MODELING SURRENDER AND LAPSE RATES WITH ECONOMIC VARIABLES

Changki Kim*

ABSTRACT

This paper presents surrender rate models with explanatory variables such as the difference between reference and crediting rates, policy age since issue, financial crises, unemployment and economy growth rates, and seasonal effects. The logit function and the complementary log-log function are used in modeling surrender rates.

This paper shows that the logit model and the complementary log-log model generally perform better than the existing surrender rate models such as the arctangent model. It also shows that the surrender rate models are different according to insurance policy types, and it finds proper surrender rate models for four insurance groups: protection plans, education plans, endowment, and annuities.

1. INTRODUCTION

When we consider asset/liability management (ALM), we should understand the characteristics of assets and liabilities to capture their cash flow profiles and interest rate sensitivities. Interest rate movements affect the cash flows of assets and liabilities, because interest-rate-dependent surrender and lapse rates and prepayment rates increase the sensitivity of the duration mismatch. When interest rates go down, the surrender and lapse rates of liabilities also go down, but the prepayment rates of assets go up. In this case the duration of liability cash flows increases, and the duration of asset cash flows decreases (see Followill 1998; Fabozzi and Fabozzi 2001). A lower lapse rate drives up the market value of the liability curve, and a high prepayment rate limits reinvestment opportunity, driving down the market value of assets. When the interest rates increase, the lapse rates also increase because of the increased gap between the interest rates and the expected product crediting rates.[1] The market value of a liability tends to decrease under high interest rates. But the convexity may be increased and affect the market value of the liability curve. For analysis of prepayment and surrender rate impacts on the duration, convexity, and value of an asset/liability, see Green and Shoven (1986), Bierwag (1987), Bierwag, Corrado, and Kaufman (1990), Asay, Bouyoucos, and Marciano (1993), Zenios ed. (1993), De Toldi, Gourieroux, and Monfort (1995), Miyazaki and Saito (1999), Carling, Jacobson, and Roszbach (2001), and Panjer et al. (1998). A recent discussion of surrender rate impacts on asset/liability management can be found in Kim (2005).

Modeling appropriate interest-rate-sensitive surrender and lapse rates is essential in managing the assets and liabilities of insurance companies. Even though a few research papers on the interest sensitivity of the cash flows exist, they usually focus on the asset side. For example, in Pesando (1974), the cash flow analysis considers the prepayment rate impacts only. But the interest sensitivity of cash flows through surrender rate fluctuations is a kind of "dual problem" to that through prepayment rate fluctuations.[2] So it is

* Changki Kim, PhD, is an Instructor in the Department of Mathematics, University of Texas, RLM 8.100, 1 University Station C1200, Austin, TX 78712-0257, changki@math.utexas.edu.

[1] Miyazaki and Saito (1999) mention that as spot interest rates go up, the duration of liabilities becomes short with anticipation of early withdrawals.

[2] For examples of prepayment models, see Kang and Zenios (1992), Sherris (1994), Stanton (1995), De Toldi, Gourieroux, and Monfort (1995), Abraham and Theobald (1997), Hall (2000), and Carling, Jacobson, and Roszbach (2001).

important to consider surrender rate impacts on cash flows using proper surrender rate models.

An old example of lapse studies is Richardson and Hartwell (1951), which considers the causes of high lapse rates and the characteristics of businesses with high or low persistency. They consider a number of factors that affect lapse rates such as income, occupation, sex, age, previous insurance, and premium frequency. They also mention that economic disasters affect most violently the terminations at durations subsequent to the second year. But they are not concerned with the mathematical techniques or lapse rate modeling. Buck (1960) derives a few results from a special study of first-year lapse rates, identifies attributes of policies, policyholders, and agents that influence lapse rates, and demonstrates ways to improve the persistency of new business. A lapse study can be found in Brzezinski (1975). Outreville (1990) tests the hypothesis that cash values are utilized by policyholders as an emergency fund. Bluhm (1993) considers a dynamic methodology for adjusting for deviations of actual lapse rates from expected lapse rates. Sharp (1996) considers lapses and termination assumptions in reserve calculations.

This paper models surrender and lapse rates with a few economic variables using the logit function and the complementary log-log function.

2. MODELING SURRENDER AND LAPSE RATES OF KOREAN INSURANCE POLICIES

First, let us summarize a few existing surrender rate models currently used by insurance companies:

Arctangent model: $q_s = a + b \text{ arctangent}(m\Delta - n)$
Parabolic model: $q_s = a + b \text{ sign}(\Delta)\Delta^2$
Modified parabolic model: $q_s = a + b \text{ sign}(\Delta)\Delta k + c^{(Cr(t-1)-CR(t))^j}$
Exponential model: $q_s = a + b \exp(m\, CR/MR)$
New York State Law 126: $q_s = a + b \text{ sign}(\Delta)\Delta k - c\, [(AV - CSV)/AV]$,

where

q_s: monthly surrender rate
a, b, c, m, n, j, k: coefficients
Δ: reference market rate − crediting rate − surrender charges
CR: crediting rate
MR: reference market rate
AV: account value of the policy
CSV: cash surrender value
sign( ): +1 if ( ) is positive, −1 if ( ) is negative.

Note that all of the existing surrender rate models are functions only of interest rates. But this investigation shows that the surrender behaviors of insurance policyholders are very complicated and even unreasonable sometimes.[3]

Because the movements of prepayment rates mirror those of surrender rates, we may refer to the factors that explain prepayments. Richard and Roll (1989) consider a prepayment model with explanatory variables such as refinancing incentive, seasoning or age of the mortgage, month of the year (seasonality), and premium burnout. Many factors can influence prepayments. Browne, Carson, and Hoyt (1999) show that economic and market variables, such as interest rates, personal income, unemployment, stock market, and real estate returns, are important predictors of life-health insurer failure rates.

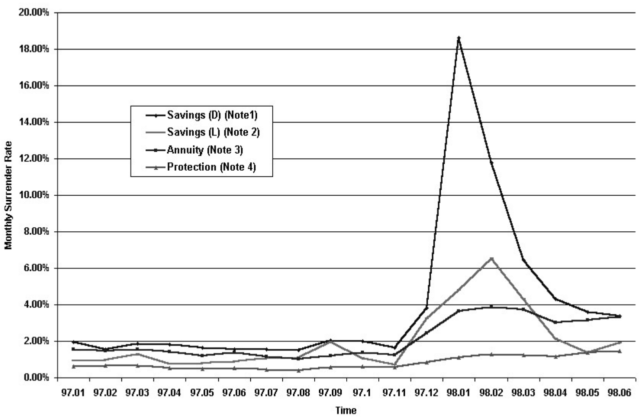
Many factors affect surrender rates as well, such as the difference between reference market rates and policy crediting rates, seasonal effect, age and gender of clients, economy growth rates, foreign exchange rates, inflation rates, policy age since issue date, and unemployment rates. All these factors play an important role in determining the surrender rates. But note that the unemployment rates and increased and volatile interest rates greatly influenced the surrender rates during the financial crises in Korea.

Figure 1 displays the monthly surrender rates experienced in a Korean life insurance company. The policyholder behaviors for the four groups of products with different interest rate sensitivities are clearly shown. The surrender rates show a sudden peak during the financial crises, from December 1997 to December 1998. Note that the savings accounts and annuities show much higher surrender rates than the protection plans.

Figure 2 shows the sudden increase in the market interest rates and the difference between the reference market rates and the policy crediting

[3] Also, for mortgage-backed securities, Green and LaCour-Little (1999) show that borrower prepayment behaviors appear to be highly irrational sometimes.

Figure 1
Monthly Surrender Rates of Four Product Groups

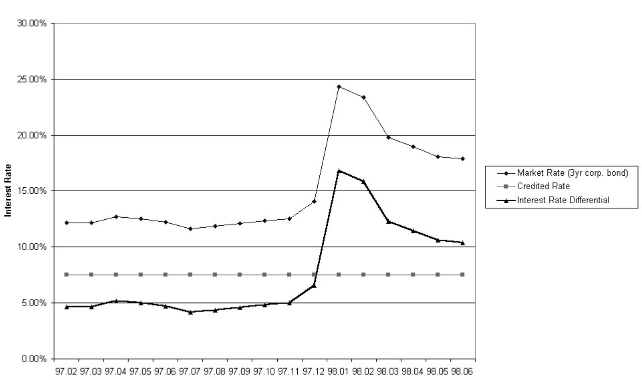


Notes: Savings (D) are mostly deposit accounts with short terms. Surrender rates are very high during financial crises with a maximum of 19%. Savings (L) are long-term deposit accounts with maximum surrender rates of 6.3%. Annuities show almost 4% surrender rates during the financial crises. The bottommost line shows the surrender rates of the traditional protection plans.

rates during the financial crises. Note that the crediting rates do not change much during the financial crises. One reason is that most of the policies guarantee fixed crediting rates at many insurance companies in Korea. After the financial crises they have been trying to develop interest-sensitive policies and hedging strategies for the interest rate risks. We see in Figures 1 and 2 that interest rate fluctuation is an important factor in determining the surrender rates.

Figure 3 shows the unemployment rates and the economy growth rates. Note that the unemployment rates soared and the economy growth rates showed a sharp decline during the financial crises. The surrender rates also skyrocketed during the financial crises, so it is possible to conjecture that the surrender rates are dependent not only on interest rates but also on other factors such as unemployment rates, economy growth rates, and seasonal effects.[4]

Figure 2
Reference Market Rates and Policy Crediting Rates



To set up proper surrender rate models, first consider the process of surrender rate generation (see Figure 4). The total insurance-related data of a Korean company are stored in an IBM mainframe computer. Face amount data for each product group were extracted with monthly observations maintained from 1997 to 2000. For explanatory purposes, we shall select the four product groups: protection plans, endowment, education plans, and annuities with more than 1,000,000 policy holders per product.[5] Then we shall calculate the monthly weighted average of the lapse rates according to the policy ages, from 1 to 10 years.

---

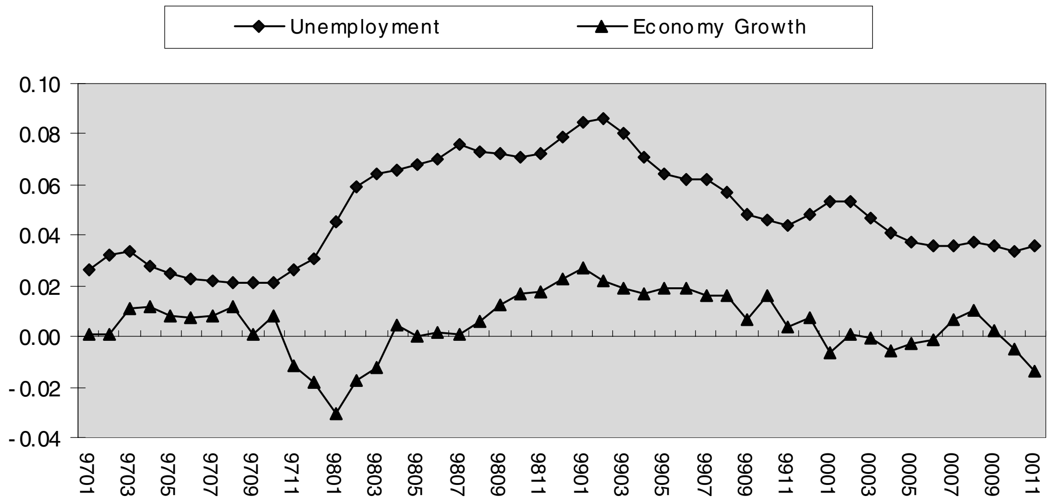
[4] As a dual problem to lapse behavior factors, a few studies look at factors of mortgage prepayment behaviors. In Böheim and Taylor (2000), negative financial surprises such as divorce or loss of employment are considered to be important factors in meeting borrowers' housing costs. Archer, Ling, and McGill (1997) investigate the effect of household financial, demographic, and location characteristics on mortgage termination behavior. Austin (1995) examines mortgage prepayment rates for white, black, and Hispanic borrowers. De Toldi, Gourieroux, and Monfort (1995) introduce a model taking into account for seasonal effects. Richard (1991) uses an empirical approach to model prepayments with factors such as refinancing incentive, age of the mortgage or seasoning, month of the year or seasonality, and premium burnout.

For insurance products, Dukes and Macdonald (1980) consider the relationship between mortality and withdrawals, but this is not attempted in the model here. Bluhm (1982) mentions that lapse rates are much higher for a given block of health policies than for life policies. But those insured who are aware of their impaired status (meaning that they have higher claim expectations) logically have a lower lapse rate than healthy policyholders, at least during early durations, when lapse rates are higher.

[5] Protection plans contain the traditional whole life, term life, and interest-indexed whole life (the crediting rates depend on the market interest rates). Endowments are mostly the traditional endowment insurance. Education plans have survival benefits when the insured's dependents (students) survive at some time (such as entrance to college) as well as death benefits. Annuities contain the traditional whole life annuities and interest-indexed annuities (the crediting rates depend on the market interest rates).

The data used in modeling surrender rates are from Korean insurance business, so there may be differences in the explanatory variables in modeling for other countries.

Figure 3
Unemployment Rates and Economy Growth Rates

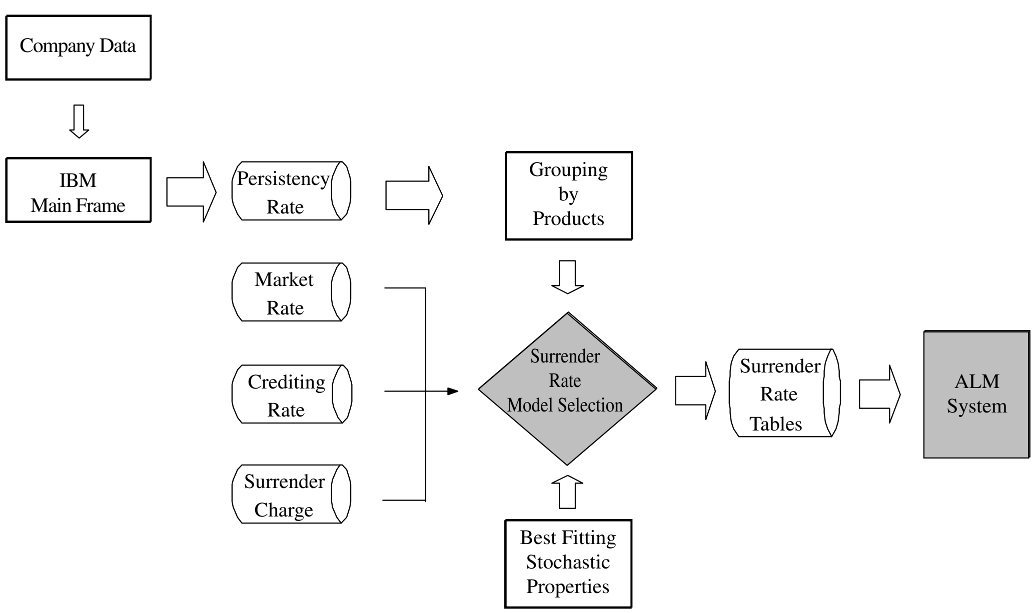


Persistency rates were calculated as the ratio of the current period face amount to the previous period face amount for each product group. The monthly surrender rates were produced for each product group from 1997 to 2000.

At the same time, reference market rates, policy crediting rates, and surrender charges were collected and input to the dynamic surrender and lapse rate generator. We shall select the best surrender models reflecting the relationship between the real surrender rates of the product and other input data. The criteria for selecting the best model are root mean squared error (RMSE) and mean absolute percentage error (MAPE), defined in Sections 3.1 and 3.2.

From the selected best dynamic surrender and lapse rate model, the future surrender and lapse rates are produced according to the projected fu-

Figure 4
Process of Surrender Rate Generation



ture market interest rates, policy crediting rates, and surrender charges. The produced surrender and lapse rates are stored in the surrender rate tables and will be used for the cash flow projection of assets and liabilities in an ALM system.

2.1 Variables and Assumptions

Next we shall determine the proper models for the surrender and lapse rates with a few explanatory variables such as the difference between reference and crediting rates, policy age since issue, financial crises, unemployment rates, economy growth rates, and seasonal effects. The explanatory variables and assumptions used in modeling surrender rates are summarized in Table 1.

For seasonal effects, we shall investigate the surrender rates from January to November. We shall consider the financial crises period because the surrender rates skyrocketed during this period and use dummy variable 1 during the period from December 1997 to December 1998, and 0 elsewhere.

Table 1

Explanatory Variables

Variable	Contents	Memo
BASEYM	Year, month of data	
DIFFLAG0	Difference of rates	Market rate—crediting rate at current time
DIFFLAG2	″	Market rate—crediting rate 2 months ago
DIFFLAG4	″	Market rate—crediting rate 4 months ago
DIFFLAG6	″	Market rate—crediting rate 6 months ago
DIFFLAG8	″	Market rate—crediting rate 8 months ago
DIFFLAG10	″	Market rate—crediting rate 10 months ago
DIFFLAG12	″	Market rate—crediting rate 12 months ago
POL-AGE	Policy age	Average policy age since issue
UNEMPLOY	Unemployment rates	
E-GROWTH	Economy growth rates	
FIN-CRISES	Financial crises period under IMF control	Period from 1997.12 to 1998.12 Dummy variable = 1 during the period
MONTH1	January	Dummy variable = 1 on current month
MONTH2	February	″
MONTH3	March	″
MONTH4	April	″
MONTH5	May	″
MONTH6	June	″
MONTH7	July	″
MONTH8	August	″
MONTH9	September	″
MONTH10	October	″
MONTH11	November	″
SUR_RATE	Real surrender rate	Dependent variables

2.2 Surrender Rate Models

We shall use the logit link function and the complementary log-log (CLL) link function.[6] For modeling programs in SAS, the generalized linear models, procedure GENMOD, logistic regression models, and procedure LOGISTIC are used.[7]

The logit function has the form

$$\ln\left(\frac{q_s}{1 - q_s}\right) = \beta_0 + \beta_1 V_1 + \cdots + \beta_n V_n, \quad (2.1)$$

and the CLL function is of the form

$$\log(-\log(1 - q_s)) = \beta_0 + \beta_1 V_1 + \cdots + \beta_n V_n, \quad (2.2)$$

where q_s is the surrender rate, $\{\beta_i; i = 1, \ldots, n\}$ are the coefficients to be estimated, and $\{V_i; i = 1, \ldots, n\}$ are the explanatory variables.

2.3 Significance of Each Explanatory Variable

We shall check the significance of each explanatory variable with logistic regression analysis. In Table 2 the p-value for all test statistics χ^2 is less than 0.0001. Since the p-value is less than 1% or 5%, each variable has its own significance for surrender rates.

For the difference between the reference market rates and the crediting rates, the estimated parameters are all positive numbers. So the surrender rates go up as the difference increases.

We see that each interest rate difference variable has its own effects on the surrender rates. In

[6] There are many examples in which logit functions are used for financial data analysis. Hall (2000) compares logit analysis of data to the results from his prepayment model. Pinder (1996) demonstrates how multinomial logit models can be used in a decision analysis framework to estimate expected monetary value. Kolari et al. (2002) use the parametric approach of logit analysis to predict large commercial bank failures. See also Johnsen and Melicher (1994) and Lo (1986).

[7] For programming with SAS, see Allison (1999), and SAS Institute (1999). For a few books on generalized linear models, see Agresti (1996, 2002), Harrell (2001), Kutner, Nachtschiem, and Wasserman (1996), McCullagh and Nelder (1989), Firth (1991), and McCulloch and Searle (2000).

Table 2
Significance of Explanatory Variables

Variables	Logit Link Function			CLL Link Function		
	Parameter	Std Error	Chi-Square	Parameter	Std Error	Chi-Square
DIFFLAG0	5.6600	0.0037	2,380,693	5.5888	0.0036	2,391,882
DIFFLAG2	6.9440	0.0036	3,693,589	6.8517	0.0036	3,716,484
DIFFLAG4	6.6217	0.0037	3,221,988	6.5291	0.0036	3,237,336
DIFFLAG6	6.5901	0.0037	3,121,563	6.4988	0.0037	3,136,179
DIFFLAG8	5.7086	0.0038	2,226,700	5.6308	0.0038	2,234,024
DIFFLAG10	4.3584	0.0040	1,205,621	4.2986	0.0039	1,206,910
DIFFLAG12	3.0710	0.0041	552,153	3.0297	0.0041	551,929
POL-AGE	−0.1076	0.0001	3,192,912	−0.1066	0.0001	3,196,557
UNEMPLOY	13.4398	0.0086	2,438,932	13.3027	0.0085	2,440,284
E-GROWTH	−5.9912	0.0139	186,882	−5.9436	0.0137	187,356
FIN-CRISES	0.7662	0.0003	5,112,065	0.7578	0.0003	5,122,041
MONTH1	0.1296	0.0006	51,236.9	0.1283	0.0006	51,292.5
MONTH2	0.1193	0.0006	43,069.9	0.1181	0.0006	43,112.8
MONTH3	0.1235	0.0006	46,335.6	0.1223	0.0006	46,383.5
MONTH4	−0.0414	0.0006	4,574.70	−0.0410	0.0006	4,573.22
MONTH5	−0.0356	0.0006	3,386.19	−0.0352	0.0006	3,385.25
MONTH6	0.0108	0.0006	326.21	0.0107	0.0006	326.24
MONTH7	−0.0507	0.0006	6,811.54	−0.0503	0.0006	6,808.86
MONTH8	−0.0667	0.0006	11,621.2	−0.0661	0.0006	11,615.2
MONTH9	−0.0652	0.0006	11,127.1	−0.0646	0.0006	11,121.5
MONTH10	−0.1009	0.0006	25,883.8	−0.1000	0.0006	25,863.9
MONTH11	−0.1191	0.0006	35,476.5	−0.1180	0.0006	35,444.7

particular, the difference of interest rates two months ago is significant with the large parameter estimate of 6.9440 (logit) and 6.8517 (CLL). So we may guess that the interest rates as of two months ago are highly influencing the surrender behaviors of the policyholders. The policyholders observe the interest rate movements for two to about six months and decide to surrender their policies. The estimated parameter for the policy age since issue is negative. So the surrender rates decrease as the policy age increases.

The positive parameter for the unemployment rates indicates that surrender rates go up when the unemployment rates increase. It is very significant to take the unemployment rates into account as an explanatory variable in modeling surrender rates considering the relatively high parameter estimate of 13.4398 (logit) and 13.3027 (CLL).

The parameter for the economy growth rates is negative, and we may think that the surrender rates go down under good economy conditions. The positive parameter for the dummy variable, financial crises under IMF control, means that the surrender rates can increase when unexpected economic or financial shocks occur.

Note that the parameters for January, February, March, and June are positive and the others are negative, but all are small. Thus season has little effect on surrender behaviors.

3. REDUCED MODELS AND MODEL COMPARISON

We shall follow three steps to find the most appropriate reduced model with the least number of explanatory variables. The first step is to select a few significant explanatory variables. The second step is to set up reduced models with the selected variables. The third step is to transform the policy age. The policy age is transformed because the surrender rates increase during the early years and decrease as time passes.

We shall also compare the three models, the arctangent, logit, and CLL models, and choose the most appropriate one for each insurance group.[8]

[8] For an example of a comparison of models for pricing mortgage-backed securities, see Dunn and McConnell (1981). Hall (2000) compares logit analysis to his modeling results.

Table 3
Model Fit Statistics for Full and Reduced Models (Protection Plan)

Criterion	Intercept Only	Intercept and Covariates
AIC	48,471,181	47,501,084
SC	48,471,198	47,501,490
−2 log *L*	48,471,179	47,501,038

	Model Fit Statistics			Increase (from Full Information Model)		
Deleted Variables	AIC	BIC	−2 × log *L*	AIC	BIC	−2 × log *L*
Interest difference	62,638,015	62,638,303	62,637,983	15,136,931	15,136,813	15,136,945
Policy age	47,551,716	47,552,105	47,551,672	50,632	50,615	50,634
Unemployment	47,509,734	47,510,123	47,509,690	8,650	8,633	8,652
Economy growth	47,501,165	47,501,554	47,501,121	81	64	83
Financial crises	47,501,149	47,501,538	47,501,105	65	48	67
Seasonal	47,561,757	47,561,952	47,561,735	60,673	60,462	60,697

3.1 Step 1: Selecting Explanatory Variables

We want to delete the variables one by one from the least significant until we get a reduced model. It is a kind of backward elimination method.[9] As a criterion for the selection of variables, let us use the $-2 \times \log$ likelihood function ($-2 \times \log L$), Akaike information criterion (AIC), and Bayesian information criterion (BIC), as well as the Schwartz criterion (SC).[10] We shall begin with the protection plans.

Table 3 shows the differences between the full information model fit statistics and reduced model fit statistics. From the second part of Table 3, the increased amounts indicate the relative significance and fit: that is, the deleted variables make contributions to the fit of the reduced model compared with the full information model as much as the increased amount.

Comparing the variables in Table 3, the relative contributions of each variable to the model are ranked as interest rate differences, policy age since issue date, unemployment rates, seasonal effects, economy growth rates, and financial crises.

Note that the financial crises and the economy growth rates influence very little the surrender rates with the increased amount of about 100. But the unemployment rates are affecting notably the surrender rate behaviors of the policyholders.

The numbers in Table 4 show the changed model fit statistics (AIC, BIC, $-2 \times \log L$) as we delete the variables one by one in each step. Figure 5 shows the decreased fit as each step is taken. Note that the fit is reduced significantly after step 3. So let us delete the first three variables, which have less significant contributions.

We find that the interest rate differences and policy age since issue are the most important factors, and the unemployment rates are also important in modeling the surrender rates of protection plans. For the three explanatory variables, we have p-values less than 0.0001 and may conclude that it is reasonable to estimate only the three parameters.

The same analysis has been done for the other three insurance groups: endowments, education plans, and annuities. For endowments we can select interest rate differences and policy age. For education plans, we can select three explanatory variables: unemployment rates, policy age, and interest rate differences. For annuities, we can

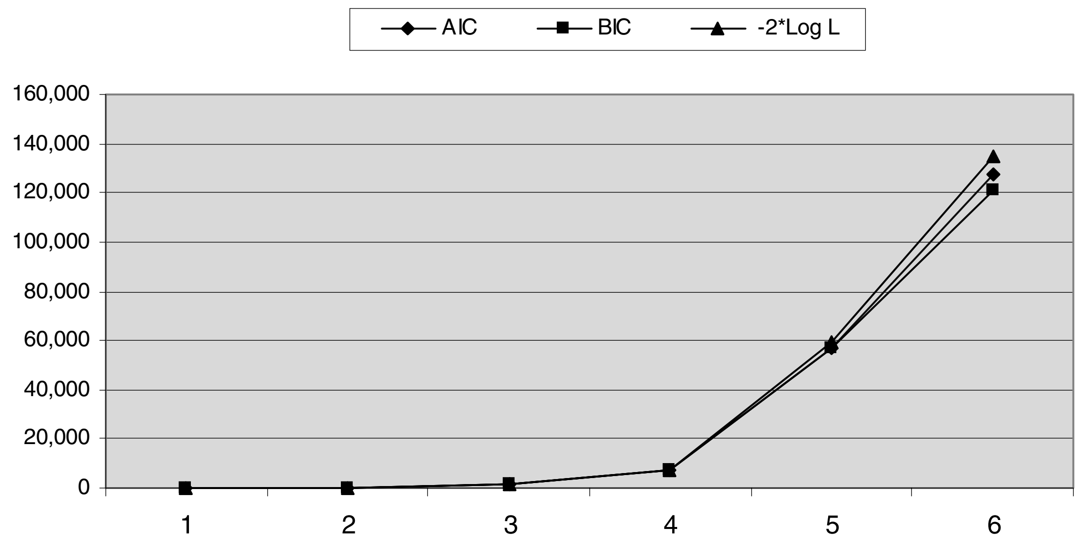
Table 4
Model Fit Statistics Changed (Protection Plan)

Step	Deleted Variable	AIC	BIC	−2 × log *L*
1	Financial crises	65	48	67
2	Economy growth	60	42	62
3	Seasonal	1,658	1,640	1,660
4	Unemployment	7,490	7,472	7,492
5	Policy age	56,883	56,866	56,885
6	Interest difference	126,766	126,749	126,768

[9] See Weisberg (1985).

[10] $AIC(r) = -2 \times \log L + 2 \times r = \log\{SSE(r)/n\} + (2/n) \times (r + 1)$, where r is the number of explanatory variables, n is the number of observation, and SSE stands for sum of squared errors. $BIC(r) = \log\{SSE(r)/n\} + \{(\log n)/n\} \times (r + 1)$. $SC = -2 \times \log L + r \times \log n$. For further explanation, see Hamilton (1994).

Figure 5
Model Fit Statistics Changed (Protection Plan)



select policy age, interest rate differences, unemployment rates, and economy growth rates as the explanatory variables.

3.2 Step 2: Reduced Models

The second step is to set up reduced models with the selected variables from step 1. Three tables show the protection plan. Tables 5 and 6 show the estimated parameters for the selected variables from the logit and CLL models. Table 7 shows the estimated errors for the arctangent, logit, and CLL models and compares the models by the differences of the estimated errors between the arctangent and logit models and the arctangent and CLL models according to the policy age since issue.[11]

For comparison purposes, RMSE and MAPE are defined as follows:

$$\text{RMSE} = \sqrt{\frac{\Sigma(y_i - \hat{y}_i)^2}{n}}, \qquad (3.1)$$

and

[11] For mortgage prepayment models, Hall (2000) compares a conventional logit analysis to the results from the model used in the paper.

Table 5
Parameter Estimates with Logit Model (Protection Plan)
Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	−5.1505	0.00343	2,259,884.65	<.0001
DIFFLAG0	1	4.1663	0.0222	35,229.3910	<.0001
DIFFLAG2	1	2.0239	0.0212	9,109.4006	<.0001
DIFFLAG4	1	0.1111	0.0183	36.8232	<.0001
DIFFLAG6	1	1.9529	0.0177	12,221.3251	<.0001
DIFFLAG8	1	2.4138	0.0168	20,731.7826	<.0001
DIFFLAG10	1	0.7572	0.0178	1,803.3140	<.0001
DIFFLAG12	1	1.2829	0.0231	3,083.9757	<.0001
POLICY AGE	1	−0.0379	0.000165	52,827.8142	<.0001
UNEMPLOY	1	6.7841	0.0785	7,466.5339	<.0001

Table 6
Parameter Estimates with CLL Model (Protection Plan)
Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	−5.1522	0.00341	2,286,440.11	<.0001
DIFFLAG0	1	4.1409	0.0220	35,409.0132	<.0001
DIFFLAG2	1	2.0070	0.0210	9,131.0851	<.0001
DIFFLAG4	1	0.1130	0.0181	38.8195	<.0001
DIFFLAG6	1	1.9361	0.0175	12,263.3299	<.0001
DIFFLAG8	1	2.3961	0.0166	20,885.3584	<.0001
DIFFLAG10	1	0.7584	0.0177	1,844.1119	<.0001
DIFFLAG12	1	1.2770	0.0229	3,102.0849	<.0001
POLICY AGE	1	−0.0377	0.000164	53,001.7031	<.0001
UNEMPLOY	1	6.7346	0.0781	7,444.7539	<.0001

$$\text{MAPE} = \frac{1}{n}\sum \frac{|y_i - \hat{y}_i|}{y_i}, \tag{3.2}$$

where y_i is the i-th real value, $\hat{y}_i$ is the i-th predicted value, and n is the sample size.

We define the terminologies used in Table 7. RMSE1 is RMSE of the arctangent model, RMSE2 is RMSE of the logit model, and RMSE3 is RMSE of the CLL model. MAPE1, MAPE2, and MAPE3 represent the MAPE of the arctangent, logit, and CLL models, respectively.

RMSEGAP1 denotes RMSE1-RMSE2, so the logit model is better than the arctangent model if RMSEGAP1 is positive. RMSEGAP2 is RMSE1-RMSE3, so the CLL model is better than the arctangent model if RMSEGAP2 is positive. MAPEGAP1 is MAPE1-MAPE2, and the logit model is better than the arctangent model if MAPEGAP1 is positive. MAPEGAP2 is MAPE1-MAPE3, and the CLL model is better than the arctangent model if MAPEGAP2 is positive.

For protection plans, the estimated errors of the arctangent model are less than those of the logit and CLL models for one and three years of policy age. But the overall estimated errors of the logit and CLL models are smaller than those of the arctangent model. We may conclude that the logit and CLL models are better than the arctangent model for protection plans. The signs of the estimated parameters look reasonable in explaining the relationship between the surrender rates and the explanatory variables.

The same analysis has been done for endowments, education plans, and annuities. For endowments, we cannot conclude that the logit model or the CLL model is better than the arctangent model. Even when we add the unemployment rates and the financial crises to the logit and the CLL models, we do not have enough evidence that one model is better than the other.

Table 7
Estimated Errors and Comparison of Models (Protection Plan)

Time	RMSE1	RMSE2	RMSE3	MAPE1	MAPE2	MAPE3	RMSEGAP1	RMSEGAP2	MAPEGAP1	MAPEGAP2
0.5	0.00835	0.00691	0.00691	0.26237	0.28027	0.28025	0.00144	0.00144	−0.0179	−0.01788
1.5	0.00434	0.00204	0.00204	0.20047	0.13811	0.13795	0.0023	0.0023	0.06236	0.06251
2.5	0.00467	0.00328	0.00328	0.25927	0.38257	0.38254	0.00139	0.00139	−0.1233	−0.12327
3.5	0.00389	0.00279	0.00279	0.22961	0.21769	0.21756	0.0011	0.0011	0.01192	0.01205
4.5	0.00353	0.00233	0.00233	0.19762	0.15368	0.15355	0.0012	0.0012	0.04394	0.04407
5.5	0.00369	0.00154	0.00154	0.18858	0.0927	0.0926	0.00215	0.00215	0.09588	0.09598
6.5	0.00375	0.00199	0.00199	0.21342	0.13171	0.13168	0.00177	0.00177	0.0817	0.08174
7.5	0.00432	0.00142	0.00142	0.22635	0.09129	0.09131	0.0029	0.0029	0.13506	0.13504
8.5	0.00649	0.00211	0.00211	0.2865	0.11086	0.1109	0.00438	0.00438	0.17564	0.1756
9.5	0.00882	0.00457	0.00457	0.33834	0.1912	0.19123	0.00425	0.00424	0.14715	0.14711

Table 8
Model Fit Statistics according to Transformed Policy Age (Protection Plan)

$(policy\ age)^{1/n}$	$-2 \times \log L$
$n = 1$	47,519,423
$n = 2$	47,484,337
$n = 3$	47,470,131
$n = 4$	47,462,804
$n = 5$	47,458,388
$n = 6$	47,455,447
$n = 7$	47,453,354
$n = 8$	47,451,788
$n = 9$	47,450,575
$n = 10$	47,449,606

Notes: Formula = $-2 \times \log L$, log(*policy age*) = 47,441,042, $-1/(policy\ age)$ = 47,390,559.

For education plans, we can conclude that the logit and CLL models are much better than the arctangent model. For annuities, we cannot conclude that the logit or CLL model is better than the arctangent model. Even when we add the unemployment rates and financial crises to the logit and the CLL models, we do not have enough evidence that one model is better than the other.

3.3 Step 3: Transformation of Period

The third step is to transform the policy age since issue. We transform the policy age because the surrender rates are increasing during early time periods and decreasing as time goes on. So there is a possibility that the fit may be decreased if we use the real policy age without transformation. Three formulas that are usually used for the transformation may be tried:[12]

$$\sqrt[n]{x},\ \log x,\ \text{and}\ -\frac{1}{x}. \tag{3.3}$$

The policy age may be transformed to

$$(policy\ age)^{1/n},\ \log(policy\ age),\ \text{or}\ -\frac{1}{policy\ age}. \tag{3.4}$$

We shall choose the best transformation formula for each insurance plan under the model fit statistics, $-2\log L$.

Let us compare the arctangent and logit models, and the arctangent and CLL models. We want to find which model is the best for each insurance plan. We shall begin with the protection plans. Table 8 and Figure 6 show the model fit statistics, $-2 \times \log L$, according to the policy age. Comparing the model fit statistics, we can conclude that the best transformation formula is

$$-\frac{1}{policy\ age}. \tag{3.5}$$

We see the analysis results in Tables 9 and 10. In these tables policy ages are transformed policy ages using the formula

$$-\frac{1}{policy\ age}.$$

[12] For more on transformation of variables, see Kutner, Nachtschiem, and Wasserman (1996).

Figure 6
Model Fit Statistics according to Transformed Policy Age (Protection Plan)

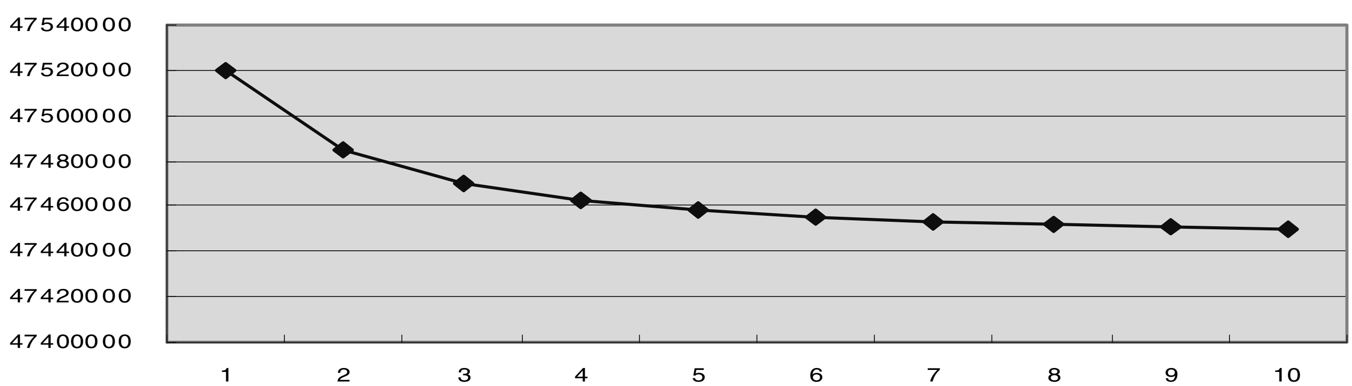


Table 9
Parameter Estimates with Logit Model under Transformation (Protection Plan)
Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	−5.5188	0.00330	2,799,348.11	<.0001
DIFFLAG0	1	4.2140	0.0222	35,894.5564	<.0001
DIFFLAG2	1	1.9623	0.0212	8,560.1585	<.0001
DIFFLAG4	1	0.0513	0.0183	7.8526	0.0051
DIFFLAG6	1	1.8964	0.0177	11,523.6609	<.0001
DIFFLAG8	1	2.3975	0.0168	20,448.0609	<.0001
DIFFLAG10	1	0.7305	0.0178	1,676.4216	<.0001
DIFFLAG12	1	1.0890	0.0231	2,228.1087	<.0001
POLICY AGE	1	−0.3251	0.000720	20,3728.130	<.0001
UNEMPLOY	1	7.5680	0.0782	9,375.3265	<.0001

Table 10
Parameter Estimates with CLL Model under Transformation (Protection Plan)
Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Error	Chi-Square	Pr > ChiSq
Intercept	1	−5.5179	0.00328	2,829,648.25	<.0001
DIFFLAG0	1	4.1885	0.0220	36,099.0236	<.0001
DIFFLAG2	1	1.9441	0.0210	8,570.8372	<.0001
DIFFLAG4	1	0.0531	0.0181	8.5646	0.0034
DIFFLAG6	1	1.8781	0.0175	11,545.8547	<.0001
DIFFLAG8	1	2.3789	0.0166	20,596.2987	<.0001
DIFFLAG10	1	0.7325	0.0177	1,719.3160	<.0001
DIFFLAG12	1	1.0820	0.0229	2,233.7300	<.0001
POLICY AGE	1	−0.3220	0.000712	204,566.026	<.0001
UNEMPLOY	1	7.5205	0.0777	9,370.5619	<.0001

Table 11
Errors and Comparison of Models under Transformation (Protection Plan)

Time	RMSE1	RMSE2	RMSE3	MAPE1	MAPE2	MAPE3	RMSEGAP1	RMSEGAP2	MAPEGAP1	MAPEGAP2
0.5	0.00835	0.00292	0.00293	0.26237	0.1128	0.11271	0.00543	0.00542	0.14957	0.14966
1.5	0.00434	0.00228	0.00228	0.20047	0.13214	0.13207	0.00206	0.00206	0.06832	0.06839
2.5	0.00467	0.00188	0.00187	0.25927	0.23508	0.23508	0.00279	0.00279	0.02419	0.02419
3.5	0.00389	0.00148	0.00148	0.22961	0.11367	0.11359	0.00241	0.00241	0.11594	0.11602
4.5	0.00353	0.00139	0.00139	0.19762	0.09254	0.09235	0.00214	0.00214	0.10508	0.10528
5.5	0.00369	0.00124	0.00124	0.18858	0.08779	0.08754	0.00245	0.00245	0.10079	0.10104
6.5	0.00375	0.00148	0.00148	0.21342	0.10069	0.10069	0.00227	0.00228	0.11273	0.11273
7.5	0.00432	0.00139	0.00139	0.22635	0.08808	0.08817	0.00293	0.00293	0.13827	0.13818
8.5	0.00649	0.002	0.002	0.2865	0.10751	0.10752	0.00449	0.00449	0.17899	0.17898
9.5	0.00882	0.00412	0.00412	0.33834	0.16507	0.16501	0.0047	0.0047	0.17327	0.17333

Figure 7
Surrender Rates with Policy Age Five Years (Protection Plan)

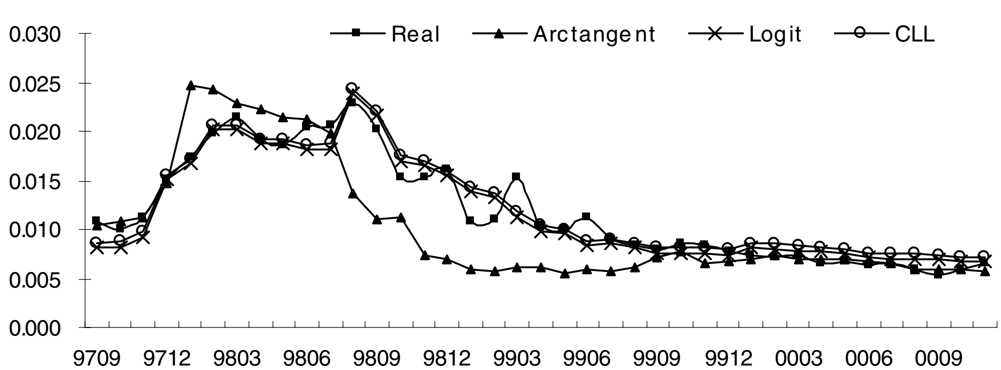


Table 12
RMSE and MAPE for Predicted Surrender Rates

	RMSE			MAPE		
	Arctangent	Logit	CLL	Arctangent	Logit	CLL
Protection	0.00091	0.00073	0.00073	0.09828	0.08935	0.08978
Endowment	0.00352	0.00327	0.00328	0.15634	0.19345	0.19431
Education	0.00106	0.00085	0.00085	0.08131	0.06046	0.06044
Annuity	0.00291	0.00206	0.00207	0.23046	0.19001	0.19164

When we do not transform the policy ages, the logit and CLL models are better than the arctangent model with the exceptions of policy ages one and three. But the logit and the CLL models are better than the arctangent model on the whole policy ages after the transformation. So we can say that the results are improved with the transformation of the policy ages.[13]

Figure 7 shows the real and estimated surrender rates for each model with policy age five years for the protection plan. We note that the logit and CLL models produce almost the same results, and the two graphs overlap.

3.4 Predicting Future Surrender Rates

In this section we shall check how well our surrender rate models predict future surrender rates. We have estimated the coefficient of each explanatory variable of our surrender rate models based on the data from January 1997 to November 2000. Using our surrender rate models, we shall predict the future surrender rates and compare them to the real surrender rate data for one year from December 2000 to November 2001.

Table 12 shows the root mean squared error (RMSE) and mean absolute percentage error (MAPE) for the predicted surrender rates under the arctangent, logit, and CLL models for the four insurance groups—protection plans, endowments, education plans, and annuities—with policy age five years.

Note that the logit and CLL models are better than the arctangent model for protection plans, education plans, and annuities. The three models have similar errors for endowment. Figure 8 shows that our models predict the surrender rates well.

4. Conclusion

We have investigated policyholder behaviors on surrendering their insurance policies. We have noted that the policyholder surrender behaviors are dependent not only on interest rates but also on other exogenous factors such as unemployment rates, economy growth rates, seasonal effects, and policy age since issue. All of the existing models on surrender rates are functions only of interest rates. So we need to consider new surrender rate models reflecting the complicated policyholder surrender behaviors with endogenous and exogenous multivariables. Using the logit and the CLL models as alternatives, we have improved the fit of the surrender rates under the logit and CLL models compared to that under the arctangent model, which is one of the existing surrender rate models with only an interest rate variable.

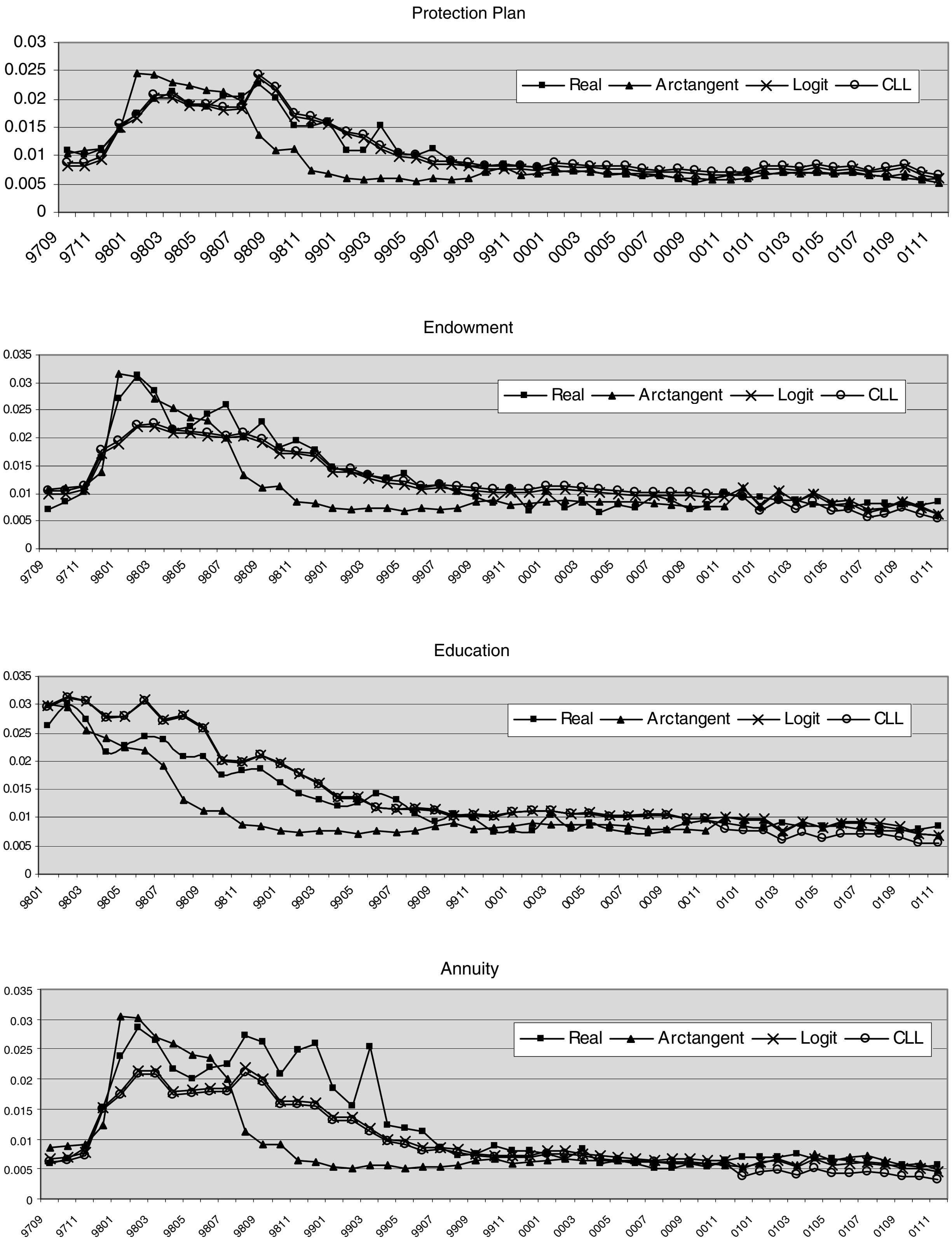
Surrender rate patterns are different according to insurance policies, so it is desirable that we

[13] The other types of plans—endowments, education plans, and annuities—are also analyzed. When we use the transformation formula, $(\textit{policy age})^{1/n}$, the fit becomes worse as n increases for endowment. Also the fit is not improved with the other transformation formulas. So we do not need to transform the policy ages for the endowment plan. Also we could not conclude that the logit or the CLL model is better than the arctangent model for endowment.

For education plans, we notice that the model fit statistics take the minimum when n is 5. The results from the other transformations are not better than that of the transformation, $\sqrt[5]{\textit{policy age}}$. So we use $\sqrt[5]{\textit{policy age}}$ for the transformation. When we do not transform the policy ages, the logit and the CLL models are better than the arctangent model with the exception of policy age three. But the logit and the CLL models are better than the arctangent model for the whole policy ages after the transformation.

Finally, we analyze annuities. Comparing the model fit statistics, we conclude that the best transformation formula is $-1/\textit{policy age}$. When we do not transform the policy ages, we do not notice that one model is better than the other ones. But the logit and the CLL models are better than the arctangent model for many policy ages after we transform the policy ages.

Figure 8
Predicted Surrender Rates



should select a proper surrender rate model for each policy. The analysis results on the proper surrender rate models for the four insurance policies are summarized as follows:

1. For protection plans, the logit and CLL models are better than the arctangent model.
2. For education plans, the logit and CLL models are better than the arctangent model.
3. For endowment plans, we do not have a superior model among the logit, CLL, and arctangent models.
4. For annuities, the logit and CLL models are better than the arctangent model.

Generally, the logit and CLL models are better than the arctangent model in modeling surrender rates. But the model selection depends on the insurance policy and the explanatory variables.

REFERENCES

ABRAHAM, J. M., AND H. S. THEOBALD. 1997. A Simple Prepayment Model of Commercial Mortgages. *Journal of Housing Economics* 6(1): 31–59.

AGRESTI, ALAN. 1996. *An Introduction to Categorical Data Analysis*. New York: John Wiley.

———. 2002. *Categorical Data Analysis*. New York: John Wiley.

ALLISON, PAUL D. 1999. *Logistic Regression Using the SAS System: Theory and Application*. Cary, N.C.: SAS Publishing.

ARCHER, WAYNE R., DAVID C. LING, AND GARY A. MCGILL. 1997. Demographic versus Option-Driven Mortgage Terminations. *Journal of Housing Economics* 6(2): 137–63.

ASAY, M. R., P. J. BOUYOUCOS, AND A. M. MARCIANO. 1993. An Economic Approach to Valuation of Single Premium Deferred Annuities. In *Financial Optimization,* ed. Stavros A. Zenios, pp. 101–35. Cambridge: Cambridge University Press.

AUSTIN, KELLY 1995. Racial and Ethnic Disparities in Mortgage Prepayment. *Journal of Housing Economics* 4(4): 350–72.

BIERWAG, GERALD O. 1987. *Duration Analysis: Managing Interest Rate Risk*. Cambridge, Mass.: Ballinger.

BIERWAG, G. O., C. J. CORRADO, AND G. G. KAUFMAN. 1990. Computing Durations for Bond Portfolios. *Journal of Portfolio Management* (fall): 51–55.

BLUHM, WILLIAM F. 1982. Cumulative Antiselection Theory. *Transactions of the Society of Actuaries* 34: 215–32; Discussions: 233–46.

———. 1993. Duration-Based Policy Reserves. *Transactions of the Society of Actuaries* 45: 11–32; Discussions: 33–53.

BÖHEIM, RENÉ, AND MARK PHILIP TAYLOR. 2000. My Home Was My Castle: Evictions and Repossessions in Britain. *Journal of Housing Economics* 9(4): 287–319.

BROWNE, MARK J., JAMES M. CARSON, AND ROBERT E. HOYT. 1999. Economic and Market Predictors of Insolvencies in the Life-Health Insurance Industry. *Journal of Risk and Insurance* 66(4): 643–59.

BRZEZINSKI, JOSEPH R. 1975. LIMRA 1971–72 Expected Lapse Tables. *Transactions of the Society of Actuaries* 27: 267–312.

BUCK, N. F. 1960. First Year Lapse and Default Rates. *Transactions of the Society of Actuaries* 12(33): 258–93; Discussions: 294–314.

CARLING, KENNETH, TOR JACOBSON, AND KASPER ROSZBACH. 2001. Dormancy Risk and Expected Profits of Consumer Loans. *Journal of Banking & Finance* 25(4): 717–39.

DE TOLDI, M., C. GOURIEROUX, AND A. MONFORT. 1995. Prepayment Analysis for Securitization. *Journal of Empirical Finance* 2(1): 45–70.

DUKES, J., AND A. M. MACDONALD. 1980. Pricing a Selected and Ultimate Annual Renewable Term Product. *Transactions of the Society of Actuaries* 32: 547–66; Discussions: 567–84.

DUNN, K. B., AND J. J. MCCONNELL. 1981. A Comparison of Alternative Models for Pricing GNMA Mortgage-Backed Securities. *Journal of Finance* 36(2): 471–84.

FABOZZI, FRANK J., ED. 2001. *Handbook of Fixed Income Securities*. 6th ed. Homewood, Ill.: Richard D. Irwin.

FIRTH, D. 1991. Generalized Linear Models. In *Statistical Theory and Modeling,* edited by D. V. Hinkley, N. Reid, and E. J. Snell. London: Chapman and Hill.

FOLLOWILL, L. 1998. An Analytical Comparison of the Durations and Price Sensitivities of Fixed-Rate, Constant Payment and Constant Amortization Mortgages. *International Review of Financial Analysis* 7(1): 51–64.

GREEN, J., AND J. B. SHOVEN. 1986. The Effect of Interest Rates on Mortgage Prepayments. *Journal of Money, Credit, and Banking* 18(1): 41–59.

GREEN, RICHARD K., AND MICHAEL LACOUR-LITTLE. 1999. Some Truths about Ostriches: Who Doesn't Prepay Their Mortgages and Why They Don't. *Journal of Housing Economics* 8(3): 233–48.

HALL, A. 2000. Controlling for Burnout in Estimating Mortgage Prepayment Models. *Journal of Housing Economics* 9(4): 215–32.

HAMILTON, JAMES D. 1994. *Time Series Analysis*. Princeton: Princeton University Press.

HARRELL, FRANK E., JR. 2001. *Regression Modeling Strategies: With Applications to Linear Models, Logistic Regression, and Survival Analysis*. New York: Springer.

JOHNSEN, THOMAJEAN, AND RONALD W. MELICHER. 1994. Predicting Corporate Bankruptcy and Financial Distress: Information Value Added by Multinomial Logit Model. *Journal of Economics and Business* 46: 269–86.

KANG, PAN, AND STAVORS A. ZENIOS. 1992. Complete Prepayment Models for Mortgage-Backed Securities. *Management Science* 38(11): 1665–85.

KIM, C. 2005. Surrender Rate Impacts on Asset/Liability Management. *Asia-Pacific Journal of Risk and Insurance*. Forthcoming.

KOLARI, JAMES, DENNIS GLENNON, HWAN SHIN, AND MICHELE CAPUTO. 2002. Predicting Large US Commercial Bank Failures. *Journal of Economics and Business* 54: 361–87.

Kutner, Michale H., Christopher J. Nachtschiem, and W. Wasserman. 1996. *Applied Linear Statistical Models*. New York: Irwin/McGraw-Hill.

Lo, Andrew W. 1986. Logit versus Discriminant Analysis: A Specification Test and Application to Corporate Bankruptcies. *Journal of Econometrics* 31: 151–78.

McCullagh, Peter, and J. A. Nelder. 1989. *Generalized Linear Models*. 2nd ed. Boca Raton, Fla.: CRC Press.

McCulloch, Charles E., and Shayle R. Searle. 2000. *Generalized, Linear, and Mixed Models*. New York: Wiley-Interscience.

Miyazaki, Kenji, and Makoto Saito. 1999. On the Market Risk Involved in the Public Financial System in Japan: A Theoretical and Numerical Investigation. *Journal of Banking & Finance* 23(8): 1243–59.

Outreville, J. Francois. 1990. Whole-Life Insurance Lapse Rates and the Emergency Fund Hypothesis. *Insurance: Mathematics and Economics* 9(4): 237–306.

Panjer, Harry H., et al. 1998. *Financial Economics: With Applications to Investments, Insurance and Pensions*. Schaumburg, Ill.: Actuarial Foundation.

Pesando, J. E. 1974. The Interest Sensitivity of the Flow of Funds through Life Insurance Companies: An Econometric Analysis. *Journal of Finance* 29(4): 1105–21.

Pinder, Jonathan P. 1996. Decision Analysis Using Multinomial Logit Model: Mortgage Portfolio Valuation. *Journal of Economics and Business* 48(1): 67–77.

Richard, S. 1991. Valuation Challenges: Mortgage-Backed Securities and Collateralized Mortgage Obligations. In *Understanding Securitized Investments and Their Use in Portfolio Management,* edited by K. M. Eades, D. R. Harrington, and R. S. Harris, pp. 20–29. Charlottesville, Va.: Association for Investment Management and Research.

Richard, S., and R. Roll. 1989. Prepayments on Fixed-Rate Mortgage-Backed Securities. *Journal of Portfolio Management* 15: 73–82.

Richardson, C. F. B., and J. M. Hartwell. 1951. Lapse Rates. *Transactions of the Society of Actuaries* 3(7): 338–74; Discussions: 375–96.

SAS Institute. 1999. *SAS/STAT User's Guide Version 8,* Cary, N.C.: SAS Institute.

Sharp, K. P. 1996. Lapses and Termination Assumptions in Reserve Calculations. *Journal of Insurance Regulation* 15(2): 194–211.

Sherris, Michael. 1994. A One-Factor Interest Rate Model and the Valuation of Loans with Prepayment Provisions. *Transactions of the Society of Actuaries* 46: 251–301; Discussions: 303–20.

Stanton, R. 1995. Rational Prepayment and the Valuation of Mortgage-Backed Securities. *Review of Financial Studies* 8(3): 677–708.

Weisberg, Sanford. 1985. *Applied Linear Regression.* 2nd ed. New York: John Wiley.

Zenios, Stavros A., ed. 1993. *Financial Optimization*. Cambridge: Cambridge University Press.

Discussions on this paper can be submitted until April 1, 2006. The author reserves the right to reply to any discussion. Please see the Submission Guidelines for Authors on the inside back cover for instructions on the submission of discussions.