Regression Case Study

Jeff Fuerte,
Virginia Commonwealth University, Richmond, Virginia, USA

1 Executive Summary

At the request of Progressive Insurance, a study was conducted by Qual-Pro consultants to see which combination of factors would increase renewal through its customers the highest. This report provides a description of the consultant team’s approach, the results of the testing of the factors, conclusions, and recommendations. The report is arranged to guide the reader from the general to the specific, with full detail provided in the technical appendices.

The consultant’s job was to find the optimal ”recipe” from a combination of twenty-six factors into fifty-six recipes. Testing occurred over a period of three months, utilizing almost 60,000 policyholders in Florida, Pennsylvania, and Colorado. The consultants split these policyholders into groups based on their riskiness (nonstandard and middle market against standard, preferred and ultra-preferred) and their price (increase against decrease). The best recipe was then found for the nonstandard or middle market customer whose price decreased, the nonstandard or middle market customer whose price decreased, the standard or preferred or ultra-preferred customer whose price decreased, the customer whose price increased, the customer whose priced increased, and for the overall customer. In each case, the important variables that influence a positive renewal were found, and thus determined to be our recommendations.

For customers that are either nonstandard or middle market who have a price decrease, no optimal solution was found. Therefore, it is best to include none of the factors, and treat these customers the same. The company should not perform any of the recipes, as we do not know how this might influence the renewal rate, either positively or negatively.

For customers that are either standard, preferred, or ultra-preferred who have a price decrease, only one factor positively influenced the renewal rate: mailing a claims comment card. Therefore, for these customers, to achieve the highest renewal rate, the company should mail a claims comment card to them.
For customers that are nonstandard or middle market and have a price increase, the best recommendation is to send a letter after the actual lapse. This factor influenced the renewal rate in a positive way, and therefore is our recommendation.

For customers that are standard, preferred, or ultra-preferred and have a price increase, no optimal solution was found. Therefore, doing nothing would be better than doing something, as we do not know how the factors might increase or decrease the renewal rate. It is best to not implement a "recipe" in this case.

For all customers that have a price decrease, no optimal solution was found. Therefore, doing nothing would be better than doing something, as we do not know how the factors might increase or decrease the renewal rate.

For all customers that have a price increase, only sending a letter after the actual lapse will influence a positive renewals. Therefore, this is our recommendation to the company: send a letter after the actual lapse to all customers that have a price increase.

For all customers, two factors were found to be important. The company should provide a postage-paid envelope with the renewal quote and send a letter after the actual lapse in order to achieve the highest renewal rate. For all customers, our recommendation is to do both of these things.

2 Main Report

2.1 Introduction

At the request of Progressive Insurance, a study was conducted by Qual-Pro consultants to see which combination of factors would increase renewal through its customers the highest. This report provides a description of the consultant team’s approach, the results of the testing of the factors, conclusions, and recommendations. The report is arranged to guide the reader from the general to the specific, with full detail provided in the technical appendices.

The consultant’s job was to find the optimal "recipe" from a combination of twenty-six factors into fifty-six recipes. Testing occurred over a period of three months, utilizing almost 60,000 policyholders in Florida, Pennsylvania, and Colorado. The consultants split these policyholders into groups based on their riskiness (nonstandard and middle market against standard, preferred and ultra-preferred) and their price (increase against decrease). The best recipe
was then found for the nonstandard or middle market customer whose price decreased, the nonstandard or middle market customer whose price decreased, the standard or preferred or ultra-preferred customer whose price decreased, the standard or preferred or ultra-preferred customer whose price increased, the customer whose price decreased, the customer whose price increased, and for the overall customer.

Using the statistical package SAS, a logistic regression was performed using a logit model to account for the binary response: yes or no to renewal. Using all twenty-six factors in the analysis, the best combination was obtained for each of the seven different customers. In this case, the response variable was if the customer renewed the policy (coded as a 1 for yes or 0 for no), and the explanatory variables were each of the twenty-six factors (coded as a 1 for if the factor was implemented and a 0 for otherwise).

### 2.2 Findings

The conclusions and recommendations proposed in this section are supported by the analysis done in the technical appendices at the end of the report. For each of the seven cases brought up (splitting the customers into riskiness and price), a logistic regression was performed in SAS. For each of the regressions, all twenty-six factors were ran against the number of renewed divided by the number of eligible renewals. The results were found, and significance was determined using an alpha level of .01.

1. What is the best "recipe" for nonstandard or middle market customers who have a price decrease?

In this case, the only significant variable was factor L. However, the estimate for this factor was a negative value. Therefore, this would impact the regression in a negative way, meaning that the renewals would actually go down if this variable was included. In essence, this variable should not be included, so no variables were positively significant. In this case, implementing none of the factors will increase the renewal rate the highest.

2. What is the best "recipe" for standard, preferred, or ultra-preferred customers who have a price decrease?

In this case, the only significant variable was factor Q. The estimate for this factor was a positive value, and the fact that it is positive means we should include this in our model. Renewals will increase if factor Q is "turned on" or effectively performed. Factor Q corresponds to mailing a claims comment card to all, so for customers that are standard, preferred, or ultra-preferred who have a price decrease a claims comment card should be mailed.
3. What is the best "recipe" for nonstandard or middle market customers who have a price increase?

In this case, the significant variables were factor D and factor U. Factor D has a negative estimate, and because of this it will be thrown out as it will decrease renewals. Factor U has a positive estimate, so by performing factor U, the renewal rate will increase. This factor corresponds to sending a letter after the actual lapse. Therefore, for customers that are nonstandard or middle market who have a price increase, the company should send a letter after the actual lapse.

4. What is the best "recipe" for standard, preferred, or ultra-preferred customers who have a price increase?

In this case, the significant variables were factor L and factor Y. The estimates for these factors were both negative, so they will influence the renewal rate in a negative way. Therefore, it is best to leave them out, and conclude that no optimal solution was found. For customers that are standard, preferred, or ultra-preferred who have a price increase, the best "recipe" would be to do nothing at all, as we don’t know how it might influence the renewal rate, positively or negatively.

5. What is the best "recipe" for customers that have a price decrease?

In this case, no significant variables are present in this case. Therefore, the best thing for the company to do with customers that have a price decrease is to do nothing at all, as one does not know how this might affect renewal rate.

6. What is the best "recipe" for customers that have a price increase?

In this case, the significant variables are factor D, factor L, and factor U. The estimates for factor D and factor L are both negative, so these factors should be omitted from the model. The estimate for factor U is positive and should be included in the model. This factor corresponds to sending a letter after the actual lapse. Therefore, for customers that have a price decrease, the company should send a letter after the actual lapse.

7. What is the best "recipe" for all customers?

In this case, the significant variables are factor D, factor L, factor N, and factor U. The estimates for factor D and factor L are both negative, so we should leave these out as they will influence the renewal rate in a negative way. Factor N and factor U have positive estimates, so these two should be performed. Because two factors were positively significant, the interaction of factor N and factor U was looked at. This regression was found to be insignificant, and it should not be included in the model. Factor N corresponds to providing
a postage-paid envelope with the renewal quote, and factor U corresponds to sending a letter after the actual lapse. Therefore, for all customers, the company should provide a postage-paid envelope with the renewal quote and send a letter after the actual lapse.

2.3 Conclusions and Recommendations

Based on these conclusions, certain recommendations should be considered when deciding what to do with certain customers. For customers that are nonstandard or middle market that have a price decrease, or for customers that are standard, preferred, or ultra-preferred who have a price increase, or for all customers who have a price decrease, no factors should be implemented. For customers that are nonstandard or middle market and have a price increase, we would recommend sending a letter after the actual lapse. For customers that have a price increase, we would recommend the same thing. Mailing a claims comment card would be the recommendation for customers that are standard, preferred, or ultra-preferred and have a price decrease. Overall to all customers in general, we would recommend providing a postage-paid envelope with the renewal quote and sending a letter after the actual lapse to all customers in order to achieve the highest possible renewal rate.

3 Technical Appendix

Using all twenty-six factors (A-D, F-Z, AA) that were given, coding them a 1 (if used) or a 0 (if not used) and the response variable of renewal and eligible renewals, a logistic regression was used with the model y/n, with y being the renewals and n being the eligible renewals. This was run for all seven of the combinations of response variables. The model was a binary logit, with the optimization technique of Fisher’s scoring. Fifty-six observations were read in, with these being the "recipes" that were used in the experiment. This procedure was the same for all seven questions.

1. What is the best "recipe" for nonstandard or middle market customers who have a price decrease?

Using the logistic regression operator in SAS, with an alpha level of .01, the significant variable was factor L (p-value = .0769). However, the estimate for this factor was -.0747. Therefore, this would impact the regression in a negative way, meaning that the renewals would actually go down if this variable was included. In essence, this variable should not be included, so no variables were positively significant. The best recipe for customers that are nonstandard or
middle market who have a price decrease would be to do nothing at all because one would not know how the variable would impact the regression because none are positively significant. In this case, no optimal solution was found.

2. What is the best "recipe" for standard, preferred, or ultra-preferred customers who have a price decrease?

Using the logistic regression operator in SAS, with an alpha level of .01, the significant variable was factor Q (p-value = .0677). The estimate for this factor was .0805, and the fact that it is positive means we should include this in our model. Renewals will increase if factor Q is "turned on" or effectively performed. Factor Q corresponds to mailing a claims comment card to all, so for customers that are standard, preferred, or ultra-preferred who have a price decrease a claims comment card should be mailed.

3. What is the best "recipe" for nonstandard or middle market customers who have a price increase?

Using the logistic regression operator in SAS, with an alpha level of .01, the significant variables were factor D (p-value = .0534) and factor U (p-value = .0006). Factor D has an estimate of -.0625, and because this is negative it will be thrown out because it will decrease renewals. Factor U has an estimate of .1108, and this is positive so by performing factor U, the renewal rate will increase. This factor corresponds to sending a letter after the actual lapse. Therefore, for customers that are nonstandard or middle market who have a price increase.

4. What is the best "recipe" for standard, preferred, or ultra-preferred customers who have a price increase?

Using the logistic regression operator in SAS, with an alpha level of .01, the significant variables were factor L (p-value = .0471) and factor Y (p-value = .0860). The estimates for these factors were -.0713 and -.0617, respectively. Because these are both negative, they will influence the renewal rate in a negative way. Therefore, it is best to leave them out, and conclude that no optimal solution was found. For customers that are standard, preferred, or ultra-preferred who have a price increase, the best "recipe" would be to do nothing at all, as we don’t know how it might influence the renewal rate, positively or negatively.

5. What is the best "recipe" for customers that have a price decrease?

Using the logistic regression operator in SAS, with an alpha level of .01, no significant variables are present in this case. Therefore, the best thing for the company to do with customers that have a price decrease is to do nothing at all, as one does not know how this might affect renewal rate.
6. What is the best “recipe” for customers that have a price increase?

Using the logistic regression operator in SAS, with an alpha level of .01, the significant variables are factor D (p-value = .0568), factor L (p-value = .0344), and factor U (p-value = .0547). The estimates for factor D and factor L are -.0557 and -.0461, respectively. Because these estimates are negative, these factors should be omitted from the model. The estimate for factor U is .0728. This factor is positive and should be included in the model. This factor corresponds to sending a letter after the actual lapse. Therefore, for all customers that have a price increase, the best way to up the renewal rate would be to send a letter after the actual lapse.

7. What is the best “recipe” for all customers?

Using the logistic regression operator in SAS, with an alpha level of .01, the significant variables are factor D (p-value = .0570), factor L (p-value = .0343), factor N (p-value = .0952), and factor U (p-value = .0019). The estimates for factor D and factor L are -.0355 and -.0395, respectively. Because both of these are negative, we should leave these out as they will influence the renewal rate in a negative way. Factor N and factor U have estimates of .0311 and .0579, respectively, so these two should be performed. Because two factors were positively significant, the interaction of factor N and factor U was looked at. This estimate was found to be insignificant (p-value = .6893), and it should not be included in the model. Factor N corresponds to providing a postage-paid envelope with the renewal quote, and factor U corresponds to sending a letter after the actual lapse. Therefore, for all customers, the company should provide a postage-paid envelope with the renewal quote and send a letter after the actual lapse.