Texas Department of Insurance
Commissioner of Insurance, *Mail Code 113-1C* ***Jose Montemayor***
333 Guadalupe • P. O. Box 149104, Austin, Texas 78714-9104
512-463-6464 telephone • 512-475-2005 fax • www.tdi.state.tx.us

January 31, 2005

The Honorable Rick Perry
Governor of Texas
P.O. Box 12428
Austin, Texas 78711

The Honorable David Dewhurst
Lieutenant Governor of Texas
The Capitol
Austin, Texas 78711

The Honorable Tom Craddick
Speaker, Texas House of Representatives
P.O. Box 2910
Austin, Texas 78768

Dear Governors and Mr. Speaker:

It is my honor to present the remainder of the credit scoring study mandated by Senate Bill 14, 78th Regular Session.

The first phase of the analysis, published December 31, 2004, indicated that credit scoring is correlated to risk. The first phase also indicated that certain age, income and race groups tended to have worse credit scores, though not all minorities have bad credit scores. Because the Texas Insurance Code does not prescribe a precise threshold or legal definition for determining disproportionate impact, I expressed disproportionate impact in terms of a relationship, i.e., over or under-representation in the various credit score categories.

The second phase of the study evaluates if, and to what extent, credit scoring enables an insurer to more accurately predict losses when used in conjunction with other variables. Overall, the second phase indicates that credit scoring significantly improves pricing accuracy when combined with other rating variables in predicting risk.

To approach the development of a credit scoring policy, it is important to understand the distinction between unfair discrimination, intentional discrimination and disproportionate impact. The Texas Insurance Code defines unfair discrimination to be the unequal treatment of individuals in the same class or hazard. Underwriting or rating classifications are not unfair, though, if they are actuarially supported. On the other hand, overt classifications or discriminations based on race, color, religion or national origin are intentional discrimination and are prohibited by law, regardless of actuarial support.

Disproportionate impact is a lack of symmetry, or unequal percentages. In other words, disproportionate impact is an uneven distribution of each racial group within a given risk factor, although the uneven distribution is not caused by one's race. Also, disproportionate impact changes over time. For credit scoring, the disproportionate impact changes as economic conditions and population distribution change. By the nature of risk-based pricing and underwriting, all factors used in insurance have a disproportionate impact to some extent. One could make a convincing argument to ban the use of all risk-related factors based solely on disproportionate impact. Effectively, we would ban risk-based pricing and underwriting and revert to a pricing system where we homogenize the risk and essentially charge everyone the same price--regardless of risk. That would be a set-back to all Texans, of all races, especially those of moderate to lower income whose risk remains low.

As Commissioner, I have the authority to end a practice that is either unfairly or intentionally discriminatory. However, I do not have a legal basis to ban a practice that has a disproportionate impact if it produces an actuarially supported result and is not unfairly or intentionally discriminatory. Prior to the study, my initial suspicions were that while there may be a correlation to risk, credit scoring's value in pricing and underwriting risk was superficial, supported by the strength of other risk variables. Hence, there would be evidence that credit scoring was a coincidental variable that served as a surrogate for an unlawful factor in rating and underwriting. If this were proven to have been the case, I would have had a legal basis to make the connection between disproportionate impact and intentional discrimination, and either ban credit scoring outright or adopt an allowable rate difference of zero, meaning no rate differences due to credit scoring.

The study, however, did not support those initial suspicions. Credit scoring, if continued, is not unfairly discriminatory as defined in current law because credit scoring is not based on race, nor is it a precise indicator of one's race. Recall that not all minorities are in the worst credit score categories. Further, its use is justified actuarially and it adds value to the insurance transaction. Without a change in statute that disallows credit scoring as a matter of public policy, any action to ban may be tied up in court for several years, further frustrating public expectation.

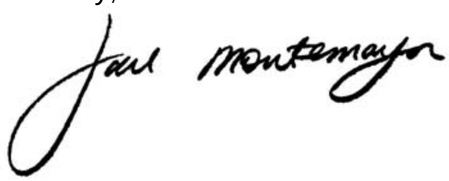
Be advised, however, that banning credit scoring overnight, by rule or law, creates pricing and availability disruptions in a market that has just stabilized and begun a rebound. The same effect would occur if a narrow rate limit, or collar, due to credit scoring were adopted with immediate effect. Premiums would go up for a very large number of policyholders if the collar on credit scoring (or any other risk variable for that matter) is set too narrow, because it would force an immediate price shock that would be unrelated to a change in risk. Further, I believe that, based on the analysis, a collar would have only visual effect, giving the public the impression that any disproportionate impact had been corrected once and for all. A collar would simply be a guess about what is publicly acceptable, and I have no valid, objective way of determining that number. Therefore, any action to ban or restrict the use of credit scoring must allow for changes in the pricing and underwriting systems to occur over a period of time, ensuring that all Texans pay a rate that is fair and based on risk.

Modern insurance pricing relies on the law of large numbers, which assumes that the more observations one makes, the greater the certainty. Credit scoring allows for a finer level of observation, but measuring propensity for risk strictly by the numbers can seem callous. Unlike other risk-related factors, credit scoring does not have that readily discernable, causal link to risk, such as driving record. As a result, credit scoring has earned the outward appearance of being a surrogate for something sinister. Unfortunately, there is no formula that reconciles economic reality with the public perception of fairness; it is a matter of guessing the right answer for the times.

Allowing credit scoring to be used, as contemplated under SB 14, 78th Regular Session, will ensure its link to risk under some of the strongest consumer protections in the nation, especially for people that suffer hardship. However, if the presence of credit scoring in insurance will only feed suspicion and divide us as Texans, its continued use to any degree may simply not be worth it. If the Legislature determines that credit scoring should be eliminated, then I recommend that it be phased out over time.

I hope the analysis and my thoughts on this matter help in your deliberations. As always, I remain at your service.

Sincerely,



Jose Montemayor
Commissioner of Insurance

c: Members of the 79th Texas Legislature

SUPPLEMENTAL REPORT TO THE 79TH LEGISLATURE

USE OF CREDIT INFORMATION BY INSURERS IN TEXAS

THE MULTIVARIATE ANALYSIS

Texas Department of Insurance
January 31, 2005

TABLE OF CONTENTS

EXECUTIVE SUMMARY

This supplemental report follows the Texas Department of Insurance's (Department) December 30, 2004 report to the Legislature (Use of Credit Information by Insurers in Texas) which noted a relationship between credit score and claim experience. This supplemental report further analyzes this relationship by considering the impact of other rating variables using a multivariate analysis. It is based on the same individual policyholder data for personal auto (bodily injury and property damage liability only) and homeowners described in the earlier report. Analysis of this type measures statistical relationships between variables and does not address the causal factors that may account for these relationships as it is beyond the scope of this study. The key findings follow:

- For both personal auto liability and homeowners, credit score was related to claim experience even after considering other commonly used rating variables. This means that credit score provides insurers with additional predictive information distinct from other rating variables. By using credit score, insurers can better classify and rate risks based on differences in claim experience.

- For personal auto liability, credit score was related to the probability of filing a claim or <u>claim frequency</u>. However, the Department found very little or no statistical evidence that credit score was related to the amount of a claim or <u>claim severity</u>.

- For homeowners, credit score was related to the probability of filing a claim. The Department found some statistical evidence that credit score was also related to the amount of a claim but was unable to draw a definite conclusion on this matter due to uncertainties caused by the unusual influences in the data (e.g., mold claims, large claims, and storm events).

- For personal auto liability, credit score varies in importance depending on the model the insurer is using, but was generally comparable in importance to territory and driving record for predicting claim experience. Only class (which reflects the age, gender and marital status of the driver combined with usage of the vehicle) was consistently a more important rating variable than credit score, driving record, or territory.

- For homeowners, credit score was one of several important rating variables for predicting claim experience. However, the Department was unable to draw definite conclusions regarding the relative ranking of rating variables due to the uncertainties mentioned above.

- For both personal auto liability and homeowners, the difference in claims experience by credit score was substantial. Typically, the claim experience for the 10 percent of policyholders with the worst credit scores was 1.5 to 2 times greater than that of the 10 percent of policyholders with the best credit scores. The magnitude of the variation noted in the earlier report remains unchanged even after considering other commonly used rating variables.

DISCUSSION

Introduction

This supplemental report follows the Texas Department of Insurance's (Department) December 30, 2004 report to the Legislature (Use of Credit Information by Insurers in Texas) which presented an analysis showing the relationship between credit score and claim experience. This analysis was univariate as it related claim experience to the single variable of credit score, without consideration of other rating variables that might impact claim experience. This univariate analysis showed that credit score was related to claim frequency, or the likelihood that a policyholder would file a claim. The relationship between credit score and the amount of the claim, or the claim severity,[1] was not as evident.

Since rating variables other than credit score are also related to claim experience, and many of them are related to credit score, it is important to ascertain whether the relationship between credit score and claim experience provides significant additional information, over and above the other rating variables. To study this issue, the Department conducted a multivariate analysis.

Data Used

The analysis presented in this report is based on data for four personal auto and two homeowners insurer groups (insurers). The origin and compilation of this data was described in the earlier report.[2] In total, the data represented the experience of almost 2 million vehicle years for personal auto liability and more than 600 thousand house years for homeowners.

Methods Used

Several multivariate methods were used to analyze the data. The methods are categorized as being either "statistical" or "actuarial."

The statistical methods involved construction of statistical models by fitting hypothetical probability distributions[3] to the individual data elements. The actuarial methods calculated relative claim experience by rating variable directly from the data and were used to cross check the conclusions of the statistical analysis. A more detailed discussion of these methods is available in the Appendix.

---

[1] For personal auto liability, all claims are limited to the minimum mandatory insurance amount ($20,000/$40,000/$15,000) in this analysis.

[2] See the section titled "Individual Policy Data Used in the Analysis" beginning on page 11 of the December 30, 2004, report.

[3] A probability distribution is a mathematical formula that specifies the probability of all possible outcomes.

Rating Variables Analyzed

The Department included several common rating variables in the multivariate analyses. These rating variables were defined by the personal lines manuals previously promulgated by the state. The variables included are summarized in Table A.

Table A: Rating Variables	
Personal Auto Liability	**Homeowners**
Territory[4]	Territory[5]
Class[6]	Protection Class[7]
Age of Driver	Construction Class[8]
Basic or Higher Limits[9]	Amount of Insurance[10]
Existence of Multi-Car Discount	Existence of Auto/Home Discount
Existence of Auto/Home Discount	Deductible
Driving Record Points[11]	Credit Score
Credit Score	

Procedure Used to Analyze the Data

The analysis was applied separately to each insurer (four insurers for personal auto and two insurers for homeowners) and line of business (i.e., auto bodily injury liability, auto property damage liability, and homeowners). The procedure was as follows:

1. Claim frequency and severity were analyzed using statistical models. Based on a review of the statistics, tentative conclusions were made regarding the statistical and practical significance[12] of the relationships between rating variables and claim frequency or severity.

2. The results of the actuarial approach were reviewed to ensure general consistency with the statistical approach in Item 1.

---

[4] Territory for personal auto refers to where the vehicle is garaged.
[5] Territory for homeowners refers to location of the home.
[6] Class reflects the age, gender and marital status of the driver combined with usage of the vehicle.
[7] Protection class refers to the quality of public fire protection facilities (e.g., proximity to fire station, fire hydrant, etc.).
[8] Construction class refers to the type of construction (brick, frame, etc.).
[9] Basic or higher limit refers to whether the policy involves the minimum required limits (i.e., basic limits) or optional higher coverage.
[10] Amount of insurance refers to the insured value of the residence structure.
[11] Driving record points reflect at-fault accidents and serious vehicle-related convictions.
[12] A statistically significant relationship refers to a relationship between two variables that has been subjected to a formal statistical test and has a p-value of less than 5 percent. A practical significance refers to the substantive importance of the relationship. For example, a relationship may be statistically significant but have such a small dollar impact as to be insignificant as a practical matter.

3. The statistical analysis of claim frequency was enhanced to exclude statistically insignificant variables. This resulted in a final statistical model involving relatively few variables, each with a strong statistical relationship to claim frequency. Again, the general conclusions from the final statistical model were compared to the results in Item 1.

4. The statistical models in Items 1 and 3 were used to measure the relative importance of the various rating variables.

Findings

The Department concluded that credit score provides insurers with additional predictive information, distinct from other rating variables, which an insurer can use to better classify and rate risks based on differences in claim experience.

Specific evidence for this conclusion was as follows:

1. In every case analyzed (i.e., all four personal auto and both homeowners insurers and for each personal auto coverage separately), the statistical approach identified credit score as being related in a statistically significant way to claim frequency. For personal auto liability, there was very little or no statistical evidence that credit score was related to claim severity. For homeowners, in some cases, the statistical models indicated a statistically significant relationship between credit score and claim severity.

2. In every case analyzed (i.e., all four personal auto and both homeowners insurers and for each personal auto coverage separately), both the statistical and actuarial approaches indicated substantial differences in claim experience over the range of possible credit scores. For both personal auto liability and homeowners, the difference in claims experience by credit score was substantial. Typically, the claim experience for the 10 percent of policyholders with the worst credit scores was 1.5 to 2 times greater than that of the 10 percent of policyholders with the best credit scores. The magnitude of the variation noted in the earlier report remains unchanged even after considering other commonly used rating variables.

For personal auto liability insurance, the Department concluded that class (which reflects the age, gender and marital status of the driver combined with usage of the vehicle) was consistently a more important rating variable for predicting claim experience. After class, credit score, driving record, and territory appeared as important rating variables. However, their ordering varied by insurer.

Specific evidence for this conclusion was obtained by ranking rating variables according to how much they added to predicting claim experience and to how much was lost when they were not used to predict claim experience.

For homeowners insurance, the Department concluded that credit score was one of several important rating variables for predicting claim experience. However, the Department was unable to draw definite conclusions regarding the relative ranking of rating variables.

The Department attempted to rank homeowners rating variables in the same manner that it did for personal auto liability but was unable to draw definite conclusions due to uncertainties caused by influences in the data. For example:

- A relatively small number of large claims accounted for a large portion of the total claim experience. This raises concerns that a small number of data points may drive the results.

- The rankings varied depending on whether water damage and weather-related claims were included in the claim experience. This is a concern because of the unusual circumstances regarding mold claims that were present in the data (but would be of much less concern if the analyses were based on current data). In addition, weather-related losses in Texas are erratic and vary substantially from year to year. For example, for claim experience caused by all covered perils (e.g., fire, wind, theft, etc.), territory was consistently the most significant rating variable when looking at claim frequency. Age of home and deductible (for perils insured other than wind and hail) follow next in importance with credit score being the fourth most important variable. When water damage and weather-related claims were excluded, the average rankings placed credit score as the most important variable followed by territory, deductible (for perils insured other than wind and hail), and age of home.

- Other variables in the analysis such as protection class, construction class, and amount of insurance, were of lower importance for claim frequency but would be expected to play a more important role in determining claim severity.

- While the time period varied somewhat by insurer, the data generally reflected experience for policies issued in 2001. More reliable patterns in the homeowners experience should be observed if more years of data in the study were available. This would help smooth the claim experience which can be erratic due to the nature of homeowners insurance (e.g., mold claims and weather related claims).

APPENDIX

This Appendix provides details of the considerations and methodologies used in the supplemental report.

Overview of the Insurance Transaction

In the insurance transaction, the insurer accepts a premium in exchange for an obligation to pay future claims incurred by the policyholder. For the insurer's business to be economically viable, the premium together with investment income earned on that premium should be sufficient to cover all the costs associated with that transaction. The accurate determination of future claim experience is of importance to the insurer as it represents a large portion of the cost of the insurance transaction.

In both personal auto and homeowners insurance, for any particular policyholder the insurer does not know what the future cost of insurance claims will be. For this reason, the insurer cannot base the premium on the claims that will actually be incurred in the future. Instead, the insurer bases the premium on the expected (or equivalently average) cost of future claims considering the probability that the particular policyholder will have one or more claims. While the insurer cannot know the future for any particular policyholder, it hopes that the claim experience of all its policyholders (in total or for each group of policyholders with the same risk characteristics) will average out and that in aggregate, actual costs will closely approximate expected costs.

The situation described above is analogous to flipping coins. Specifically, there is no way of knowing ahead of time whether a particular coin will turn up a head or a tail. However, if a large number of coins are flipped, it is possible to predict with a reasonable degree of certainty that approximately 50 percent will be heads.

Risk Classification

To determine the expected cost of claims, an insurer typically identifies a number of risk factors that it believes are related to the expected cost of future claims. Examples of risk factors commonly used for personal auto insurance include previous claims history, garage location, type and use of vehicle, driving record, age, marital status, credit score and existence of companion homeowners policy. Examples of risk factors commonly used for homeowners include amount of coverage, previous claim history of the policyholder, previous claim history on the insured property, quality of local fire department and facilities, type of construction and credit score.

The risk factors are then incorporated into a risk classification plan consisting of a schedule of all applicable risk factors that is used to classify policyholders. For example, a hypothetical auto insurer might classify the risk for bodily injury coverage in Texas as follows:

- Territory (50 groups based on garage location)
- Class (20 groups based on age, gender and marital status of the driver combined with usage of the vehicle)
- Driving record (5 groups based on past driving record of the driver)
- Credit score (10 groups based on credit score)
- Multi-car (2 groups based on number of vehicles insured)
- Companion policy (2 groups based on whether there is a homeowners policy issued to the policyholder)
- Past claims history (3 groups)
- Type of vehicle (3 groups)
- New or renewal policy (2 groups)

This hypothetical plan would thus classify Texas policyholders into 3.6 million different categories (= 50 territories x 20 classes x 5 driving record groups x 10 credit score groups x 2 multi-car groups x 2 companion policy groups x 3 past claims history groups x 3 types of vehicle x 2 new/renewal policy groups).

In addition to classifying an applicant for insurance (or in some cases a current policyholder), the insurer applies underwriting rules to determine if the applicant is eligible for coverage, and under what terms. For example, the insurer may have rules that make certain applicants ineligible for coverage based on their previous driving record. If an applicant is eligible for coverage, the insurer has a rating plan to determine a rate for that applicant. The rating plan consists of various formulas and factors that determine the rate for every possible combination of risk classifications written by the insurer. The factors that enter into the rate calculations are called rating variables.

Analyzing Risk Factors / Rating Variables

The analysis of risk factors and rating variables should involve a determination of the relationship (if any) between the factor and variable in question to claim experience (i.e., claim frequency and/or severity). Because there are usually many variables that simultaneously impact claim experience and these variables are often related among themselves, this analysis should consider the impact of a variable when all other variables are held constant. In other words, the analysis should be multivariate and consider the impact of all variables simultaneously. Regression methods including variants such as generalized linear models are statistical tools specifically designed for such purposes. Actuarial approaches are also designed for such purposes.

Overview of Regression Methods

In a regression analysis, one variable (known as the dependent variable) is related to the values of other "explanatory" or independent variables. In this way, the relationship (if any) between the dependent and independent variables can be examined and future values of the dependent variable can be predicted based on the independent variables. For example, a football fan trying to predict college results may study the relationship between the win/loss record of a team (i.e., the dependent variable) based on the weather, record of their opponent, the team's recent record, the salary of the coaches and other explanatory variables. A regression analysis may enable the football fan to determine whether a particular variable is related to the outcome of the games, to measure that variable's effect, and to better predict the results of future games. For the case at hand, the dependent variable is claim experience and the explanatory variables are the rating variables.

Regression analysis can be used to:

- Test whether the rating variables are in fact related to claim experience.
- Measure the impact (if any) of the rating variables on claim experience.
- Test whether a particular rating variable adds new information over and above information regarding claim experience that could be gleaned from some combination of all the other rating variables.
- Produce forecasts of future expected claim costs based on the rating variables.

Steps in a Multivariate Regression Analysis

Conducting a regression analysis involves the following steps:

1. Identification of the specific question to be answered.
2. Specifying the form of the regression model to be used.
3. Fitting the model to the actual data and reviewing statistics that measure how well the model fits the data.
4. Reviewing summary statistics to evaluate the nature of the relationships between the variables.

Sources for Further Information on Regression

For the interested reader, the theory behind regression analysis is discussed in detail in a large number of standard statistical texts.[13] In addition, The Reference Manual on Scientific Evidence (published by Federal Judicial Center, West Group) provides a good overview of regression analysis, how it should be evaluated, and questions it

[13] For example, Intermediate Business Statistics: Analysis of Variance, Regression and Time Series, by Miller and Wichern, and published by Holt, Rinehart and Winston provides an introduction to regression analysis. Generalized Linear Models with Applications in Engineering and Sciences, by Myers, Montgomery and Vining and published by Wiley provides an introduction to more advanced techniques.

raises. It is written for participants in the judicial process who are not statisticians but who have to evaluate expert statistical testimony.

Statistical versus Causal Relations

Regression analysis identifies and measures the relationship between the dependent variable and explanatory variables based on data. In other words, it identifies statistical correlations between the variables.[14] Variables may be correlated to each other but neither may be causally related to the other. For example, two variables may be correlated because they are both related to a third unexamined variable. In the case at hand, the analysis examines statistical correlations between credit score and claim experience. It does not provide any causal information on how this relationship arises.

The distinction between statistical and causal relationships can be appreciated by considering the contrast between credit score and driving record as rating variables. Specifically, the causal link between a bad driving record and a high number of insurance claims is obvious to most people. The causal link (if any) between a bad credit score and a high number of insurance claims is not as obvious.

The Department's Analysis of Rating Variables

The Department's analysis of rating variables proceeded as follows:

1. The Department assembled a team to conduct the analysis. This consisted of technical staff within the Department as well as an outside actuarial consultant. The Department contracted with consultants from the University of Texas and Texas A&M University who are experts in statistics, insurance and related fields. These consultants provided technical advice to the Department as well as critical review of the Department's work. In addition, the Department periodically discussed its approaches with, and received input from, staff at the Office of Public Insurance Counsel.

2. The Department performed a number of alternative exploratory analyses. These included the approaches finally selected (and discussed later) as well as logistic regressions. Many variations in the precise way variables could be included in the model were also performed. For example, some variables can be included either as a categorical variable (i.e., a variable that takes integer values) or as a continuous variable (i.e., a variable that takes any numerical value). Methods were also included that combined coverages as opposed to separating them. In addition, claim frequency and severity were considered separately as opposed to their combined impact. The methods that were finally used were selected over other possible alternatives for the following key reasons:

[14] Two variables are correlated positively if, on average, they move in the same direction; two variables are correlated negatively if, on average, they move in opposite directions.

a. For personal auto liability, the bodily injury and property damage liability coverages were analyzed separately rather than combined because the data indicated that the relationships between claim experience and rating variables varied by coverage.

b. Claim frequency and severity were analyzed separately because the data indicated the rating variables impact these quantities in different ways.

c. The generalized linear model using a Poisson probability distribution[15] was used to analyze claim frequency because it is a standard approach to problems of this type and has been used in other insurance and non-insurance applications. It was also found to be the best approach to account for varying policy periods.

d. The generalized linear model using a gamma distribution was used to analyze claim severity because it is one of several models commonly used for such purposes in the insurance industry. In addition, the exploratory analysis indicated that if other reasonable distributions were used, similar results were obtained.

e. All rating variables except credit score were included in the model as categorical variables. Credit score was included as a continuous variable. This approach was taken as the Department wished to measure the incremental impact on claim experience from each unit of credit score. In addition, the exploratory analysis indicated that if other variable structures were used, similar results were obtained (e.g., driving record points included as a continuous rather than categorical variable, credit score grouped into deciles and included as a categorical variable, etc.).

f. For personal auto liability, the analysis was performed based on the territory structure previously promulgated in Texas (based on counties or groups of counties), as well as a more refined territory structure based on groups of ZIP codes. These approaches were taken to test the impact of territory structure on the results (it had no significant impact on the conclusions regarding credit score). Some exploratory analysis was performed to test that credit score remained significant under even more refined territory structures. For homeowners, the Department analyzed the data using the previously promulgated territory structure in Texas, based on counties or groups of counties.

g. For personal auto liability, the analysis was performed based on the class structure previously promulgated in Texas. As there was some indication

[15] A Poisson probability distribution provides a formula for the probability of zero claims, the probability of one claim, the probability of two claims, etc.

that this system could be refined, driver age was also included as a rating variable.

h. The analysis was performed for each insurer separately due to differences in the credit score models used by the insurers.

i. For homeowners, the severity analysis was refined to include additional separate analysis of large and small claims.

j. For homeowners, the analysis was refined to include separate analyses for claim experience caused by all covered perils (e.g., fire, wind, theft, etc.) and for claim experience caused by all covered perils excluding water damage and weather-related claims.

k. For personal auto liability, the analysis was conducted with all losses limited to the minimum or basic mandatory limits of coverage. This was found to be the most practical way to combine data for policyholders that purchased different limits of coverage.

3. The statistical method used to analyze claim frequency was based on the assumption that the number of claims for a given policyholder is random with a Poisson probability distribution. This distribution implies that the expected number of claims is proportional to the length of a policy period and is commonly used in insurance to model the number of claims for a policy. It is also commonly used in other non-insurance applications for analogous situations. The statistical model assumed that, in addition to the length of the policy period, the number of claims for a given policyholder is a specified[16] function of the rating variables (e.g., class, territory, credit score) for that policyholder. With these assumptions, statistics were calculated (using SAS software[17]) to determine the parameters of the formula relating the rating variables to the expected number of claims, as well as to test the statistical significance of the relationships. The Department noted that some statistics (deviance divided by degrees of freedom) indicated the actual data were less dispersed than expected based on the Poisson assumption. The Department did exploratory analysis to make adjustments for this situation and determined that the degree of statistical significance indicated by the statistical models used in the analysis tends, if anything, to be understated.

4. The statistical method used to analyze claim severity was based on the assumption that the amount of a given claim is random with a gamma probability distribution (with adjustments reflecting the impact of deductibles and policy

[16] In the Department's analysis, it was assumed that the logarithm of the expected number of claims can be expressed as a linear function of the rating variables. This is equivalent to the assumption that rating variables should be applied in a multiplicative manner.
[17] Using the procedure PROC GENMOD

limits). This distribution is commonly used to model the severity of insurance claims.[18] The statistical model assumed that the average claim amount for a given policyholder is a specified[19] function of the rating variables (e.g., class, territory, credit score) for that policyholder. With these assumptions, statistics were calculated (using SAS software[20]) to determine the parameters in the formula relating rating variables to the expected claim severity as well as to test the statistical significance of the relationships.

5. A Chi-square statistical test was used to evaluate whether a set of rating variables (e.g., territories) were statistically significant as a group, or individually. This approach tests the null hypothesis that the variable(s) in question has a coefficient(s) of zero in the regression equation and so is not related to the dependent variable (i.e., claims experience). The p-value measures the probability that a Chi-square statistic equal or greater than that actually observed could arise under the null hypothesis. If the p-value is small, the null hypothesis was rejected and the rating variable accepted as being statistically significant. In the analysis, the Department used a threshold of 5 percent (or 0.05) to determine statistical significance (in practice, the p-values were often much smaller than this threshold by many orders of magnitude).

6. The statistics used to evaluate the models were "deviance" and "Chi-square." The reader may be familiar with a statistic known as the R-square statistic that is often used in different types of regression analysis. This statistic was not used in the Department's analysis because it is not applicable to the type of regression that was employed. However, because it is so commonly used, the Department offers the following observations. The R-square statistic ranges from zero to one and represents the portion of variation in the dependent variable (i.e., claim experience) that is accounted for by all explanatory variables (i.e., rating variables). The statistical models used in the Department's analysis generally had very low R-square statistics. The Department considered this situation and concluded that it should not raise concerns about the procedures used in the analysis or of the entire rating system for the following reasons:

    a. The models used by the Department were based on data for individuals rather than aggregate data. The low R-square statistic indicated that at the individual vehicle or house level, the claim experience in any given year contains a relatively large random component. This situation is analogous to considering the chance of flipping a head using a coin that may be fair (i.e., 50 percent chance of getting a head) or biased with a 55

[18] The Department did some exploratory analysis with other probability distributions that could also reasonably be used to model claim severity to verify that, within reason, the precise distribution assumed does not crucially impact the conclusions.
[19] In the Department's analysis, it was assumed that the log of the expected average severity can be expressed as a linear function of the rating variables. This is equivalent to the assumption that rating variables should be applied in a multiplicative manner.
[20] Using the procedure PROC LIFEREG

percent chance of getting a head. Knowing whether or not the coin is biased does not meaningfully improve the ability to predict whether the next flip will yield a head or not. This is because the inherent randomness in flipping a coin is a much more significant factor than knowing the relatively small difference in the probability of getting a head.

b. The R-square statistic in the context above measures the amount of variation explained at the individual level. It does not address the amount of variation that is explained at different levels of aggregation. For example, if the data were first aggregated into deciles (i.e., ten equal groups) based on credit score and then fit to a statistical model, a high R-square would result. This is because the action of aggregating the data effectively averages out the effects of variations in individual behavior. Continuing the coin analogy from (a) above, knowing whether or not the coin is biased meaningfully improves the ability to predict the proportion of heads in the next 1,000 flips.

c. A review of literature (e.g., Reference Manual on Scientific Evidence) indicated that a low R-square is typically encountered in other situations where analysis is conducted at the individual level and differences in individual behavior are involved.

d. The Department simulated hypothetical claims experience using an assumed territory and class structure together with assumed probability distributions and a random number generator. The Department then analyzed the data using regression methods. These simulations illustrated that very low R-square statistics typically resulted even if the regression model exactly mirrored the process used to generate the data. The examples also illustrated that the regression models accurately measured the underlying variables and their impact even when the R-square was very low.

7. A traditional actuarial method was used to calculate relativities for each of the rating variables in order to validate and confirm the results of the statistical method. The actuarial method was applied to claim experience per exposure[21] directly (rather than separately to claim frequency and severity) and involved calculating a set of relativities[22] for each rating variable describing the relative level of claim experience per exposure. For example, a territorial relativity of 0.80 would indicate that claim experience per exposure in that territory is 20 percent less than the statewide average. Similarly, a territorial relativity of 1.15 would indicate that claim experience per exposure is 15 percent more than the statewide average. Techniques are available to simultaneously calculate relativities for all rating variables. In this way, relativities are defined for each

[21] Here "exposure" means vehicle years for auto insurance or house years for homeowners.
[22] A relativity of 1.00 indicates an average level of claim experience per exposure.

policyholder so that they exactly capture the difference in claim experience when each rating variable is considered separately. Once the manner in which the rating variables are to be combined has been decided,[23] this determination is mechanical. The only judgment that is required concerns how the calculated results for rating classes with small numbers of exposures (i.e., “low credibility”) should be interpreted. Because the method does not require judgments concerning what model or probability distribution to use, it provides an objective method to confirm the conclusions from the more involved statistical approach.

8. The claim frequency models were enhanced to include variables describing industrywide[24] relativities for those rating variables (e.g., territory, protection class) where the volume of insurer specific experience may be too small to produce reliable estimates (as indicated by statistical evidence from the models). In the initial run, the variables were maintained in the model to account for any situations where the insurer’s claim experience might differ significantly from the industrywide average. The initial run was compared to the findings in Item 3 to verify that results of the different approaches were generally similar. The model was then refined by eliminating all variables that were found to be statistically insignificant and combining the age and class variables as appropriate and indicated by the statistical analysis. This resulted in a final model involving relatively few variables, each with a strong statistical relationship to the claim frequency. Again, the general conclusions from the final model were compared to the earlier results to ensure that they remained valid.

9. The final models in Item 8 were used to measure the relative importance of the various rating variables by re-running the model several times and each time excluding just one variable. Once a variable was removed, the predicted claims for that variable (using the model that excluded that variable) were compared to the actual claims. In this process, a large difference between the actual and predicted claims indicated that the absence of the rating variable in the model results in substantial aggregate estimation errors. By measuring these errors we were able to rank the importance of the various rating variables.

10. An alternative ranking of the relative importance of the various rating variables was obtained by considering the univariate relationship between the average claim frequency for the variable under consideration and the statewide average claim frequency per exposure. A large difference indicated an important rating variable as it exhibits substantial variation from the average.

11. By the measures in Items 9 and 10, the Department concluded that for personal auto, class (which refers to age, gender and marital status of the driver combined with usage of the vehicle) appears to be the most important rating

[23] The Department assumed that rating variables are to be multiplied together.
[24] The industrywide relativities were based on information collected under the Texas statistical plans for several years.

variable. After class, credit score, driving record, and territory appeared as important rating variables. Credit score appears to be much more important than the minor rating variables (i.e., multi-car discount, type of policy – basic versus higher limits, auto/home discount).