Does Credit Score Really Explain Insurance Losses? Multivariate Analysis from a Data Mining Point of View

Cheng-Sheng Peter Wu, FCAS, MAAA, and
James C. Guszcza, ACAS

Does Credit Score Really Explain Insurance Losses? Multivariate Analysis from a Data Mining Point of View

by

Cheng-Sheng Peter Wu and James Guszcza

Abstract

One of the most significant developments in insurance ratemaking and underwriting in the past decades has been the use of credit history in personal lines of business. Since its introduction in late 80's and early 90's, the predictive power of credit score and its relevance to insurance pricing and underwriting have been the subject of debate [1-3]. The fact that personal credit is widely used by insurers strongly suggests its power to explain insurance losses and profitability. However, critics have questioned whether the apparently strong relationship between personal credit and insurance losses and profitability really exists. Surprisingly, even though this is a hot topic in the insurance industry and in regulatory circles, actuaries have not been actively participating in the debate. To date, there have been few actuarial studies published on the relationship of personal credit to insurance losses and profitability. We are aware of only two such studies: one published by Tillinghast, which was associated with the NAIC credit study [4], and the other by Monaghan [5]. A possible reason for the lack of published data is that many insurers view credit scores as a confidential and cutting-edge approach to help them win in the market place. Therefore, they might be reluctant to share their results with the public. In this paper, we will first review the two published studies and comment on their results. We will then share our own experience on this topic. We have conducted a number of comprehensive, large-scale data mining projects in the past that included credit information as well as an extensive set of traditional and non-traditional predictive variables. Because our projects have been true multivariate studies, conducted using rigorous statistical methodology on large quantities of data, our experience should add value to the debate. Our experience does suggest that such a relationship exists even after many other variables have been taken into account.

About the Authors
Cheng-Sheng Peter Wu, F.C.A.S, A.S.A., M.A.A.A., is a director in the Advanced Quantitative Services practice of Deloitte & Touche's Actuarial and Insurance Consulting Group. He is based in the Los Angeles, CA office. Mr. Wu received his masters degrees in chemical engineering and statistics from the Pennsylvania State University. Mr. Wu has published several papers in automotive engineering, tribology (lubrication engineering), statistics, and actuarial science.

James Guszcza, A.C.A.S., M.A.A.A., is a manager in the Advanced Quantitative Services practice of Deloitte & Touche's Actuarial and Insurance Consulting Group. He is based in the Los Angeles, CA office. Mr. Guszcza received his Ph.D. in Philosophy from the University of Chicago.

Messrs. Wu and Guszcza's address is: Deloitte & Touche LLP, 350 South Grand Avenue, Los Angeles, CA 90071

Acknowledgments
The authors would like to thank Matt Carrier, Ravi Kumar, Jan Lommele, Sara Schlenker, and an anonymous referee for their thoughtful comments.

Introduction

One of the more important recent developments in the U.S. insurance industry has been the rapidly growing use of credit scores to price and underwrite personal auto and homeowners insurance. But this development has not come without controversy. Perhaps the most important criticism raised is that there exists no convincing causal picture connecting poor credit history with high insurance loss potential [1-5]. Partly for this reason, many insurance regulators and consumer advocates have expressed doubts that the observed correlations between credit scores and insurance loss history truly reflect an underlying reality. Some critics have suggested that these correlations might be spurious relationships that would not survive more sophisticated (multivariate) statistical analyses.

Given the business significance and statistical nature of this topic, it is curious that actuaries have not participated more actively in the debate. We are aware of only two actuarial studies that have been published so far: one published by Tillinghast, which was associated with the NAIC credit study [4], and the other by Monaghan [5].

The aim of this paper is to review these studies and complement them with a qualitative description of our own experiences in this area. For reasons of confidentiality, we are not able to share detailed quantitative results in this forum. Our focus will be on the use of credit in the line of personal auto, but many of our comments will hold true for other lines of insurance. We will begin with several historical comments on the development of auto classification ratemaking in the United States, and with comments on the actuarial issues relating to the use of credit in auto ratemaking.

The Development of Auto Classification Ratemaking in the United States

Personal auto ratemaking came a long way in the 20th century [6]. Prior to World War II, auto ratemaking involved only three classes: adult, youthful operator, and business use. The three decades after the war saw a proliferation of new class categories such as vehicle characteristics (symbol, model year) and refined driver classifications.

Today, a typical personal auto rating plan contains hundreds, if not thousands of classes involving the following variables:

- *Territorial Characteristics*: insurers define intra-state rating territories that reflect such relevant aspects of the physical environment as population density and traffic conditions.
- *Vehicle Use*: examples include business use, pleasure use, and driving more or less than a certain number of miles per year.
- *Driver characteristics*: examples are age, gender, marital status, and good student status
- *Driving Record*: this is reflected by a point system based on accidents and violations.

- *Vehicle Characteristics*: this typically includes a vehicle symbol system as well as a model year rating structure.
- *Miscellaneous surcharges/discounts*: this is where rating plans vary the most from company to company. Special surcharges or discounts are used to reflect policy characteristics or advances in motor vehicle technology. Commonly seen discounts include multi-car discounts, homeowner discounts, safe driver discounts, anti-lock brake discounts, anti-theft discounts, affinity group factors, and so on.

In addition to the above class variables, a typical rating plan is not complete without a *tier rating structure*. A tier structure is designed to address rating inadequacies that an insurer believes exists in a class plan. For example, an insurer might create three companies for its preferred, standard, and high-risk books, and the rate differential for such companies can range from –20% to 20%. Such differentials are typically applied at the policy level, across all coverages. Tier rating factors can include characteristics that are not used in the class plan, such as how long an insured has been with the insurer. They can also include certain interactions of class factors, such as youthful drivers with poor driving records.

As class plan structures have become more complex, the problem of estimating rates for each combination of class variables has become more difficult. This is because many of the variables used to define rating factors are not statistically independent. For this reason, factors based on univariate analyses of the variables are not necessarily appropriate for a multi-dimensional rating structure. Some form of multivariate analysis is called for.

To take a concrete example, suppose that an existing rating plan charges youthful drivers 3 times that of mature drivers. Furthermore, we analyzed loss (pure premium) relativities by driver age group, and noticed that the youthful driver group has losses per exposure 4 times that of the mature driver group. But it does not follow that the youthful driver rating factor should be raised to 4. This is because other variables used in the class plan might be correlated with age group variable. For example, youthful drivers have more accidents and violations; they are more likely to drive sports cars; they are more likely to be unmarried, and so on. They are therefore likely to be surcharged along these other dimensions of the rating plan. To give them a driver age rating factor of 4 would possibly be to over-rate them.

This issue -- that non-orthogonal rating variables call for multivariate statistical analyses -- lies at the heart of the debate over credit. In addition, this issue is perhaps the key theme in the *methodological* development of classification ratemaking since the 1960's.

McClenahan's ratemaking chapter [7] in *The Foundations of Casualty Actuarial Science* outlines the univariate approach to ratemaking, an approach still employed by many insurance companies. Appealing to examples like the one just given, Bailey and Simon [8,9] pointed out that the univariate approach could lead to biased rates if the individual rating factors are non-orthogonal. Their proposed solution to this problem, the *minimum* *bias* procedure, involves assuming a mathematical relationship between the rating factors and pure premium.

The mathematics of minimum bias is pure algebra: Bailey and Simon derived their models without positing statistical models. In his 1988 paper, Robert Brown [10] showed that commonly used minimum bias formulas could be derived from statistical models via maximum likelihood. Stephen Mildenhall's 1999 paper [11] is the most rigorous examination to date of the statistical underpinnings of the minimum bias method. Thanks to Brown, Mildenhall, and others [12, 13], it is now abundantly clear that Bailey-type actuarial analyses are in fact special cases of Generalized Linear Models. Multi-dimensional classification ratemaking projects should therefore be viewed as exercises in multivariate statistical modeling.

The lesson is obvious: a multivariate statistical analysis is necessary to establish the importance of credit for personal auto ratemaking.

How Credit is Currently Used in Personal Auto Ratemaking

During 1970's and 1980's, when classification ratemaking was undergoing its methodological development, no major rating *variables* were introduced. This changed in the late 1980's and 1990's when credit scores were introduced to personal lines insurance [1].

Raw credit information is supplied by several major credit bureaus, including Choice Point, TransUnion, and Experian. These companies collect individuals' credit data and in turn sell this data in the form of credit reports. Credit reports contain a wealth of information that can be grouped into four classifications:

- General information
- Trade line information
- Inquiries
- Public Records and Collections

The raw fields on these reports can be combined in many ways to create a plethora of random variables. Examples include number of trades, months since oldest trade, amount past due, trade line balance-to-limit ratio, number of inquiries, number of collections, and number of lawsuits. Using various statistical techniques (such as multiple regression, principal components analysis, clustering, Classification and Regression Trees) these random variables can in turn be combined to create *credit scores.*

Using credit scores to segment risks is hardly a new idea. For many years the lending industry has used such scores to underwrite loan applications. The Fair, Isaac Company is a leading vendor of one such score, called the FICO score.

Linking credit scores to personal auto and homeowners profitability, however, *was* a new idea, when they were introduced to the insurance industry approximately15 years ago. A typical credit score used in personal lines insurance might be calculated based on 10 to 30 variables. Conning's latest report [1] indicates that today more than 90% of insurance companies use credit scores or credit information in one way or another.

As noted above, the growing use of credit scores in insurance underwriting and ratemaking has garnered controversy along many fronts. We will set aside the political and social aspects of the debate and focus on the more purely actuarial issue: *do credit scores really help explain insurance profitability?* As we will discuss further, answering this question in the affirmative involves more than simply demonstrating a correlation between credit and loss ratio.

In the remainder of this paper, we will review the answers given to this question by the Tillinghast [4] and Monaghan [5] studies, and then add our own perspective. But first, it would be good to briefly discuss some general actuarial and statistical issues.

Background Actuarial and Statistical Considerations

Loss (Pure Premium) Relativity vs. Loss *Ratio* (Profitability) Relativity: The distinction between these concepts might not be clear to a non-actuarial audience, but it is absolutely critical. Because premium reflects all of the components of a rating plan, a correlation between a new variable (say, credit score) and loss ratio indicates the degree to which this variable can explain losses not already explained by the existing rating plan. For example, a critic might question the power of credit scores by claiming that credit is correlated with driver age. Since driver age is already in the class plan, there is no need to include credit as well. This argument would have some validity if it were in response to a pure premium relativity analysis. However, it would have much less validity if the relativity were based on loss ratios. Returning to the above example, the premium for youthful drivers is already 3 times that of mature drivers. Therefore a correlation between credit and loss ratio indicates the extent to which credit explains losses not already explained by the youthful driver surcharge.

Non-Independent Rating Variables: We believe that this is the key issue of the debate over the explanatory power of credit score. Intuitively, independence means that knowing the probability distribution of one variable tells you absolutely nothing about the other variable. Non-independence is common in insurance data. For example, youthful drivers have more accidents and violations than do mature drivers; mature drivers have more cars on their policies than do youthful drivers; number of drivers are correlated with number of vehicles. We can therefore expect that credit score will exhibit dependences with other insurance variables, such as driver age, gender, rating territory, auto symbol, and so on.

Univariate v. Multivariate Analyses: In the case of independent random variables, univariate analyses of each variable are entirely sufficient -- a multivariate analysis would add nothing in this case. Failure of independence, on the other hand, demands multivariate analysis. Furthermore, the results of multivariate analyses can be surprising. Below, we will give a hypothetical example in which an apparently strong relationship between credit and loss disappears entirely in a multivariate context.

Credibility vs. Homogeneity: paying attention to the credibility and homogeneity of one's data is important when we review any actuarial study and is essential in this debate for the usefulness of credit scores. Sparse data present the danger that one's model will fit noise rather than signal, leading to non-credible results. Non-homogenous data present the danger that extrapolating from one sub-population to another will lead to inaccurate predictions.

With these general remarks in hand, let us turn to the Tillinghast [4] and Monaghan [5] studies.

Tillinghast's Study

Tillinghast's credit study was undertaken on behalf of the Fair, Isaac Company for use in its discussions with the National Association of Insurance Commissioners (NAIC). The purpose of the study was to establish a relationship between Insurance Bureau credit scores with personal auto and homeowners insurance. Tillinghast received the following information for each of nine personal lines insurance companies:

- Credit score interval
- Interval midpoint
- Earned premium
- Loss ratio relativity

For the most part, the credit score intervals were constructed to contain roughly equal amounts of premium. The results for these 9 companies are given in Exhibit 1.

Clearly, the information provided to Tillinghast only allowed for a univariate study, and this is all Tillinghast set out to perform. Tillinghast's report displays tables containing each interval's loss ratio relativity alongside the interval's midpoint. These numbers are also displayed graphically. The report comments, "From simply viewing the graphs… it seems clear that higher loss ratio relativities are associated with lower Insurance Bureau Scores."

No detailed information is provided on the data used, or about the 9 companies that provided the data. Therefore we cannot comment on how credible the results are. The loss ratio relativity curves are somewhat bumpy for certain of the 9 companies; and the loss ratio spreads varies somewhat from company to company. But the patterns are clear enough to strongly suggests that the relativity spreads are robust, and not merely company-specific fluctuations in the data.

Furthermore, the relativities produced by credit are fairly large. The 10% of the companies' books with the best credit have anywhere from –20% to –40% loss ratio relativities. The worst 10% have relativities ranging from +30% to +75%. These loss ratio spreads compare favorably with those resulting from traditional rating variables. For example, based on our experience, about 20% to 30% of a standard auto book will have point surcharges for accidents or violations. The average surcharge might range from 15% to 40%. Therefore, the loss ratio spread indicated in the study is no less than the accident and violation point surcharge. In addition, the credit loss ratio spread can largely support the commonly seen rate differentiation for the tier rating. Examples such as this make it clear why insurers are embracing the use of credit scores.

In additional to displaying tabular/graphical evidence, Tillinghast computed regression slope parameters and their associated p-values. The p-values were all below 0.1, and often well below 0.05. (The p-value is defined as the probability of observing the actual slope parameter -- or a greater slope parameter -- given that the "true" slope parameter is zero.) The Tillinghast report concluded: "from the data and P-Values, we conclude that the indication of a relationship between Insurance Bureau Scores and loss ratio relativities is highly statistically significant."

Simpson's Paradox and the Perils of Univariate Analyses

This is reasonable as far as it goes. Unfortunately, univariate statistical studies such as Tillinghast's do not always tell the whole story. A statistical phenomenon knows as *Simpson's Paradox* [14,15] illustrates what can go wrong. A famous example of Simpson's Paradox is the 1973 study of possible gender bias in graduate school admissions at the University of California at Berkeley [16]. We will stylize the numbers for ease of presentation, but the point will remain the same.

Suppose it was reported 1100 men and 1100 women applied for admission to Berkeley in 1973. Of these people, 210 men were accepted for admission, while only 120 women were accepted. Based on this data, 19% of the men were accepted, while only 11% of the women were accepted. This is a univariate analysis (somewhat) analogous to Tillinghast's, and it seems to prove decisively that there was serious gender bias in Berkeley's 1973 graduate admissions.

But in fact this univariate analysis does not tell the whole story. When the admissions were broken down by division (suppose for simplicity that there were only two divisions: Arts & Sciences and Engineering) the data looked more like this:

	Applicants			# Accepted			% Accepted		
	Arts	Eng.	Total	Arts	Eng.	Total	Arts	Eng.	Total
Women	1000	100	1100	100	20	120	10%	20%	11%
Men	100	1000	1100	10	200	210	10%	20%	19%

Now our analysis is multivariate, by virtue of the fact that we are including *division applied to*, in addition to gender. The multivariate analysis quite clearly shows that the acceptance rate for men and women *within each division* was identical. But because a greater proportion of women applied to the division with the lower admission rate (Arts & Sciences), fewer women overall were accepted.

This is a very simple example of what can go wrong when one's data does not contain all relevant variables: *an apparent correlation between two variables can disappear when a third variable is introduced.*

In order to make the link to regression analysis, let us analyze this data at the un-grouped level. The reader can reproduce the following results with a simple spreadsheet exercise. Create 2200 data points with a {0,1}-valued target variable (ACCEPTED) and two {0,1}-valued predictive variables (MALE, ENGINEERING). 1000 of the points are males who applied to engineering {MALE=1, ENGINEERING=1}. For 200 of these points ACCEPTED=1, for the remaining 800 ACCEPTED=0, and so on.

If we regress ACCEPTED on MALE, we get the following results:

	Beta	**t-statistic**
Intercept	.1091	10.1953
MALE	.0818	5.0689

As expected, this univariate regression analysis indicates that gender is highly predictive of acceptance into graduate school, and indeed it is: a greater proportion of males *were* accepted! However this analysis is potentially misleading because it does not help *explain why* males are accepted at a higher rate.

When we regress ACCEPTED on MALE *and* ENGINEERING, we get quite different results:

	Beta	**t-statistic**
Intercept	.1	9.1485
MALE	0	0
ENGINEERING	.1	3.8112

When the truly relevant variable is introduced, the spurious association between gender and acceptance goes away (the beta and t-statistics for MALE are both 0). This multiple regression approach on un-grouped data is illustrative of our data mining work involving credit and other predictive variables.

(Of course logistic regression is usually a more appropriate way to model a binary target variable such as application acceptance or auto claim incidence. But such an analysis could not easily be replicated in a spreadsheet. Because ordinary multiple regression gives the same results in this simple case, it is sufficient for our illustrative purpose.

Readers are encouraged to try logistic regression, from which precisely the same conclusion will be reached.)

Returning to the Tillinghast study, consider the following scenario: suppose our credit variable has two levels (good/bad). Rather than academic division, suppose that the "true" confounding variable is urban/rural (territory). Thus good/bad correspond to male/female in the Berkeley example, and urban/rural corresponds to arts/engineering. Rather than acceptance into school, the target variable is now having a personal auto claim. Now our data is:

	Exposures			# Claims			Claim Freq		
	Rural	Urban	Total	Rural	Urban	Total	Rural	Urban	Total
Good credit	1000	100	1100	100	20	120	10%	20%	11%
Poor credit	100	1000	1100	10	200	210	10%	20%	19%

If we similarly re-label the terms of our regressions, we will again see that (in this purely hypothetical example) the GOOD_CREDIT indicator loses its apparent significance once the URBAN indicator is introduced.

These considerations make it clear that a multivariate analysis is needed to assess whether credit history bears a *true* relation with insurance loss experience. A univariate analysis might produce a statistical illusion, not true insight.

Of course, given our discussion of the difference between a pure premium study and a loss ratio study, it is not entirely fair to call the Tillinghast study "univariate". Recall that Tillinghast's target variable was *loss ratio relativity*, not claim frequency. In the above example, suppose all claims have a uniform size of $1000, and further suppose that the territorial rates are $2000 for urban territories, and $1000 for rural territories. Now the loss ratio relativity in each cell will be exactly 1.0. In this (again, purely hypothetical) case, Tillinghast's methodology would (correctly) show no relationship between credit and loss ratio relativity.

In other words, to the extent that all possible confounding variables are perfectly accounted for in premium, Tillinghast's "univariate" analysis *is* implicitly a multivariate analysis, and is therefore convincing. But realistically, this may not be the case. For example, in our work we regularly regress loss ratio on such zip code-based variables as population density and median population age. If territory were entirely accounted for in premium, such variables would never appear statistically significant. But in fact they sometimes do. Therefore a true multivariate study is desirable even if loss ratio is used as the target variable.

Monaghan's Study

James Monaghan's paper on "The Impact of Personal Credit History on Loss Performance in Personal Lines" is an advance over the Tillinghast study partly because he addresses the multivariate issue. Monaghan asks: if the correlation between credit and loss ratio exists, "is it merely a proxy, i.e., is the correlation actually due to other characteristics (which may already be underwritten for or against, or rated for)?" And, "are there dependencies between the impact of credit history on loss performance and other policyholder characteristics or rating variables?"

Monaghan's study for auto is based on three calendar years of data (1993-95). Each record in his database contains premiums and losses accumulated over this entire three-year period. So each record may have different length for the term. Losses are evaluated at 6/30/1995. For this reason, losses on different records might be evaluated at varying states of maturity. Losses include reserves, salvage and subrogation recoveries, and allocated loss adjustment expenses. The credit information used in this study was a "snapshot view" taken at the policy inception date. Approximately 170,000 records were used in the analysis. The total premium and loss in these records were $393 million and $300 million, respectively.

The amount of data in Monaghan's study is very large. While we don't know all the details about the data, the large amount of premium indicates that it is probably based on a countrywide population. Our experience on auto data indicates that on average there will be 150 to 400 claims per $1 million in premium, depending on the geographic concentration, program type, and policy type (liability only vs. full coverage) represented in the data. This suggests that there will be on the order of a hundred thousand claims in Monaghan's study. According to actuarial credibility theory [17], Monaghan's data should provide very credible results.

Monaghan discusses credit variables from a number of angles. First, he performs a number of univariate studies comparing *individual* credit variables (such as Amounts Past Due, Derogatory Public Records, Collection Records, Age of Oldest Trade, Number of Inquiries, Account Limits, and Balance-to-Limit Ratios) with fitted loss ratio relativity. In each case, there exists a positive correlation. This part of Monaghan's study is much like Tillinghast's study. The difference is that Monaghan analyses individual credit variables, whereas Tillinghast analyses a composite credit *score*.

While not conclusive for the reasons given above, this part of Monaghan's study is helpful in that it unpacks credit score into its component variables. The relationship between credit score is not entirely the result of some mysterious or proprietary interaction of the components credit variables. Rather, each of these component variables is individually somewhat predictive of insurance losses. For the record, the results Monaghan reports in this section are consistent with our experience working with credit data.

Note that these univariate results -- as well as Monaghan's multivariates to be described below -- are in terms of loss *ratio* relativity. Therefore, Monaghan's work (like the Tillinghast study) indicates the degree to which credit is able to capture loss variation not captured by the existing rating plan.

Next, Monaghan studies credit in conjunction with several traditional underwriting characteristics. Monaghan uses the above credit variables to profile policies into four groups, A, B, C, and D. For example, group A (the profile with the worst loss ratio) is characterized by one or more derogatory public records, high amounts past due, and so on. Group D (the profile with the lowest loss ratios) is characterized by long credit histories, low balance-to-limit ratios, and so on. Consistent with Monaghan's earlier results and Tillinghast's study, Monaghan shows that group A has a loss ratio relativity of 1.33; and group D has a relativity of 0.75.

Monaghan displays several two-way tables showing loss ratio relativity by credit group and an underwriting variable. The auto underwriting variables he displays in conjunction with credit include past driving record, driver age, territory, and classical underwriting profile. The last variable is a composite variable combining marital status, multicar, homeowner, and clean driving record. (Monaghan supplies similar tables for homeowners rating variables. We will not review the specifics of these tables here.)

In no case did Monaghan's inclusion of the rating factor cause the relationship of credit with loss ratio to disappear (as in the Simpson illustration above). Indeed, Monaghan's tables contain some very telling relationships. For example, the loss ratio relativity of drivers with clean driving histories and poor credit was 1.36. In contrast the relativity for drivers with good credit and poor driving records was only 0.70!

It is possible to reinforce Monaghan's conclusions by performing multivariate calculations on his data. Rather than use Bailey's iterative minimum bias equations, we performed equivalent Generalized Linear Model calculations using the PROC GENMOD facility in SAS. Recall [11,12] that the multiplicative Bailey model is equivalent to a GLM with the Poisson distribution and log link function; the additive Bailey model is equivalent to a GLM with the normal distribution and the identity link function. Note also that this latter model is simply a classical multiple regression model. Exhibits 2-4 contain GLM analyses of credit group by Driver Record, Driver Age, and Classical Underwriting Profile.

The results of the GLM analyses are striking, and they buttress Monaghan's claims. For example, the multiplicative Bailey factors arising from the credit/driving record analysis are 1.709, 1.339, 1.192, and 1.0 for credit groups A-D. These are quite close to the univariate loss ratios relativities that can be calculated from Monaghan's data (1.757, 1.362, 1.204, 1.0). This is excellent confirmation that credit is largely uncorrelated with driving record: the multiplicative Bailey factors are almost the same as the factors that would arise from a univariate analysis!

Furthermore, the GLM parameter estimates are quite large relative to their standard errors. Also, the Chi-squared statistics for the four credit groups are high, and the associated p-values are very low. These observations add statistical rigor to the claim that the loss ratio "lift" resulting from credit score is "real". These observations hold equally well for the other two variables as well. Finally, performing an additive Bailey analysis (normal/identity GLM – not shown) produces qualitatively similar results.

Monaghan reports that he produced such two-way tables for a large number of other traditional underwriting characteristics. He says, "there were no variables that produced even roughly uniform results across the credit characteristics."

Applying Data Mining Methodology to Credit Data

For several years, we have applied data mining methodology and a range of predictive modeling techniques to build insurance profitability and underwriting models for writers of both commercial and personal lines insurance. Credit variables and credit scores are typically included along with a comprehensive set of other traditional and non-traditional insurance variables. Because of the truly multivariate context in which we employ credit information, our findings lend further support to the conclusions reached in the Tillinghast and Monaghan studies. For reasons of confidentiality, we are not at liberty to share quantitative results in this paper. However, we shall describe our methodology and modeling results in a qualitative way.

We follow a standardized, disciplined methodology when embarking upon a data mining project. The first several steps involve studying internal and external data sources and generating predictive variables. Typical internal data sources include statistical records for premiums and losses, "snapshot" data for policyholder characteristics from legacy systems or a data warehouse, driver data, vehicle data, billing data, claims data, an agent database, and so on. Typically, several years of the company's relevant data sources will be utilized in the study. Commonly used external data sources include credit reports of the kind used by Monaghan, MVR (Moving Violation Records) data and CLUE (Claims Loss Underwriting Exchange) data. But other external data is available. For example, useful predictive variables at the zip-code level can be generated from data available from the US Census Department and the US Weather Bureau.

By the end of this process, literally *hundreds* of predictive variables will have been created from the internal and external data sources. The goal is to create upfront as many variables as possible that might be related to insurance loss and profitability. These variables represent a wide range of characteristics about each policyholder.

Typically we design our analysis files in such a way that each data record is at a policy-term level. For example, personal auto policies usually have a six-month term. If a policy has two years of experience in our study, we will generate four 6-month term data points in the study. This design, which is different from that of Monaghan's study, will give each record equal weight for the term in the analysis process. All of the predictive variables, including the credit variables, are evaluated as of the beginning of the term-effective date.

Target variables, including loss ratio, frequency, and severity, are created in parallel with the predictive variables. Losses are usually evaluated a fixed number of months from the term effective date. The reason for this is to minimize any chance of bias appearing in the target variables due to varying loss maturities. In addition, we will incorporate various actuarial techniques that we deem necessary to adjust the target information. Such adjustments include loss trending, premium on-leveling, re-rating, loss capping, cat loss exclusion, and so on.

Once the generation of target and predictive variables has been accomplished, we will merge all the information together to produce a policy-term level database. This database contains all of the predictive variables, as well as such target information as claim frequency, claim severity, loss ratio, capped loss ratio, and so forth. The database is then used to produce univariate reports showing the relationship of each predictive variable with the target information. This is essentially a collection of reports containing one Tillinghast-type study for each of the hundreds of predictive variables. This database is a useful exploratory data analysis (EDA) prelude to the multivariate modeling phase of our projects.

This database of univariate results also provides invaluable information for multivariate modeling regarding (1) whether to discard the variable right away because it has no/little distribution or because there is any business or other reason to do so; (2) how to cap the variable either above or below; (3) what to do with missing values; and (4) whether to treat the variable as a continuous or categorical random variable. Other needed transformations might be suggested by this univariate study.

Once the Exploratory Data Analysis stage is completed, we are ready to begin the modeling process. The first sub-phase of this process is to search for an optimal multiple regression model. Criteria used to judge "optimality" include (but are not limited to) strong t-statistics, parameter estimates that agree with business intuition, and not overfitting data used to estimate the parameters. This model serves as a useful benchmark for comparison purposes. In addition, the parameter estimates, and the t- and F-statistics generated by regression models are useful for such interpretive issues as the topic of this paper.

Once the optimal regression model has been selected, we turn to more advanced model building techniques such as Neural Networks [18-20], Generalized Linear Models [8-13], Classification and Regression Trees (CART) [21] and Multivariate Adaptive Regression Splines (MARS) [22]. These more advanced techniques can potentially provide more accurate predictions than a multiple regression model, but this additional predictive power often comes at a cost: more complex models can be harder to interpret and explain to underwriters, upper management, and insurance regulators.

We use a *train/test* methodology to build and evaluate models. This means that the modeling dataset is randomly divided into two samples, called the training and test samples. A number of models are fit on the training sample, and these models are used to "score" the test sample. The test sample therefore contains both the *actual* loss ratio (or any other target variable) as well as the *predicted* loss ratio, despite the fact that it was not used to fit the model. The policies in the test sample are then sorted by the score, and then broken into (for example) ten equal-sized pieces, called *deciles*. Loss ratio, frequency, and capped loss ratio are computed for each decile. These numbers constitute *lift curves*. A model with a low loss ratio for the "best" decile and a very high loss ratio for the "worst" decile is said to have "large lift". We believe that the lift curves are as meaningful for measuring the business value of models as such traditional statistical measures as mean absolute deviation or R^2. The purpose of setting aside a test set for model evaluation is to avoid "*overfit*". (Of course a lift curve can also be computed on the training dataset. Naturally, this lift will be unrealistically large.) A third sample, called a *validation* sample, sometimes will also be set aside to produce an unbiased estimate of the future performance of the final selected model.

We have performed several large data mining projects that included credit variables and credit scores. Similar to the Tillinghast study and Monaghan's study, we have studied data from various sources, different distribution channels, and different geographic concentrations. Our studies are very large in size, similar to Monaghan's study, usually with several hundred thousand data points that contain a total of hundreds of millions of dollars of premium. Our approach is tailored to the use of large datasets, the use of train/test methodology, the use of lift curves to evaluate models, and the exploratory use of a variety of modeling techniques. These are all hallmarks of the data mining approach to statistical problems. We believe that our analyses are true multivariate analyses that yield very robust and credible results. It is precisely this kind analysis that makes it possible to decisively answer the question: does credit *really* help explain insurance losses and profitability?

Our Findings: the Importance of Credit Variables in a Data Mining Context

First, through our univariate databases we note that composite credit score and many of its associated credit variables invariably show strong univariate relationships with frequency, severity, and loss ratio. Our univariate experience is entirely consistent with that of Tillinghast and Monaghan.

Turning to our multivariate modeling work, the estimates and statistics coming from our multiple regression models are useful for evaluating the importance of credit relative to the other variables considered in our model building process. Several points are worth making. First, credit variables consistently show up as among the most important variables *at each step* of the modeling process. As noted by Tillinghast and Monaghan, they dependably show strong univariate relationships with loss ratio. Furthermore, they are typically among the first variables to come out of a stepwise regression analysis.

Second, the parameter estimates for credit variables are consistently among the strongest of the parameters in our regression models. As illustrated in the Simpson's paradox example, credit score would have a small beta estimate and t-statistic were it a mere proxy for another variables or some combination of other variables. But this is not the case. Rather, we have repeatedly seen that credit adds predictive power *even in the presence of a comprehensive universe of traditional and non-traditional predictive variables, all used in conjunction with one another, on a large dataset.*

We are basing our conclusion in part on the t-statistics of the credit variables in our underwriting/pricing regression models. To this one might object: "but one of the assumptions of regression analysis is a normally distributed target variable. It is obvious that loss ratio is not normally distributed, therefore your t-statistics are meaningless." In response, it is true that loss ratios are not normally distributed. Nevertheless, the models we build using regression analysis reliably produce strong lift curves on test and validation data. Therefore, our models do "work" (in the sense of making useful predictions) in spite of the lack of normality.

It is also true that because of the lack of normality, we cannot use our models' t-statistics to set up traditional hypothesis tests. But neither our analyses nor our conclusions are based on hypothesis tests. We interpret t-statistics as measures of the *relative importance* of the variables in a model. Consider ranking the variables in a regression model by the absolute value of their t-statistics. The resulting order of the variables is the same as the order that would result from ranking the variables by their marginal contribution to the model's R^2 (in other words the additional R^2 that is produced by adding the variable after all of the other variables have been included in the model). This interpretation of t-statistics does *not* depend on the normality assumption.

To summarize, our reasoning is as follows:

- Our models effectively predict insurance losses. The evidence for this is repeated, unambiguous empirical observations: these models dependably distinguish profitable from unprofitable policies on out-of-sample data. In other words, they produce strong lift curves on test and validation datasets.

- Furthermore, credit variables are among the more important variables in these models. This is evidenced by the following observations: (i) the univariate relationship between credit and loss ratio is as strong or stronger than that of the other variables in the model. (ii) Credit variables reliably appear in a stepwise regression performed using all of the available variables. (iii) Credit variables typically have among the largest t-statistics of any of the variables in the model.

- Supporting the above observations, removing the credit variable(s) from a model generally results in a somewhat dampened lift curve.

- The implication of the above two bullets is that credit variables add measurable and non-redundant predictive power to the other variables in the model. Therefore, we believe that the observed correlation between credit and loss ratio cannot be explained away as a multivariate effect that would go away with the addition of other available variables.

Furthermore, this is true not just of the final selected regression model, but of most or all of the models produced along the way. In addition, we have noticed this result applies in all different lines of insurance, in both personal lines and commercial lines. For this reason, we feel comfortable saying that credit bears an unambiguous relationship to insurance loss, and is not a mere proxy for other *available* kinds of information.

But *Why* is Credit Related to Insurance Losses?

It is important to emphasize the word *available* because poor credit is obviously not in itself a *cause* of poor loss experience. In this sense, it is analogous to territory. Presumably credit is predictive because it reflects varying levels of "stress", planning and organization, and/or degrees of risk-taking that cannot be directly measured by insurers. These specific conjectures have been offered many times and they are intuitively plausible. However it is less conjectural to say that whatever credit might be a proxy for, it is not a proxy for any other variable (or combination of variables) practically available to insurers. In our data mining projects we explicitly set out to generate the most comprehensive universe of predictive variables possible. In this sense, we therefore use credit in the "ultimate" kind of multivariate analysis. Even in this truly multivariate setting, credit is indicated to have significant predictive power in our models.

It is beyond the scope of this paper to comment on the societal fairness of using credit for insurance pricing and underwriting. From a statistical and actuarial point of view, it seems to us that the matter is settled: credit *does* bear a real relationship to insurance losses.

Conversely: Can We Predict Insurance Losses without Credit? Can We Go beyond Credit?

Our experience does indicate that credit score is a powerful variable when it is used *alone* for a standard rating plan. In addition, our large-scale data mining results suggest that just about any model developed to predict insurance profitability will be somewhat stronger with credit than without credit. Typically credit score, when added to an existing set of non-credit predictive variables, will be associated with a relatively large beta estimate and t-statistic. Consistent with this, the resulting model will have higher "lift" than its counterpart without credit.

The results we have described might create an impression that credit variables are an essential part of any insurance predictive modeling project. But this would be an exaggeration. *Our experience also shows that pricing and underwriting models created without credit variables can still be extremely good.* The key to building a non-credit predictive model is to fully utilize as many available internal data sources as possible, incorporate other types of external information, use large amount of data, and apply multivariate modeling methodologies. Given all the regulatory and public policy issues surrounding insurers' use of credit, such non-credit models provide the insurance industry with a valuable alternative to using credit scores for pricing and underwriting.

Conclusion: Predicting the Future

Our data mining projects are multivariate predictive modeling projects that involve hundreds of variables being used to analyze many thousands of records. Many of these variables are credit variables, which play an important role even in this broad context. Our experience using credit scores and credit variables in a truly multivariate statistical setting has allowed us to add a new perspective to the debate over credit.

The use of credit in insurance underwriting and ratemaking might seem like a rather specialized topic. But we believe the issue reflects two important trends in the development of actuarial science. First, credit scores come from a non-traditional data source. The advent of the Internet makes it likely that other new data sources will become relevant to actuarial practice. Credit information is probably just the beginning.

The second issue is the increasingly multivariate nature of actuarial work. Credit scores themselves are inherently "multivariate" creatures in that they are composites built from several underlying credit variables. In addition, recall that we have reviewed and discussed three ways of studying the relationship between credit scores and insurance losses and profitability. Each study has been progressively more multivariate than its predecessor. This reflects the methodological development of classification ratemaking from univariate to multivariate statistical analyses (Generalized Linear Modeling).

In our opinion, the adoption of modern data mining and predictive modeling methodologies in actuarial practice is the next logical step in this development. Bailey's minimum bias method might seem like actuarial science's in-house answer to multivariate statistics. On the contrary, Mildenhall's paper makes it clear that conceptually, nothing separates minimum bias from work done by mainstream statisticians in any number of other contexts. But why stop at Generalized Linear Modeling?

We live in an information age. The availability of new data sources and cheap computing power, together with the recent innovations in predictive modeling techniques allow actuaries to analyze data in ways that were unimaginable a generation ago. To paraphrase a famous logician, actuaries inhabit "a paradise of data". This, together with our insurance savvy and inherently multivariate perspective, puts us in an excellent position to benefit from the data mining revolution.

Given the success of credit scores and predictive modeling, we expect actuaries to be enlisted to push this type of work even further. Here are examples of future questions we anticipate being asked of actuaries:

- Are we currently getting the most predictive power out of the internal and external information/data sources that we are currently using? Are we really analyzing data in a rigorous multivariate fashion?

- What other powerful variables and data sources are "out there" that we are not aware of? How do we go beyond credit?

- Are there other ways insurance companies (and indeed other kinds of companies) can leverage predictive modeling? For example, predictive modeling has a proven record of success in such applications as target marketing, customer retention/defection analysis, predicting cross-sales, customer profiling, and customer lifetime value. These are all important projects at which actuaries can excel. Furthermore, they are not insurance-specific. An actuary with expertise in these areas could transfer his or her skills to other industries.

To conclude, our multivariate predictive modeling work supports the widely held belief that credit scores help explain insurance losses, and that they go beyond other sources of information available to insurers. However it is unclear to what extent insurers will be permitted to used credit for future pricing and underwriting. For this reason insurers might want to consider non-credit scoring models as an alternative to traditional credit scores. For actuaries, the use of credit scores and predictive modeling is the beginning of a new era in insurance pricing and underwriting.

References

1. "Insurance Scoring in Personal Automobile Insurance – Breaking the Silence", *Conning Report,* Conning, (2001).

2. "Insurers Battling Credit-Scoring", *National Underwriter*, March 5th Issue, (2002).

3. "Insurers Lose a Credit Scoring Battle", *National Underwriter*, February 21^{st} Issue, (2002).

4. "Credit Reports and Insurance Underwriting", *NAIC White Papers*, National Association of Insurance Commissioners, (1997).

5. Monaghan, J. E., "The Impact of Personal Credit History on Loss Performance in Personal Lines", *CAS Forum*, Casualty Actuarial Society, (2000).

6. Finger, R. J., "Risk Classification", *Foundations of Casualty Actuarial Science*, Chapter 5, Casualty Actuarial Society, (1990).

7. McClenahan, C.L., "Ratemaking", *Foundations of Casualty Actuarial Science*, Casualty Actuarial Society, Chapter 2, (1990).

8. Bailey, R. A., "Insurance Rates with Minimum Bias", *Proceedings of Casualty Actuarial Society,* Vol. L, Casualty Actuarial Society, (1960).

9. Bailey, R. A. and Simon, L. J., "Two Studies in Auto Insurance Ratemaking", *Proceedings of Casualty Actuarial Society,* Vol. XLVII, Casualty Actuarial Society, (1963).

10. Brown, R. L., "Minimum Bias with Generalized Linear Models", *Proceedings of Casualty Actuarial Society,* Vol. LXXV, Casualty Actuarial Society, (1988).

11. Mildenhall, S.J., "A Systematic Relationship Between Minimum Bias and Generalized Linear Models," *Proceedings of Casualty Actuarial Society,* Vol. LXXXVI, Casualty Actuarial Society, (1999).

12. Holler, K.D., Sommer, D.B.; and Trahair, G., "Something Old, Something New in Classification Ratemaking with a Novel Use of GLMS for Credit Insurance," *CAS Forum*, Casualty Actuarial Society, (1999).

13. Brockman, M. J., Wright, T. S., "Statistical Motor Rating: Making Effective Use of Your Data," *Journal of the Institute of Actuaries*, Vol. 119, Part III.

14. Simpson, E. H., "The Interpretation of Interaction in Contingency Tables," *Journal of the Royal Statistical Society, Series B*, 13: 238-241, (1951).

15. Pearson, K., Lee, A., Bramley-Moore, L., "Genetic (Reproductive) Selection: Inheritance of Fertility in Man", *Philosophical Transactions of the Royal Society A*, 73: 534-539, (1899).

16. Bickel, P. J., Hamel, P.A., O'Connell, J.W., "Sex Bias in Graduate Admissions: Data from Berkeley", Science, (1975).

17. Mahler, H.C., and Dean, C.G., "Credibility", *Foundations of Casualty Actuarial Science,* Casualty Actuarial Society, Chapter 8, (2001).

18. Francis, L., "Neural Networks Demystified," *CAS Forum*, Casualty Actuarial Society, (2001)

19. Wu, C. P., "Artificial Neural Networks – the Next Generation of Regression" *Actuarial Review,* Casualty Actuarial Society, Vol. 22, No. 4, (1995).

20. Zizzamia, F., Wu, C. P., "Driver by Data: Making Sense Out of Neural Networks", *Contingencies*, American Academy of Actuaries, May/June, (1998).

21. Breiman, L., Friedman, J. H., Olshen, R. A., and Stone, C. J., "Classification and Regression Trees", Monterrey: Waldsworth and Brooks/Cole, (1984).

22. Friedman, J. H., "Multivariate Adaptive Regression Splines," *Annals of Statistics*, Vol. 19, (1991).

Exhibit 1
Tillinghast -NAIC Study of Credit Score [4]

Company 1
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
813 or more	850.0	10.2%	0.657
768-812	790.0	9.9%	0.584
732-767	749.5	11.0%	0.692
701-731	716.0	10.9%	0.683
675-700	687.5	10.4%	1.184
651-674	662.5	9.8%	0.793
626-650	638.0	9.9%	1.332
601-625	613.0	10.0%	1.280
560-600	580.0	9.4%	1.214
559 or less	525.0	8.6%	1.752

Company 2
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
840 or more	854.0	10.0%	0.607
823-839	831.0	10.0%	0.813
806-822	814.0	10.0%	0.626
789-805	797.0	10.0%	1.342
771-788	779.5	10.0%	1.059
748-770	759.0	10.0%	1.019
721-747	734.0	10.0%	1.322
686-720	703.0	10.0%	0.810
635-685	660.0	10.0%	0.986
635 or less	592.0	9.9%	1.417

Company 3
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
826 or more	845.0	10.0%	0.723
803-826	814.5	10.0%	0.903
782-803	792.5	10.0%	0.895
759-782	770.5	10.0%	0.795
737-759	748.0	10.0%	1.073
710-737	723.5	10.0%	0.941
680-710	695.0	10.0%	0.912
640-680	660.0	10.0%	1.115
583-640	611.5	10.0%	1.221
583 or less	535.0	10.0%	1.421

Company 4
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
832 or more	859.0	10.0%	0.672
803-832	817.5	10.0%	1.027
767-803	785.0	10.0%	0.823
739-767	753.0	10.0%	1.036
720-739	729.5	10.0%	0.775
691-720	705.5	10.0%	1.000
668-691	679.5	10.0%	1.041
637-668	652.5	10.0%	1.023
602-637	619.5	10.0%	1.251
602 or less	571.0	10.0%	0.135

Company 5
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
845 or more	857.0	10.0%	0.800
830-845	837.5	10.0%	0.919
814-830	822.0	10.0%	0.740
798-814	806.0	10.0%	0.733
779-798	788.5	10.0%	0.855
757-779	768.0	10.0%	0.889
730-757	743.5	10.0%	0.993
695-730	712.5	10.0%	1.143
643-695	669.0	10.0%	1.300
643 or less	600.0	10.0%	1.628

Company 6
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
810 and up	837.5	19.7%	0.656
765-809	777.0	20.1%	0.795
715-764	739.5	20.8%	0.911
645-714	679.5	20.2%	1.066
Below 645	600.0	19.2%	1.593

Company 7
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
750 and up	795.0	21.3%	0.783
685-749	717.0	25.8%	0.900
630-684	657.0	19.6%	1.083
560-629	594.5	19.3%	1.150
Below 560	520.0	13.9%	1.200

Company 8
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
755 or more	775.0	8.9%	0.767
732-754	743.0	9.3%	0.798
714-731	722.5	9.6%	0.859
698-713	705.5	9.9%	0.969
682-697	689.5	10.3%	0.922
666-681	673.5	9.7%	0.978
647-665	656.0	10.5%	1.070
625-646	635.5	10.2%	1.107
592-624	608.0	10.7%	1.122
591 or less	562.0	10.8%	1.324

Company 9
Scores & Loss Ratio Relativity Summary

Score Interval	Midpoint	Earned Premium	Loss Ratio Relativity
780 and up	815.0	16.8%	0.637
745-779	762.0	13.7%	0.715
710-744	727.0	13.9%	0.734
670-709	689.5	15.0%	0.807
635-669	652.0	12.1%	0.909
590-634	612.0	11.2%	1.241
530-589	559.5	9.8%	1.357
Below 530	495.0	7.5%	2.533

Exhibit 2
Bailey Analysis of Monaghan's Two-Way Study
Credit Score vs. Driving Record

	Credit Group A		Credit Group B		Credit Group C		Credit Group D			Overall		**Bailey**
Prior Driving Record	Prem	LR	Prem	LR	Prem	LR	Prem	LR	Prem	LR	LR Rel	**Factor**
No incidents	28.4	93%	66.0	71%	30.70	64%	45.80	53%	170.90	68.6%	1.000	**1.000**
1 minor	8.0	94%	17.3	68%	7.50	68%	8.40	50%	41.20	69.4%	1.012	**0.987**
1 at-fault accident	3.7	101%	7.7	74%	4.10	68%	5.90	65%	21.40	75.0%	1.094	**1.096**
1 non-fault accident	6.6	109%	14.8	81%	7.30	70%	9.90	70%	38.60	80.9%	1.180	**1.176**
2 minors	2.5	86%	6.0	59%	1.90	41%	2.40	43%	12.80	58.6%	0.855	**0.827**
2 incidents (any)	6.5	108%	13.5	96%	6.60	82%	7.90	64%	34.50	88.3%	1.287	**1.268**
All other (> 2 incid.)	18.6	114%	33.7	95%	10.80	83%	11.50	66%	74.60	93.5%	1.364	**1.289**
Overall	74.3	101%	159	79%	68.9	69%	91.8	58%				
LR Rel		1.757		1.362		1.204		1.000				
Bailey Factor		**1.709**		**1.339**		**1.192**		**1.000**				

Generalized Linear Model Details

		estimate	exp* estimate	Bailey factor	s.e.	Chi Squared	*p* -value
Credit Group	A	0.1631	1.177	1.709	0.026	40.60	0.000
	B	-0.0807	0.922	1.339	0.023	12.13	0.001
	C	-0.1970	0.821	1.192	0.031	40.37	0.000
	D	-0.3727	0.689	1.000	0.031	148.23	0.000
Prior Driving Record	No incidents	-0.2540	0.776	1.000	0.026	96.80	0.000
	1 minor	-0.2667	0.766	0.987	0.038	50.12	0.000
	1 at-fault accident	-0.1624	0.850	1.096	0.047	11.93	0.001
	1 non-fault accident	-0.0922	0.912	1.176	0.037	6.34	0.012
	2 minors	-0.4438	0.642	0.827	0.065	46.44	0.000
	2 incidents (any)	-0.0169	0.983	1.268	0.037	0.21	0.647
	All other (> 2 incid.)	0	1.000	1.289	0	--	--

* Because the log link function was used, the GLM parameter estimate must be exponentiated

Exhibit 3
Bailey Analysis of Monaghan's Two-Way Study
Credit Score vs. Driver Age

	Credit Group A		Credit Group B		Credit Group C		Credit Group D		Overall			**Bailey**
Age of Driver	Prem	LR	Prem	LR	Prem	LR	Prem	LR	Prem	LR	LR Rel	**Factor**
<25	3.8	121%	23.6	75%	1.40	51%	1.90	53%	30.70	78.2%	1.000	**1.000**
25-34	21.1	103%	55.8	79%	22.60	66%	8.90	63%	108.40	79.6%	1.018	**1.023**
35-39	13.0	100%	21.8	81%	12.90	65%	13.00	54%	60.70	75.9%	0.970	**1.007**
40-44	12.4	109%	18.5	82%	10.40	76%	15.60	52%	56.90	78.6%	1.004	**1.055**
45-49	9.8	93%	14.6	83%	8.20	76%	14.80	58%	47.40	76.1%	0.972	**1.036**
50-59	9.2	97%	14.4	78%	7.90	68%	16.50	53%	48.00	71.4%	0.913	**0.985**
60+	3.8	110%	8.3	75%	4.90	81%	20.00	67%	37.00	75.1%	0.959	**1.129**
Overall	73.1	103%	157	79%	68.3	70%	90.7	58%				
LR Rel		1.775		1.367		1.202		1.000				
Bailey Factor		**1.805**		**1.394**		**1.220**		**1.000**				

Generalized Linear Model Details

		estimate	exp* estimate	Bailey factor	s.e.	Chi Squared	*p*-value
Credit Group	A	0.1214	1.129	1.805	0.052	5.44	0.020
	B	-0.1372	0.872	1.394	0.050	7.66	0.006
	C	-0.2707	0.763	1.220	0.055	24.11	0.000
	D	-0.4692	0.626	1.000	0.048	94.19	0.000
Age	<25	-0.1214	0.886	1.000	0.067	3.30	0.069
	25-34	-0.0985	0.906	1.023	0.053	3.52	0.061
	35-39	-0.1146	0.892	1.007	0.057	4.10	0.043
	40-44	-0.0674	0.935	1.055	0.057	1.42	0.234
	45-49	-0.0865	0.917	1.036	0.059	2.15	0.142
	50-59	-0.1366	0.872	0.985	0.059	5.28	0.022
	60+	0	1.000	1.129	0	--	--

* Because the log link function was used, the GLM parameter estimate must be exponentiated

Exhibit 4
Bailey Analysis of Monaghan's Two-Way Study
Credit Score vs. Classical Underwriting Profile

	Credit Group A		Credit Group B		Credit Group C		Credit Group D		Overall			**Bailey**
Prior Driving Record	Prem	LR	Prem	LR	Prem	LR	Prem	LR	Prem	LR	LR Rel	**Factor**
MMH, Clean**	10.2	97%	22.3	77%	14.50	76%	20.20	57%	67.20	73.8%	1.000	**1.000**
MMH, Other	10.6	102%	20.2	85%	13.50	76%	16.00	58%	60.30	78.8%	1.068	**1.051**
not MMH, Clean	27.8	92%	62.9	69%	24.40	58%	34.40	50%	149.50	67.1%	0.909	**0.877**
not MMH, Other	25.6	113%	53.4	88%	16.70	74%	21.20	70%	116.90	88.2%	1.195	**1.125**
Overall	74.2	101%	158.8	79%	69.1	69%	91.8	58%				
LR Rel		1.761		1.365		1.202		1.000				
Bailey Factor		**1.739**		**1.354**		**1.196**		**1.000**				

Generalized Linear Model Details

		estimate	exp* estimate	Bailey factor	s.e.	Chi Squared	*p*-value
Credit Group	A	0.1268	1.135	1.739	0.028	20.56	0.000
	B	-0.1237	0.884	1.354	0.024	26.83	0.001
	C	-0.2479	0.780	1.196	0.035	51.46	0.000
	D	-0.4265	0.653	1.000	0.034	162.37	0.000
Prior Driving Record	MMH, Clean	-0.1175	0.889	1.000	0.035	11.02	0.000
	MMH, Other	-0.0679	0.934	1.051	0.036	3.59	0.000
	not MMH, Clean	-0.2485	0.780	0.877	0.029	75.81	0.001
	not MMH, Other	0	1.000	1.125	0.000	0.00	0.012

* Because the log link function was used, the GLM parameter estimate must be exponentiated

**MMH = Married Multicar Homeowner
Clean = Clean driving record