Check the Score

Credit Scoring and Insurance Losses: Is There a Connection?

by Bruce Kellison

*Associate Director
Bureau of Business Research*

and Patrick Brockett

*Gus Wortham Chair in Risk
Management and Insurance,
Professor
Department of Management
Science and Information
Systems, and
Director, Risk Management
Program,
University of Texas at Austin*

Over the past decade, the insurance industry has begun using credit histories to create credit scores for individuals who apply for or, sometimes, renew automobile and other insurance policies. A "credit score," typically a number between 200 and 1,000, is a measure of a person's financial responsibility. Insurers use these scores ("high" if a person's credit history is good, "low" if not) in rate-making decisions, raising premiums for individuals with poor credit history and lowering premiums for those with good credit history. Additionally, some insurers may use credit scores in underwriting procedures, including placement of policyholders within groups. Insurance companies vary widely in whether, where, and the extent to which they use credit scoring.[1]

Is there, in fact, a connection between credit history and the potential to incur insurance loss? The practice of credit scoring in insurance has fueled much public policy debate, in part because it is not immediately understood why one's credit history should be related to one's automobile insurance claims. In addition, some evidence suggests strong public agreement with the concept that an insurance applicant should be charged a premium reflective of the actual risk he or she brings to the insurance company.[2] Thus, in examining the relationship between credit scoring and loss history, researchers must first ask whether a statistically significant relationship indeed exists between the two. And, second, is the information contained in the credit score "new" information not already used in pricing the insurance?

Underwriting Variables and Insurance Classification

Insurance is, by its very nature, the substitution of group-based ex-ante premiums (a small loss known with certainty in advance) for the uncertain or fortuitous loss that may be realized by individuals within the group. The financial strength of a group allows individuals to pay a relatively small deterministic amount (a premium) to avoid the potential of catastrophic loss. A homogeneous group of exposed individuals, facing the same potential low probability but high severity event, can pool their expected losses through advance premium payments and accurately predict the average loss to be experienced by the group. Some groups of policyholders, however, will be predicted to have larger-than-average losses (e.g., young male drivers), and others, lower-than-average losses.

Therefore, to ensure homogeneity within individual groups, a number of groups are created, with the characteristics of members within each group being internally consistent. The variables used to create these groups are called classification or underwriting variables. Individuals are placed into appropriate classes to achieve the homogeneity necessary to promote rate equity. A fundamental purpose of this classification is to allow the insurer to charge an insured a premium proportional to the expected cost and the associated underwriting risk that it is perceived to bring to the insurer. What has been termed "unfair discrimination" between risks occurs when price and average expected cost are not proportional.

Insurance regulation generally recognizes that insurance is not unfairly discriminatory when insurers distinguish among risks and group them into homogeneous pools on the basis of distinct loss-cost projections—provided that this is not done on the basis of certain "protected" variables such as race, national origin, and ethnicity. The newest variable to be proposed for use in insurance classification is the credit score.

The Texas Study of Credit Scoring

In the summer of 2002, then-Lieutenant Governor Bill Ratliff asked the Bureau of Business Research (BBR), as a nonpartisan and independent research unit, to investigate whether a statistically significant relationship exists between credit score and insurance loss. To effect this assessment, the research team obtained a random sample of automobile insurance policies, including loss histories, premiums, and other variables, from several of the largest companies writing automobile insurance coverage in Texas. These policies were matched with the credit history of the named insured on the policy to create a database that included both policy information and credit information (including a summary "credit score"), but excluded information about race, ethnicity, income, or geographic location.

In this study, companies selling in the Texas automobile market were ranked according to the number of premiums written in the state. The insurers comprising the top 70 percent of the market were asked to provide a random sample of new or renewing automobile policies from the first quarter of 1998.[3] Five insurers, including those with both standard and nonstandard subsidiaries (County Mutuals), supplied data for the study, with the number of policies produced by each insurer corresponding to its market share. Included in the data obtained were the age of the named insured on the policy, the annualized premium for the policy, and the incurred losses (including actual and reserved for losses) for the policy.

The insurance companies submitted information on 175,647 separate policies to Choicepoint, a commercial firm that provides underwriting information products for the insurance industry. Prior to transmittal to the BBR, Choicepoint obtained the credit history for the policies' named insured by matching on name, address, or Social Security number, and removed all individual policyholder identifying characteristics and company marketing-sensitive information. Choicepoint created a summary credit score for each policy. A small number (.0012 percent) of clearly anomalous policies were deleted from the database. The net sample on which tests were conducted was 175,433 policies: 22,284, or 12.7 percent, had no credit scores (the "no hit group") because these policies did not have sufficient or matchable information or credit history to create a credit score, and 153,149 had credit scores matched.
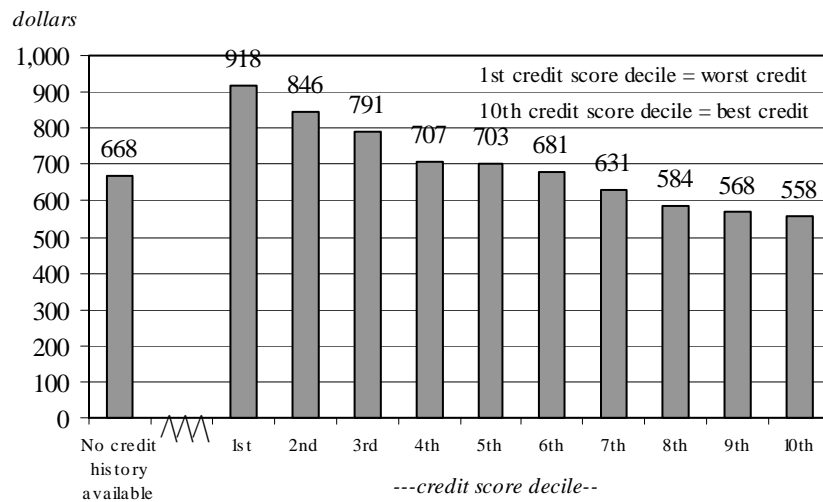
Is There a Relationship Between Incurred Losses and Credit Score?

The data obtained were broken into deciles according to the credit score. Figure 1 shows the average incurred dollar loss for each policy in each decile. Over the entire dataset, the average loss per policy was $695. For those policies in the lowest 10 percent of credit scores, this average loss was $918, whereas within the highest credit score decile, the average loss per policy was $558. Thus, the average loss per policy is higher for the lowest credit score deciles and lower for the highest credit score deciles. As the figure illustrates, the statistical correlation between incurred losses and credit score is extremely high and statistically significant.

Does Credit Scoring Add New Information?

For every dollar in premiums received, automobile insurance companies allocate a certain amount to pay claims and loss adjustment expenses. The remaining amount goes for administration costs, taxes, profit, and commissions. The ratio of incurred losses plus loss adjustment expenses to earned premiums—known as the *loss ratio*—is a frequently used measure of performance for a group of automobile insurance policies. For the companies examined here, the average company loss ratio varied from 58 percent to 74 percent, with an average of 61 percent across all companies. Because different insurers have different underwriting guidelines and different risk profiles for their businesses, the "target" loss ratio differs from insurer to insurer. The insurer sets premiums (using such underwriting criteria as age, type of automobile, coverage, deductible, territory where driven, and age and gender of driver) in such a manner as to target the insurer's desired loss ratio. If the

Figure 1
Average Incurred Losses Within Each Group for
Policies Grouped by Credit Score Decile

dollars

1st credit score decile = worst credit
10th credit score decile = best credit

No credit history available: 668
1st: 918
2nd: 846
3rd: 791
4th: 707
5th: 703
6th: 681
7th: 631
8th: 584
9th: 568
10th: 558

---credit score decile--

Note: each column represents 15,500 policyholders.

The loss ratio incorporates the multitude of underwriting variables and is an appropriate variable for assessing the statistical usefulness of a new potential underwriting variable such as credit score.

underwriting characteristics for the group indicate that an expected loss will exceed the target loss ratio, the insurer raises the premium for this group. Similarly, the premium will be lowered if the expected loss ratio falls short of the target.

Because of the random nature of individual accidents, it makes sense to measure the average loss ratio for large groups of policies and not individual policyholders. Some groups may exhibit higher accident frequencies than other groups and submit claims at a higher rate. If premiums were not adjusted upward for, say, younger drivers, the loss ratio for this group would be higher than the target ratio. However, the loss ratio for younger drivers as a group adjusts downward when their premiums are raised. Theoretically, this adjustment process continues until the insurance company achieves its target loss ratio, at which point the loss ratio for younger drivers should approximate the loss ratio for older drivers inasmuch as increased premiums have already compensated for increased losses.

If done correctly, this adjustment process makes the loss ratio for the insurer constant across all groups of drivers, with no group being charged premiums disproportionate to its anticipated losses. This is why the loss ratio is used as the benchmark for determining equity among policyholder groups. The expected loss ratio for policies within a class of policies defined by their underwrit-

ing characteristics has already, to the best ability of the insurer, accounted for the effect of existing underwriting variables on losses. Indeed, if systematic deviations from the target loss ratio appear for a given underwriting class, the premiums for this class would be adjusted to remove this bias. Any variation in loss ratio within the class should be due strictly to random or nonsystematic error. Conversely, if an analysis of a particular potential underwriting variable shows that it is significantly related to the loss ratio for the insurer, then this variable's influence on losses has not been accounted for by previous adjustments in premiums, and its inclusion as another underwriting classification variable adds value in determining premiums. Thus, the usefulness of adding an additional underwriting variable can be assessed by seeing whether the variable is significantly related to the loss ratio.

Accounting for Company Differences

As mentioned earlier, different insurers have different target markets, different risk profiles, and, consequently, different target loss ratios. The above discussion implies that for any one particular insurer, the loss ratio incorporates the multitude of underwriting variables and is an appropriate variable for assessing the statistical usefulness of a new potential underwriting variable such as credit score.

This study found that poor credit history strongly relates to insurance losses in the automobile insurance industry.

However, if one insurer or group of insurers had both a lower average credit score for its clientele and a higher average loss ratio than the automobile insurance industry as a whole, then an examination of credit scores versus loss ratios might indicate a relationship due to an insurer effect rather than to an intrinsic relationship between credit score and loss ratio. For the data examined, the credit scores for the standard market are significantly higher than the credit scores for the nonstandard market. This most likely represents the "safety valve" role that the nonstandard market insurers play in Texas, providing insurance for those unable, for one reason or another, to obtain insurance in the standard market. The way to avoid confounding the statistical analysis due to the possible correlation of company loss ratios with both incurred losses and credit scores is to use a *relative loss ratio* for each policy, where relative loss ratio is defined as the loss ratio for the policy divided by the average loss ratio for the insurer issuing the policy. This avoids potentially spurious findings due solely to insurer differences.

Figure 2 illustrates the findings. For the named insureds in the lowest 10 percent of the credit scores, the relative loss ratio for their policies averaged 53 percent higher than expected, whereas for the named insureds in the highest 10 percent of the credit scores, the relative loss ratio averaged 25 percent lower than expected (where by construction, a relative loss ratio of one

was the expected loss ratio over all policies without using credit scoring). The no credit history group has an average loss ratio of 1.07, or 7 percent higher than expected.

Statistical analyses confirmed the visual relationship apparent in figure 2. A regression analysis of the relative loss ratio on credit score was highly significant (p<.0001). This indicates that there is less than a 1 in 10,000 chance that the relationship observed between credit score and relative loss ratio could be due to chance alone.[4]

Findings, Limitations, and Implications

This study found that poor credit history strongly relates to insurance losses in the automobile insurance industry. Insurers' actuaries can use this relationship to better predict an individual's loss costs, to differentiate more effectively between classes of insureds, and to price their firms' policies more commensurate with the risks that the policyholder brings to the insurer. This will result in fewer cross-subsidies from one class or group of policyholders to another, reinforcing the goal of equitable rates among various groups of policyholders.

However, this study was not designed to, nor does it, answer a number of important public policy questions. Certain critics argue that credit information collected by the three main credit bureaus (TransUnion, Experian, and Equifax) can contain inaccurate information on consumers and their credit histories. Such inaccuracies would then compromise any subsequent credit

Figure 2
Average Relative Loss Ratios Versus Credit Scores
for Total Market Dataset
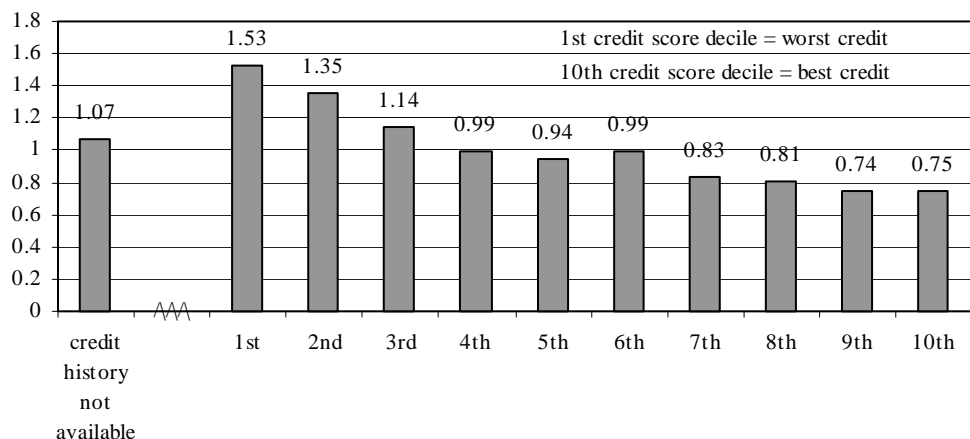
score created by third-party commercial firms like Choicepoint for use in the insurance industry, not to mention credit scores created by insurance companies themselves. In the present study, if the credit information contains inaccuracies, and if the inaccurate variables happen to be among the subset of variables used to create the credit score, then the credit scores generated will be inaccurate as well. Although it was beyond the scope of this study to examine the accuracy of the credit report supplied, it should be noted that random and occasional errors in credit data would not significantly weaken the statistical correlation found in the results.

An important proviso regarding inferences that can be drawn from this study relates to the credit score itself. The analysis in the study used the credit score created by Choicepoint. Individual insurance companies can (and do) use individual credit histories to create their own models and credit scores. If an individual insurance company can create a "better" (more predictive) credit score model, the relationship between credit scoring and losses will be even stronger than that found in this study. Conversely, if an insurance company produces a less predictive credit scoring model, the relationship found here may be stronger than that observed by such an insurer. Thus, results presented here should be viewed as illustrative of the relationship that can be determined between credit scoring and losses.

As is the general practice in the insurance industry, the credit score generated by Choicepoint was based on a credit match with the identifying characteristics of the named insured. For multiple driver policies, each driver might have a different credit score and different incurred losses. Consequently, a named insured (a father, for example) could have a very good credit history, but the young son driving on the policy could have many incurred losses. In such a case, a "good" credit score would be associated with a policy having high incurred losses. In this regard, the current study should be interpreted as showing a significant relationship between the credit score of the named insured and losses for everyone on the policies and not as showing a relationship between the credit score of an individual driver and the losses of that particular driver.

A common criticism of credit scoring and its use in underwriting decisions is that it may discriminate against low-income and/or minority applicants, and that its use, in effect, amounts to "red lining." Some within the insurance industry maintain that their underwriting and rate-making practices are blind with regard to ethnicity and income. The database used in this study did not contain information on named insured income, ethnicity, or physical address (other than rather gross delineation of rating territory for some but not all insurers), so the results of this study cannot and do not address this issue.

Notes

1. The credit score described here is tailored to insurance losses and is often called an insurance credit score. It is distinct from the use of credit history by, for example, banks to assess whether or not to approve a loan.

2. In response to the statement "people should pay different rates for car insurance based upon the degree of risk they represent to the insurance company," 1,000 persons polled had an average score of 3.36 on a one-to-four scale, with one denoting strong disagreement and four representing strong agreement (Gallup March 1991). See *Best Review* (Property/Casualty Ed.), vol. 91, no. 11 (March, 1991), p. 12.

3. This examination period was chosen because at this time most of the insurers were not using credit scoring in rate-making decisions, so the premium data collected were not confounded by credit history. Also, more complete policy loss information, including paid losses and reserves for losses, could be obtained (it takes time for claims to pay out).

4. Breaking the loss into frequency of loss and severity of loss, two additional analyses were performed. A logistic regression analysis was conducted to determine whether the existence of a positive claim (incurred loss greater than zero) was significantly related to credit score. Each policy was classified as to whether a positive loss or no loss was experienced. This was then related to credit score using logistic regression. It was found that there was a statistically significant relationship between credit score and the likelihood of a positive claim being filed ($p<.0001$). Another analysis was performed to ascertain if the size of the claim was related to credit score. For this, a regression of the relative loss ratio on credit score was performed using only those policies having a strictly positive relative loss ratio. Again, credit score was significant ($p<.0001$), indicating that the size of the loss is also significantly related to credit score. Finally, using the data grouped by credit score deciles exhibited in figure 2, the correlation between credit score and relative loss ratio was calculated. The correlation $|r|$ was .95, which is statistically and substantively significant. Thus, both the likelihood of a claim, and the size of the claim should it occur, are significantly related to credit score.◆

Random and occasional errors in credit data would not significantly weaken the statistical correlation found in the results of this study.

BBR

BUREAU OF BUSINESS RESEARCH
The University of Texas at Austin
1 University Station B8500
Austin, Texas 78712

Change Service Requested

Glossary of Terms
for the Bureau of Business Research
Credit Score Study

Credit score--a measure, usually a number between 200 and 1,000, of a person's financial responsibility based on different variables in a person's credit report and credit history. It is developed to relate to insurance losses and is distinct from the use of credit history for loans.

Incurred loss--a loss, in dollars, paid by an insurance company on a claim by an insured on an insurance policy. Incurred losses in the study included actual losses paid and reserves for losses not yet paid.

Insured--the person or policyholder covered by an insurance policy.

Loss ratio--the sum of actual paid losses, loss expenses, and loss reserves divided by the premiums paid by the insured for a given period of insurance coverage.

Nonstandard market company--an insurance company in Texas (in the automobile insurance industry these are County Mutuals) whose rates are not regulated by the state and who usually insures higher-risk drivers than do standard market companies.

Standard market company--an insurance company in Texas whose rates are regulated by the state.

Announcement

The 2003 edition of **Texas Trade and Professional Associations** will be released in May 2003 by the Bureau of Business Research. Since 1951, this directory has provided information on associations with headquarters in Texas. The new edition lists association name, address, phone and fax numbers, web address, executive officer, publication titles, and number of members for more than 900 associations. Listings are alphabetical by name, and the directory also includes city and subject indexes. The database is also available on CD-ROM or diskette. For ordering information, contact Rita Wright at 888-212-4386 or email at Rita.Wright@bus.utexas.edu.◆