2024 Annual BAS Case Competition

Team 1: Nathan Wei Chan, Derek Younger, Dalton Bassler-Haynes, Beloved Maina

Agenda

01 Introduction

02 Evaluating Four Key Predictors

03 Lapse Rate Formula Proposal and Projection

04 Additional Considerations in regards to GLB

05 Results and Recommendations

Introduction

Goals

01 Proposing new lapse rate formula for Sgt. Pepper Financial Group

02 Narrow down 4 key factors to include in lapse rate equation

03 Predict future lapse rates given sample product data

04 Recommend additional factors for the addition of a GLB to the product

Evaluating Factors for Lapse Formula

Methodology

01 **Linear Modeling in RStudio**
Linearly modeled all predictors given

02 **Leap Analysis in RStudio**
Obtained 4 most contributing factors

03 **Match Analyzed Equation**
Compare equation with historical data

04 **Evaluation of Comparison**
Examined R^2, MAPE, BIC of data

Impactful Factors

Mortality Rate

High mortality rate→high lapse rate

Crediting Rate

High crediting rate→retention of policyholders

10 Yr. Treasury Rate

Long term indicator of future interest rate

Surrender Charge

Penalty discouraging early lapse/withdrawal

Proposed Lapse Formula

Lapse Rate as a function of the 4 proposed variables.

$$f(a, b, c, d) = 0.1135 + 0.516a - 0.876b + 1.08c - 1.06d$$

a = Mortality Rate

b = Crediting Rate

c = 10-Year Treasury Rate

d = Surrender Charge

$\square_0$ = 0.1135

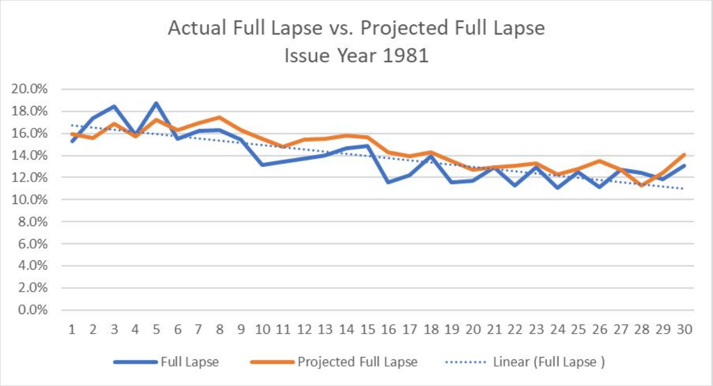
$\square_1$ = 0.516

$\square_2$ = -0.876

$\square_3$ = 1.08

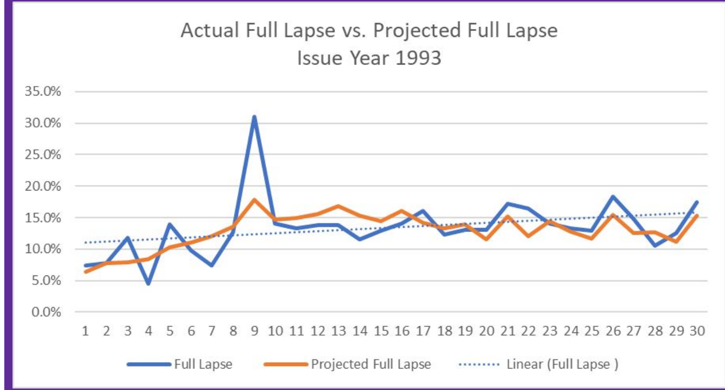
$\square_4$ = -1.06

Matching Equation with Historical Data

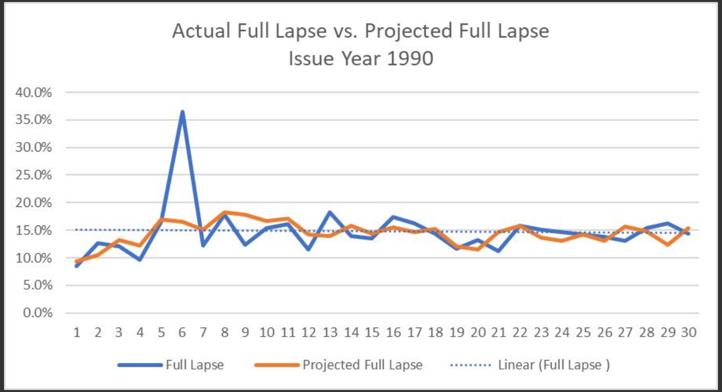


MAPE (including outlier): 8.614%
(no outlier)

Cumulated R^2: 0.3154
BIC (8 Variables): -572.27
BIC (4 Variables): -592.13



MAPE (including outlier): 18.594%
MAPE (excluding outlier): 17.772%



MAPE (including outlier): 14.163%
MAPE (excluding outlier): 12.765%

Cumulated R^2 and BIC solved by using R Studio

MAPE: $M = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{A_t - F_t}{A_t}\right|$

Lapse Rate Predictions with Projected Data

Projected Lapse Using Equation

a = Mortality Rate
b = Crediting Rate
c = 10-Year Treasury Rate
d = Surrender Charge

$$f(a, b, c, d) = 0.1135 + 0.516a - 0.876b + 1.08c - 1.06d$$

Year	Mortality Rate	Crediting Rate	10 Yr. Treasury Rate	Surrender Charge	Projected Lapse
2023	0.06%	4.50%	3.96%	12%	3.00%
2024	0.46%	4.50%	3.95%	10%	3.00%
2025	1.79%	4.50%	4.00%	10%	3.00%
2026	0.30%	4.70%	4.19%	7%	4.49%
2027	1.05%	4.80%	4.28%	5%	7.02%
2028	1.64%	4.80%	4.93%	4%	9.08%
2029	1.81%	5.00%	5.07%	3%	10.21%
2030	0.34%	5.00%	5.64%	2%	11.13%
2031	0.32%	5.00%	5.80%	0%	13.40%
2032	0.18%	5.20%	5.50%	0%	12.83%
2033	1.77%	5.20%	6.16%	0%	14.37%
2034	0.26%	5.50%	5.90%	0%	13.04%
2035	0.37%	5.50%	7.42%	0%	14.74%
2036	0.78%	5.50%	6.60%	0%	14.07%
2037	0.14%	5.50%	7.00%	0%	14.17%

Account for 17.09%
increase outlier

Additional proposed factors

Additional Proposed Factor, Monthly Unemployment Rates

01 Assume policyholders schedule GLB to correlate with retiring.

02 Unemployment Rate doesn’t account for retirees

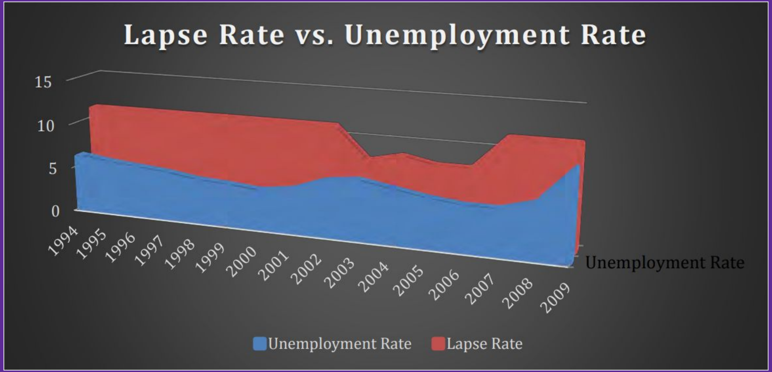
03 Focused on policyholders before GLB occurs

Unemployment Example

- Study by SOA, showing that Lapse Rate vs. Unemployment Rate are weakly correlated. (Modeling the Unemployment Risk in Insurance Products, SOA)

Table 23

Statistic	Result (%)
Correlation	4.4
Covariance	8.0
R-Squared	0.2
Slope	3.8



Results and Recommendations

MAPE Interpretation

We solved for the Mean Absolute Percentage Error, a good estimate of the data is generally accepted to have an MAPE of ≤ 20%.

R^2 Interpretation

Obtained an R^2 value of 0.3154 using RStudio. Relatively accurate to the cumulative data given a large set of variables to consider.

BIC Interpretation

Bayesian Information Criterion is a measure of the simplicity and relative accuracy of the linear representation. Measures absolute difference between 2 BIC; ≥ 6 is significant.

Overall Recommendation

Our equation builds a relatively good prediction using Mortality Rate, Crediting Rate, 10 Yr. Treasury Rate, and Surrender Charge. These 4 variables and our equation should be of good use to Sgt. Pepper Financial Group.

Why does Unemployment Matter?

01

Previous unemployment rates predict future months lapse rates

02

Unemployment targets policyholders who haven't received GLB.

03

If unemployment occurs, long term priorities change.

RStudio Appendix

(A) Summary of Linear Regression

```
#Using all the data to make the main model with all variables
model_all_pred = lm(Full.Lapse ~ MVA + Mortality.Rate + Crediting.Rate +
                    X5yr.Treasury.Rate + General.Account.Portfolio.Yield +
                    Statutory.Reserves....billions. + X10.Yr.Treasury.rate +
                    Surrender.charge, data = master_data)
summary(model_all_pred)
# now finding best subset of 4
# Using leaps data package to systematically check every combination for the best adjusted R^2 value
best_subset = regsubsets(Full.Lapse~.,
                         data = master_data, nbest = 1, nvmax = NULL, force.in = NULL, force.out = NULL,
                         method = "exhaustive")
best_subset

summary_best_subset <- summary(best_subset)
as.data.frame(summary_best_subset$outmat)
which.max(summary_best_subset$adjr2)
# Which 4 are the best variables that contribute the most to R^2
summary_best_subset$which[4,]
# Final formula with the 4 best predictors
model_all_pred_2 = lm(Full.Lapse ~ Mortality.Rate + Crediting.Rate +
                      X10.Yr.Treasury.rate +
                      Surrender.charge, data = master_data)
# Final model with all the intercepts
summary(model_all_pred_2)
```

(B) Summary of 8 Variable Model

```
> summary(model_all_pred)

Call:
lm(formula = Full.Lapse ~ MVA + Mortality.Rate + Crediting.Rate +
    X5yr.Treasury.Rate + General.Account.Portfolio.Yield + Statutory.Reserves....billions. +
    X10.Yr.Treasury.rate + Surrender.charge, data = master_data)

Residuals:
      Min        1Q    Median        3Q       Max
-0.053673 -0.012104 -0.004833  0.008350  0.201799

Coefficients:
                                 Estimate Std. Error t value Pr(>|t|)
(Intercept)                      0.083132   0.071554   1.162   0.2487
MVA                             -0.074025   0.163031  -0.454   0.6510
Mortality.Rate                   0.470492   0.218905   2.149   0.0346 *
Crediting.Rate                  -0.962730   0.528502  -1.822   0.0722 .
X5yr.Treasury.Rate               0.227750   0.780656   0.292   0.7712
General.Account.Portfolio.Yield  0.526794   1.832566   0.287   0.7745
Statutory.Reserves....billions.  0.001037   0.002402   0.432   0.6671
X10.Yr.Treasury.rate             0.776866   1.235763   0.629   0.5313
Surrender.charge                -1.151413   0.236218  -4.874 5.35e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03323 on 81 degrees of freedom
Multiple R-squared:  0.3639,    Adjusted R-squared:  0.301
F-statistic: 5.792 on 8 and 81 DF,  p-value: 7.743e-06

> # Which 4 are the best variables that contribute the most to R^2
> summary_best_subset$which[4,]
             (Intercept)              ï..Policy.Year                             MVA                  Mortality.Rate
                    TRUE                       FALSE                           FALSE                            TRUE
          Crediting.Rate          X5yr.Treasury.Rate General.Account.Portfolio.Yield Statutory.Reserves....billions.
                    TRUE                       FALSE                           FALSE                           FALSE
    X10.Yr.Treasury.rate            Surrender.charge
                    TRUE                        TRUE
```

(C) Summary of Reduced Linear Regression

```
> # Final model with all the intercepts
> summary(model_all_pred_2)

Call:
lm(formula = Full.Lapse ~ Mortality.Rate + Crediting.Rate + X10.Yr.Treasury.rate +
    Surrender.charge, data = master_data)

Residuals:
     Min       1Q   Median       3Q      Max
-0.05497 -0.01493 -0.00713  0.01288  0.19969

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)           0.11355    0.01646   6.897 8.83e-10 ***
Mortality.Rate        0.51615    0.21186   2.436   0.0169 *
Crediting.Rate       -0.87614    0.51976  -1.686   0.0955 .
X10.Yr.Treasury.rate  1.08137    0.25696   4.208 6.35e-05 ***
Surrender.charge     -1.06032    0.22319  -4.751 8.16e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.03289 on 85 degrees of freedom
Multiple R-squared:  0.3462,	Adjusted R-squared:  0.3154
F-statistic: 11.25 on 4 and 85 DF,  p-value: 2.259e-07
```

(D) BIC Calculations

```
# Finding BIC, grading on how well the model models data and simplicity
residuals = residuals(model_all_pred_2)
residuals
# RSS- residual sum of squares
RSS = sum(residuals^2)
RSS
# Get sample size (n) and number of parameters (k)

n <- length(residuals)
k <- length(coefficients(model_all_pred_2))
# Estimate variance of residuals
sigma_hat_squared <- RSS / (n - k)
sigma_hat_squared
# Calculate BIC
BIC_4_var <- n * log(sigma_hat_squared) + k * log(n)
BIC_4_var
# BIC for the 4 variable model
# Find BIC for model with all variables
residuals_8 = residuals(model_all_pred)
RSS_8 = sum(residuals_8^2)
n_8 = length(residuals_8)
k_8 = length(coefficients(model_all_pred))
sigma_8 = RSS_8 / (n_8 - k_8)
BIC_8_var = n_8 * log(sigma_8) + k_8 * log(n_8)
BIC_8_var
# Since the BIC difference is greater than 10, therefore there is a significant
# Preference over the 4 variable model vs the 8 variable model.
```

(E) BIC Implementation

```
> RSS
[1] 0.09194491
> n
[1] 90
> k
[1] 5
> sigma_hat_squared
[1] 0.001081705
> BIC_4_var
[1] -592.1305
> RSS_8
[1] 0.08945228
> n_8
[1] 90
> k_8
[1] 9
> sigma_8
[1] 0.001104349
> BIC_8_var
[1] -572.2666
```

Thank You! Any questions?