration vectors involve both the interest rate and unemployment rate. Furthermore, the adjustment of lapse rate to deviations from the long-term relationship is only partial. Apart from the long-term relationship, the estimated error-correction model also reveals interesting short-term lapse rate dynamics. The highly significant coefficient of $\Delta U_{t-1}$ in the equation for $\Delta L_t$ as shown in Table 3 suggests that the unemployment rate can influence the lapse rate in the short run. This result is basically consistent with that from Outreville (1990) and supports the emergency fund hypothesis. In contrast, the interest rate is only marginally significant, which is also consistent with the result from Outreville. Therefore, according to the short-term dynamics of lapse rate equation, our results display a higher explaining power of unemployment rate over the interest rate and are generally consistent with the findings of Outreville.

To sum up, changes in the lapse rate depend on the deviation from the cointegration relationship among the lapse rate, interest rate, and unemployment rate, which represents the influence in the long run, as well as the changes in the lagged lapse rate, interest rate, and unemployment rate, which represents the influence in the short run. We find that the emergency fund hypothesis is substantiated in both the long and short run and that the interest rate hypothesis is valid in the long run but obscure in the short run. Our results are thus in line with those of Outreville (1990) and provide a further insight into the long-term relationship among the lapse rate, interest rate, and unemployment rate.

## Impulse Response Analysis

The above statistically significant relations among the lapse rate, unemployment rate, and interest rate do not necessarily imply economically significant effects. We therefore

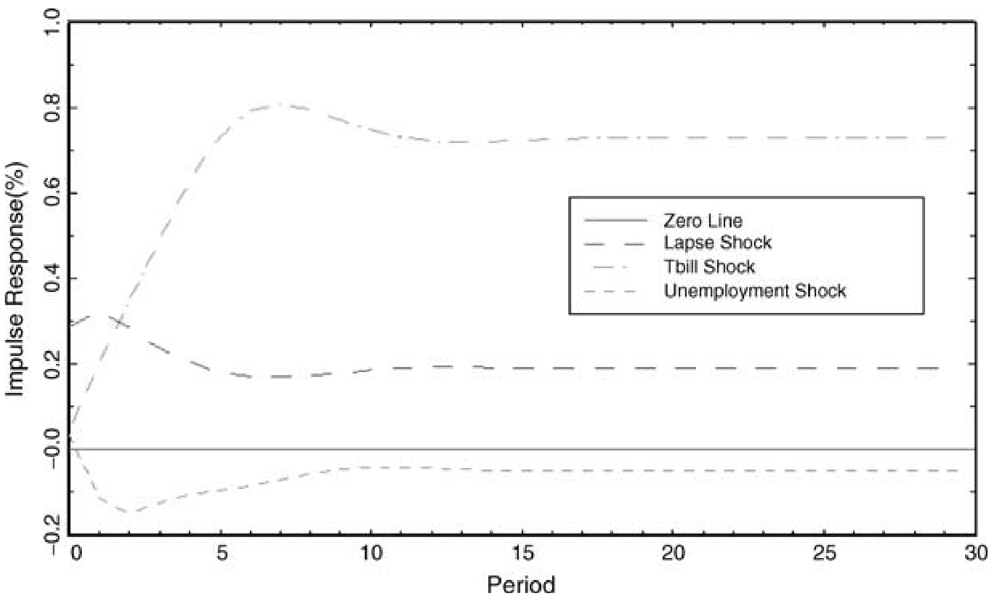
**TABLE 4**
Impulse Response of the Lapse Rate, Interest Rate, and Unemployment Rate (Percent)

| | Lapse Rate | | | Interest Rate | | | Unemployment Rate | | |
|---|---|---|---|---|---|---|---|---|---|
| | Shock Variable | | | Shock Variable | | | Shock Variable | | |
| Period | L | I | U | L | I | U | L | I | U |
| 0 | 0.288 | 0.047 | 0.034 | 0.241 | 1.485 | −0.881 | 0.099 | −0.497 | 0.839 |
| 1 | 0.325 | 0.209 | −0.120 | 0.185 | 1.683 | −0.804 | 0.177 | −0.004 | 0.397 |
| 2 | 0.283 | 0.350 | −0.149 | 0.177 | 1.388 | −0.345 | 0.140 | 0.467 | −0.049 |
| 3 | 0.237 | 0.492 | −0.124 | 0.226 | 1.140 | −0.063 | 0.085 | 0.628 | −0.159 |
| 4 | 0.203 | 0.630 | −0.105 | 0.272 | 1.037 | −0.005 | 0.066 | 0.608 | −0.100 |
| 5 | 0.181 | 0.736 | −0.096 | 0.292 | 1.012 | −0.033 | 0.074 | 0.543 | −0.035 |
| 6 | 0.170 | 0.794 | −0.085 | 0.295 | 1.012 | −0.061 | 0.091 | 0.485 | −0.010 |
| 7 | 0.169 | 0.807 | −0.071 | 0.291 | 1.021 | −0.073 | 0.104 | 0.443 | −0.009 |
| 8 | 0.173 | 0.794 | −0.057 | 0.285 | 1.035 | −0.077 | 0.112 | 0.415 | −0.014 |
| 9 | 0.179 | 0.770 | −0.047 | 0.281 | 1.052 | −0.078 | 0.115 | 0.400 | −0.019 |
| 10 | 0.185 | 0.747 | −0.043 | 0.278 | 1.065 | −0.078 | 0.115 | 0.400 | −0.023 |
| 11 | 0.190 | 0.730 | −0.042 | 0.276 | 1.073 | −0.078 | 0.114 | 0.399 | −0.027 |
| 12 | 0.192 | 0.721 | −0.044 | 0.275 | 1.077 | −0.076 | 0.112 | 0.406 | −0.030 |
| 13 | 0.192 | 0.718 | −0.047 | 0.276 | 1.076 | −0.074 | 0.110 | 0.413 | −0.031 |
| 14 | 0.191 | 0.719 | −0.049 | 0.276 | 1.074 | −0.073 | 0.109 | 0.418 | −0.031 |
| 15 | 0.190 | 0.722 | −0.050 | 0.277 | 1.071 | −0.073 | 0.108 | 0.421 | −0.030 |
| 16 | 0.189 | 0.725 | −0.051 | 0.278 | 1.069 | −0.072 | 0.108 | 0.422 | −0.030 |
| 17 | 0.189 | 0.728 | −0.051 | 0.278 | 1.068 | −0.073 | 0.108 | 0.421 | −0.029 |
| 18 | 0.189 | 0.730 | −0.051 | 0.278 | 1.067 | −0.073 | 0.108 | 0.420 | −0.029 |
| 19 | 0.189 | 0.730 | −0.050 | 0.278 | 1.067 | −0.073 | 0.109 | 0.419 | −0.028 |
| 20 | 0.189 | 0.730 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.419 | −0.028 |
| 21 | 0.189 | 0.730 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.418 | −0.028 |
| 22 | 0.189 | 0.729 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.418 | −0.029 |
| 23 | 0.189 | 0.729 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.418 | −0.029 |
| 24 | 0.189 | 0.729 | −0.050 | 0.278 | 1.069 | −0.073 | 0.109 | 0.418 | −0.029 |
| 25 | 0.189 | 0.729 | −0.050 | 0.278 | 1.069 | −0.073 | 0.109 | 0.418 | −0.029 |
| 26 | 0.189 | 0.728 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.418 | −0.029 |
| 27 | 0.189 | 0.729 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.419 | −0.029 |
| 28 | 0.189 | 0.729 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.419 | −0.029 |
| 29 | 0.189 | 0.729 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.419 | −0.029 |
| 30 | 0.189 | 0.729 | −0.050 | 0.278 | 1.068 | −0.073 | 0.109 | 0.419 | −0.029 |

carry out an impulse response analysis of (6) to examine the responses of lapse rate to random shocks from the interest and unemployment rates.[12] Table 4 and Figure 3 illustrate the impulse responses of the lapse rate to one standard deviation shock in the unemployment rate and interest rate for up to 30 periods. The magnitude of the lapse rate response to the shock from the unemployment rate is close to zero over the

[12] We also perform impulse response analyses for the unemployment and interest rates. The responses of the interest rate and unemployment rate to the random shock of one standard deviation in the lapse rate is miniscule, which corresponds to the insignificant coefficients of lapse rate in both the interest rate and unemployment rate equations in Table 3.

**FIGURE 3**
Impulse Response of the Lapse Rate to the Shocks from Itself, the Interest Rate, and the Unemployment Rate




entire 30 periods. This suggests that an unexpected change in the unemployment rate does not have an economically significant impact upon the future lapse rate.

Compared to the response to the unemployment rate shock, the response of lapse rate to the shock from the interest rate is far more appreciable in terms of both strength and endurance. As shown in Figure 3, the response increases sharply initially, reaches the maximum in the 7th period, and then declines slightly afterward. Specifically, if the interest rate rises unexpectedly about 3 percent, the increase in the lapse rate would be about 0.2 percent after 1 year, reach 0.8 percent after 7 years, descend slightly thereafter and stay at 0.7 percent afterward. This phenomenon implies that the interest rate overwhelms the unemployment rate in influencing the lapse rate, even though the unemployment rate has higher statistical explanatory power than the interest rate in the short run as suggested in the estimated error-correction model. In view of this evidence, we conclude that the interest rate hypothesis is economically more important than the emergency fund hypothesis in explaining the dynamics of the lapse rate.

### Further Analyses

The only significant coefficients in the unemployment rate equation are the intercept term and *ECM*2. This suggests that the lapse rate generates no feedback for the unemployment rate in both the short and long run and that the interest rate affects the unemployment rate only through a long-term mechanism. Similar results can be found in the interest rate equation. The insignificant coefficients of $\Delta L_{t-1}$ on $\Delta I_t$ and $\Delta U_t$ in the interest rate and unemployment rate equations reinforce our confidence in the specification for our error-correction model because the lapse rate intuitively does not play an important role in the working of the economic system.

Based on the insignificance of the lapse rate in the interest rate and unemployment rate equations, we decide to treat the interest and unemployment rates as exogenous $I(1)$ variables in the cointegration system and follow the method developed by Johansen (1991) and Pesaran, Shin, and Smith (2000) to reestimate the error-correction model under the presumption that the interest and unemployment rates are long-term forcing variables as defined by Granger and Lin (1995). We find a highly significant cointegration vector and estimate the corresponding error-correction model as follows:

$$\begin{aligned}\Delta L_t = \underset{(5.570)}{0.006^{***}} \underset{(-5.579)}{-0.278^{***}} \left( L_{t-1} - \underset{(-7.919)}{0.537^{***}} I_{t-1} - \underset{(-2.218)}{0.310^{**}} U_{t-1} \right) \\ + \underset{(3.989)}{0.400^{***}} \Delta L_{t-1} - \underset{(-2.005)}{0.089^{*}} \Delta I_{t-1} - \underset{(3.607)}{0.210^{***}} \Delta U_{t-1} + \zeta_t.\end{aligned} \quad (7)$$

This error-correction model passes relevant misspecification tests, as (6) does. Equation (7) possesses similar features regarding the long- and short-term dynamics of the lapse rate to those implied by (6), except that the interest rate has more significant explanatory power in the short run in (7). Overall, this reestimated error-correction model confirms our previous conclusions and suggests that our results are robust for different specifications of the error-correction model.

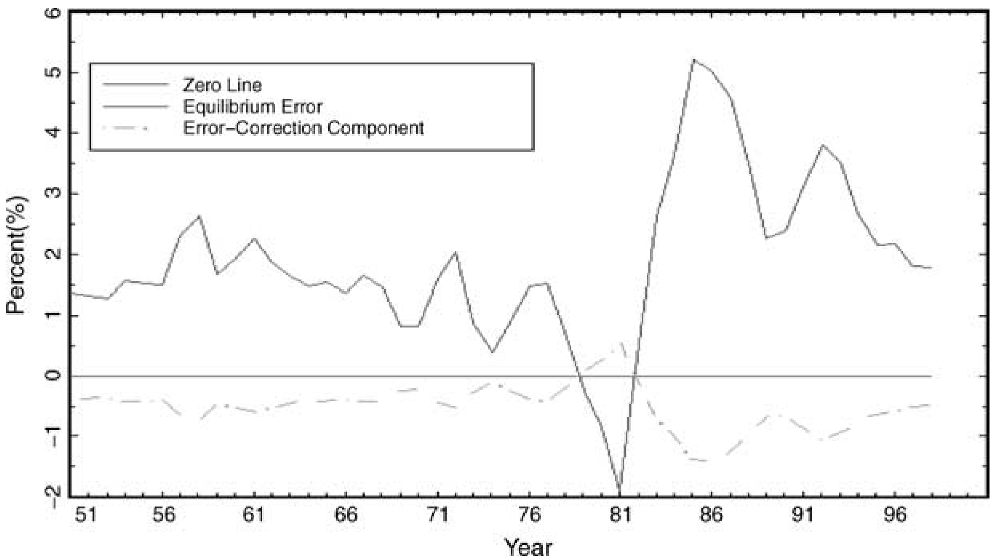
To complete our cointegration analysis, we investigate how the deviation from the long-term cointegration relation contributes to the variation in the lapse rate.[13] Figure 4 depicts the time series of the deviation and the corresponding adjustments of lapse rate during the sample period. We see that the equilibrium error (the nonzero cointegration vector) remained steady around 1.45 percent and the adjustment of the lapse rate was about −0.41 percent until 1978. However, the story was quite different from 1979 through 1987.[14] During this period, the lapse rate was less than its equilibrium level by 1.92 percent in 1981 and then overshot to a sky-high level of 5.21 percent in 1985. Responding with such volatile deviations, the adjustment of the lapse rate increased by 0.53 percent in 1982 and decreased by 1.45 percent in 1986. After 1987, the magnitude of both the equilibrium error and the lapse rate adjustment slid gradually to 1.78 percent and −0.50 percent, respectively. According to these results, we conclude that the long-term component does play an important role in the variation in the lapse rate.

We also decompose the expected lapse rate for time $t$ based on the available information at time $t-1$ into two components: the long-term component ($LC$) and the short-term component ($SC$). By comparing the magnitude of $LC$, $SC$, and the realized change in the lapse rate ($\Delta L$), we may assess the relative importance of the long-term

[13] This analysis is based on Equation (7).

[14] Note that the Federal Reserve announced in October 1979 that it would deemphasize the federal funds rate target and place a higher priority on hitting the money growth rate targets. Some economists dubbed the new procedure implemented in the period from 1979 to 1982 the "monetarist experiment." (Thomas, 1997, p. 460)

**Figure 4**
Time Series of Equilibrium Error and the Corresponding Adjustment of the Lapse Rate




and the short-term forces in terms of their contribution to the variation in the lapse rate. Table 5 shows the sizes of the equilibrium errors, *LC*, *SC*, and $\Delta L$. From 1953 through 1978, $\Delta L$ fluctuated primarily with *SC* rather than with *LC* in the sense that $\Delta L$ and *SC* moved together in the same direction in most years during this period. The correlation coefficient between *SC* and $\Delta L$ was 0.435 whereas that between *LC* and $\Delta L$ was 0.226. In contrast, *LC* dominated its counterpart in determining the variation in $\Delta L$ during the latter period from 1979 to 1998. $\Delta L$ moved almost side by side with *LC* during this period. The correlation coefficient between *LC* and $\Delta L$ was as high as 0.822 that was much higher than 0.314, the correlation coefficient between *SC* and $\Delta L$. In other words, since 1979 *LC* became more important to policyholders in forming their expectations about the future lapse rate.[15] This result again indicates the importance of the long-term component in determining the variation in the lapse rate and also demonstrates the advantages of using the more flexible cointegration method.

## Conclusions and Discussions

Understanding the determinants of the lapse rate is important because policy lapse can have negative impacts on the insurer's profitability and liquidity. Furthermore, policy lapse could cause the cash flow of the insurance policy to be sensitive to the interest rate and significantly change the duration, convexity, and value of the insurance policy. Despite the importance of policy lapse, most insurers do not have a reliable model to specify lapse behavior, especially involving the sensitivity of the lapse rate to the

[15] During the entire sample period, the correlation coefficients between *SC* and $\Delta L$ and between *LC* and $\Delta L$ were 0.341 and 0.720, respectively. This result indicates that *LC* is more important than its counterpart in affecting $\Delta L$.

**TABLE 5**
Decomposition of Expected Lapse Rate (%)

| Year ($t$) | $ECM_{t-1a}$ | $LC_{t-1}$ | $SC_{t-1}$ | $\Delta L_t$ | Year ($t$) | $ECM_{t-1}$ | $LC_{t-1}$ | $SC_{t-1}$ | $\Delta L_t$ |
|---|---|---|---|---|---|---|---|---|---|
| 1953 | 1.326 | −0.369 | 0.034 | 0.000 | 1976 | 0.887 | −0.247 | −0.322 | −0.100 |
| 1954 | 1.266 | −0.352 | 0.006 | 0.600 | 1977 | 1.480 | −0.411 | 0.199 | 0.000 |
| 1955 | 1.561 | −0.434 | −0.232 | 0.000 | 1978 | 1.515 | −0.421 | 0.109 | 0.000 |
| 1956 | 1.511 | −0.420 | 0.185 | 0.400 | 1979 | 0.728 | −0.203 | 0.025 | 0.600 |
| 1957 | 1.491 | −0.415 | 0.129 | 1.200 | 1980 | −0.252 | 0.070 | 0.013 | 0.600 |
| 1958 | 2.311 | −0.643 | 0.390 | 0.300 | 1981 | −0.870 | 0.242 | −0.173 | 0.700 |
| 1959 | 2.618 | −0.728 | −0.282 | −0.500 | 1982 | −1.917 | 0.533 | −0.079 | 1.000 |
| 1960 | 1.663 | −0.462 | −0.055 | 0.000 | 1983 | 0.529 | −0.147 | 0.308 | 0.900 |
| 1961 | 1.915 | −0.532 | 0.027 | 0.400 | 1984 | 2.613 | −0.726 | 0.574 | 0.900 |
| 1962 | 2.248 | −0.625 | −0.033 | −0.500 | 1985 | 3.629 | −1.009 | 0.711 | 0.300 |
| 1963 | 1.864 | −0.518 | −0.003 | 0.000 | 1986 | 5.215 | −1.450 | 0.384 | −1.100 |
| 1964 | 1.628 | −0.453 | −0.050 | −0.100 | 1987 | 5.033 | −1.399 | −0.258 | −0.800 |
| 1965 | 1.467 | −0.408 | 0.026 | 0.100 | 1988 | 4.600 | −1.279 | −0.129 | −0.800 |
| 1966 | 1.543 | −0.429 | 0.139 | 0.100 | 1989 | 3.509 | −0.976 | −0.259 | −0.500 |
| 1967 | 1.356 | −0.377 | 0.107 | 0.000 | 1990 | 2.248 | −0.625 | −0.290 | −0.100 |
| 1968 | 1.652 | −0.459 | 0.041 | 0.300 | 1991 | 2.391 | −0.665 | −0.057 | −0.100 |
| 1969 | 1.459 | −0.406 | 0.083 | 0.100 | 1992 | 3.118 | −0.867 | −0.098 | −0.200 |
| 1970 | 0.829 | −0.230 | −0.070 | 0.300 | 1993 | 3.804 | −1.058 | −0.035 | −0.700 |
| 1971 | 0.823 | −0.229 | −0.167 | −0.100 | 1994 | 3.525 | −0.980 | −0.118 | −0.400 |
| 1972 | 1.582 | −0.440 | −0.052 | 0.200 | 1995 | 2.684 | −0.746 | −0.105 | 0.000 |
| 1973 | 2.036 | −0.566 | 0.178 | 0.300 | 1996 | 2.144 | −0.596 | −0.009 | −0.300 |
| 1974 | 0.895 | −0.249 | −0.001 | 0.200 | 1997 | 2.174 | −0.604 | −0.037 | −0.500 |
| 1975 | 0.391 | −0.109 | −0.160 | 0.200 | 1998 | 1.790 | −0.498 | −0.105 | −0.300 |

*Note:* The decomposition is based on Equation (7). In particular, based on the information at time $t-1$, the expected lapse rate at time $t$ can be decomposed into two components, a long-term component and a short-term component, as follows:

$$E_{t-1}(\Delta L_t) = \mu + LC_{t-1} + SC_{t-1}$$

where $\mu = 0.006$ is the intercept term, $LC_{t-1} = -0.278(L_{t-1} - 0.537I_{t-1} - 0.310U_{t-1})$ is the long-term component, and $SC_{t-1} = 0.400\Delta L_{t-1} - 0.089\Delta I_{t-1} - 0.210\Delta U_{t-1}$ is the short-term component.

[a] $ECM_{t-1} = L_{t-1} - 0.537I_{t-1} - 0.310U_{t-1}$ is the cointegration relation in Equation (7). $LC_{t-1}$ is the long-term component at time $t-1$. $SC_{t-1}$ is the short-term component at time $t-1$. $\Delta L_t$ is the realized change in the lapse rate at time $t$.

interest rate. Insurers have not tracked or organized their lapse data in a manner that allows them to accurately model the lapse rate (Santomero and Babbel, 1997).

Aiming to construct a robust lapse rate model, our article extends the literature by using a more comprehensive method and a longer data period. Although previous studies focused exclusively on the short-term dynamics, our study investigates both the short- and long-term lapse behavior using the cointegration model developed by Engle and Granger (1987). Our sample period spans almost 50 years, but the samples in the literature either missed the important era of interest rate movements in the last two decades or had sampling periods shorter than 10 years.

We find that the influence of the unemployment rate upon the lapse rate in the short run is statistically significant whereas the short-term impact of the interest rate is only marginally significant. This evidence seems consistent with the emergency fund hypothesis as well as with the findings of Outreville (1990). In addition, we discover a long-term relationship among the lapse rate, interest rate, and unemployment rate that is not identified in Outreville's paper. Both the unemployment and interest rates have statistically significant power in explaining long-term behavior of the lapse rate.

We speculate that the long-term causality from the interest and unemployment rates on the lapse rate could occur through two mechanisms. The first mechanism suggested by Engle and Granger (1987) assumes that there is a long-term equilibrium relationship among the lapse rate, interest rate, and unemployment rate and any equilibrium error will be corrected gradually. The causality from the interest and unemployment rates on the lapse rate reflects the partial adjustment of a temporary disequilibrium economy system. On the other hand, Campbell and Shiller (1988) suggested an alternative mechanism, arguing that causality resulted from the optimal decision making process of policyholders in the sense of rational expectation.

We further perform an impulse response analysis to examine the economical significance of the influence of unemployment and interest rates on the lapse rate. We find that the lapse rate responds far more strongly to the random shock from the interest rate than to the shock from the unemployment rate. In other words, the interest rate has a more significant economic impact upon the lapse rate than the unemployment rate. We therefore conclude that the interest rate hypothesis dominates the emergency fund hypothesis in interpreting the dynamics of the lapse rate, although the latter may outweigh the former in terms of statistical significance.

Under the presumption that a long-term equilibrium relationship and/or the rational expectations of policyholders are the underlying forces for the causal relationship among the lapse rate, interest rate, and unemployment rate, we now need an equilibrium model and/or a rational expectation model to accommodate our findings. A promising future research topic is to establish such a model. Another path for future work could be to explicitly compare the forecasting performance of our error-correction model with exogenous $I(1)$ variables with that of Outreville's model. The final, but not the last, interesting research topic is applying our empirical model to quantify the reserve risk of policy issuers with respect to variations in the interest rate.

## REFERENCES

Albizzati, M., and H. Geman, 1994, Interest Rate Risk Management and Valuation of the Surrender Option in Life Insurance Policies, *Journal of Risk and Insurance*, 61: 616-637.

American Council of Life Insurance, 1999, *Life Insurance Fact Book* (Washington, DC: American Council of Life Insurance).

Babbel, D. F., 1995, Asset-Liability Matching in the Life Insurance Industry, in: E. I. Altman and I. T. Vanderhoof, eds., *The Financial Dynamics of the Insurance Industry* (New York: Irwin Professional Publishing).

Black, K., Jr., and H. D. Skipper Jr., 2000, *Life and Health Insurance* (Upper Saddle River, NJ: Prentice Hall).

Briys, E., and F. de Varenne, 1997, On the Risk of Insurance Liabilities: Debunking Some Common Pitfalls, *Journal of Risk and Insurance*, 64: 673-694.

Campbell, J. Y., and R. J. Shiller, 1988, Interpreting Cointegrated Models, *Journal of Economic Dynamics and Control*, 12: 505-522.

Cox, S. H., P. D. Laporte, S. R. Linney, and L. Lombardi, 1992, Single-Premium Deferred Annuity Persistency Study, *Transactions of Society of Actuaries Reports*, 281-332.

Engle, R. F., and C. W. J. Granger, 1987, Co-integration and Error Correction Representation, Estimation and Testing, *Econometrica*, 55: 251-276.

Gonzalo, J., 1994, Five Alternative Methods of Estimating Long-Run Equilibrium Relationships, *Journal of Econometrics*, 60: 203-233.

Granger, C. W. J., and J. L. Lin, 1995, Causality in the Long Run, *Econometric Theory*, 11: 530-536.

Grosen, A., and P. L. Jorgensen, 2000, Fair Valuation of Life Insurance Liabilities: The Impact of Interest Guarantees, Surrender Options, and Bonus Policies, *Insurance: Mathematics and Economics*, 26: 37-57.

Hamilton, J. D., 1994, *Time Series Analysis* (Princeton, NJ: Princeton University Press).

Johansen, S., 1991, Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models, *Econometrica*, 59: 1551-1580.

Ljung, G., and G. Box, 1979, On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 66: 265-270.

MacKinnon, J. G., 1991, Critical Values for Cointegration Tests, in: R. F. Engle and C. W. J. Granger, eds., *Long-run Economic Relationships: Readings in Cointegration* (New York: Oxford University Press).

Otero, J., and J. Smith, 2000, Testing for Cointegration: Power versus Frequency of Observation—Further Monte Carlo Results, *Economics Letters*, 67: 5-9.

Outreville, J. F., 1990, Whole-Life Insurance Lapse Rates and the Emergency Fund Hypothesis, *Insurance: Mathematics and Economics*, 9: 249-255.

Pesaran, M. H., Y. Shin, and R. Smith, 2000, Structural Analysis of Vector Error Correction Models with Exogenous I(1) Variables, *Journal of Econometrics*, 97: 293-343.

Santomero, A. M., and D. F. Babbel, 1997, Financial Risk Management by Insurers: An Analysis of the Process, *Journal of Risk and Insurance*, 64: 231-270.

Schwarz, G., 1978, Estimating the Dimension of a Model, *Annals of Statistics*, 6: 461-464.

Thomas, L. B., 1997, *Money, Banking, and Financial Markets* (Boston, MA: Irwin McGraw-Hill).