The Impact of Credit-Based Insurance Scoring on the Availability and Affordability of Insurance

Lawrence S. Powell, Ph.D.
Whitbeck-Beyer Chair of Insurance and Financial Services
University of Arkansas-Little Rock
326 Reynolds Center
2801 S. University Avenue
Little Rock, AR 72204

lspowell@ualr.edu
501.773.7577

Research Fellow
The Independent Institute
1319 18th St. NW
Washington, D.C. 20036

Introduction:[1]

Insurance companies face an unusual challenge. They must set prices for the products they sell before they know all of the costs. To meet this challenge, they employ complex pricing methods developed by actuaries using applied economic and statistical techniques. It should then come as no surprise that some aspects of actuarial science and insurance pricing are puzzling to people who have not developed substantial expertise in this field.

Insurance scoring, the use of credit information in insurance underwriting and pricing, is an example of a beneficial practice that is sometimes misunderstood. Insurance scoring benefits consumers in several ways, all of which stem from its accuracy as a predictor of insured losses.

The purpose of my testimony is to present comprehensive information about insurance scoring in a non-technical format. In Section 1, I present a brief conceptual summary of insurance pricing and insurance scoring. In Section 2, drawing from existing studies, I present evidence that insurance scores are powerful and accurate predictors of insurance losses. In Section 3, I conclude with discussion of the appropriateness of insurance scoring.

Section 1: Insurance Pricing and Insurance Scoring

An insurance company facilitates risk pooling, reducing the uncertainty of individual pool members. Uncertainty decreases because the ultimate value of the group's losses is more predictable than that of an individual. Swiss mathematician Jacob Bernoulli first proved this phenomenon, known as the law of large numbers, around 1690. Relying on the law of large numbers, a group of pool participants can each pay the average or expected loss of the group, rather than paying for a much less predictable and potentially larger individual loss on one's own.

Risk pooling is most effective when all members of the pool have the same expected loss. Insurance companies rely on risk classification systems to ensure that groups of insureds pay premiums commensurate with their exposures to risk. When insurers pool exposures with unequal expected losses, the low-risk group must subsidize

---

[1] Much of this testimony is drawn from a study I am writing for the Independent Institute.

the high-risk group. This creates an incentive for low-risk pool members to purchase less insurance than high-risk pool members, a scenario called adverse selection. Adverse selection can break down the risk pooling mechanism and, in extreme cases, lead to insolvency of the pool.

Insurance companies use information about applicants for insurance to classify them into groups with very similar expected loss. Of course, no risk classification system is perfect. In addition to other restrictions, insurers can only use rating information if it is cost effective, meaning the cost of obtaining the information is less than the difference in expected loss between groups. For example, assume there are only two types of drivers, low-risk and high-risk. The low-risk group has expected loss of $500 and the high-risk group has expected loss of $700. If it costs more than $100 to classify a driver, it will be more cost effective to simply pool them together and charge both groups $600. However, if an insurer can identify low-risk drivers for, say, $20, it benefits the low-risk drivers to charge them $520, and charge the high-risk drivers $720. On the other hand, insurers could be more precise in risk classification if they hired private investigators to follow each driver for six months before offering an insurance policy. Obviously, this would cost more than $100, and raise privacy concerns. To have enough money in the risk pool to cover expected losses, low-risk drivers would have to pay more than $600. In this case, there is no justification for such an unfair classification.

There are many variables insurers use to classify drivers based on expected loss. These include, but are not limited, to geographic location, age, gender, marital status, miles driven, type of vehicle, use of vehicle, driving record, and insurance score. An insurance score is a numerical prediction of propensity for loss estimated using certain information from a driver's credit history. The actuarial literature shows it is one of the most accurate and cost effective loss predictors available (EPIC, 2003).

There are several apparent misconceptions about insurance scores. To understand why insurance scores are beneficial to insurance systems, it is important to start with an accurate description that is free of incorrect assumptions. The variables commonly used to estimate insurance scores include measures of performance on credit obligations, credit-seeking behavior, use of credit, length of credit history, and types of credit used

(FTC, 2007). They do not include income, wealth, race, ethnicity, or any prohibited factor.

Insurance scores and credit scores are calculated using some of the same information, but they are not equivalent. The important difference is that credit scores use these variables (and others) to estimate the probability of a borrower defaulting on a financial obligation, while insurance scores estimate the probability of having insured losses.

An important fact often overlooked in the debate about insurance scoring is that the only way including insurance scores in an insurance rating model can result in higher premiums is for the sample population with lower scores to have more insured losses. As I describe in more detail in Section 3, any deviation from using the most accurate, cost effective predictors results in unfair outcomes and damage to the insurance mechanism.

One observed barrier to understanding insurance scoring is manifest in the common criticism that there is not an intuitive link between insurance scores and driving ability. While several studies develop potential causal links between insurance scores and driving, I find it more compelling to recognize an alternative relation. The use of insurance scores does not rely on a link between credit information and "driving ability." Rather, it is a link between insurance scores and insured losses.

There are many factors unrelated to driving ability that increase the likelihood of insured losses. For example, someone who always makes debt payments on time to avoid higher interest rates the next time they borrow may also choose not to file a small insurance claim to prevent an increase in insurance premiums in the future. It may also be the case that insurance scores measure hazards other than lack of driving ability.

Section 2: Predictive Accuracy of Insurance Scores

The correlation between driving outcomes and credit information appears in academic literature as early as 1949 (Tillman and Hobbs, 1949). Over time, evidence of the empirical relation between automobile insurance losses and insurance scores has developed to address not only the simple correlation between insurance costs and insurance scores, but also the additional predictive power and accuracy insurance scores contribute to insurance pricing models containing traditional pricing variables.

In this section, I review methods and results from several studies investigating the relation between insurance scores and insurance losses. The findings consistently and conclusively demonstrate that insurance scores are highly correlated with losses. The studies also show that insurance scores supply information about insurance losses not contained in other underwriting and rating variables.
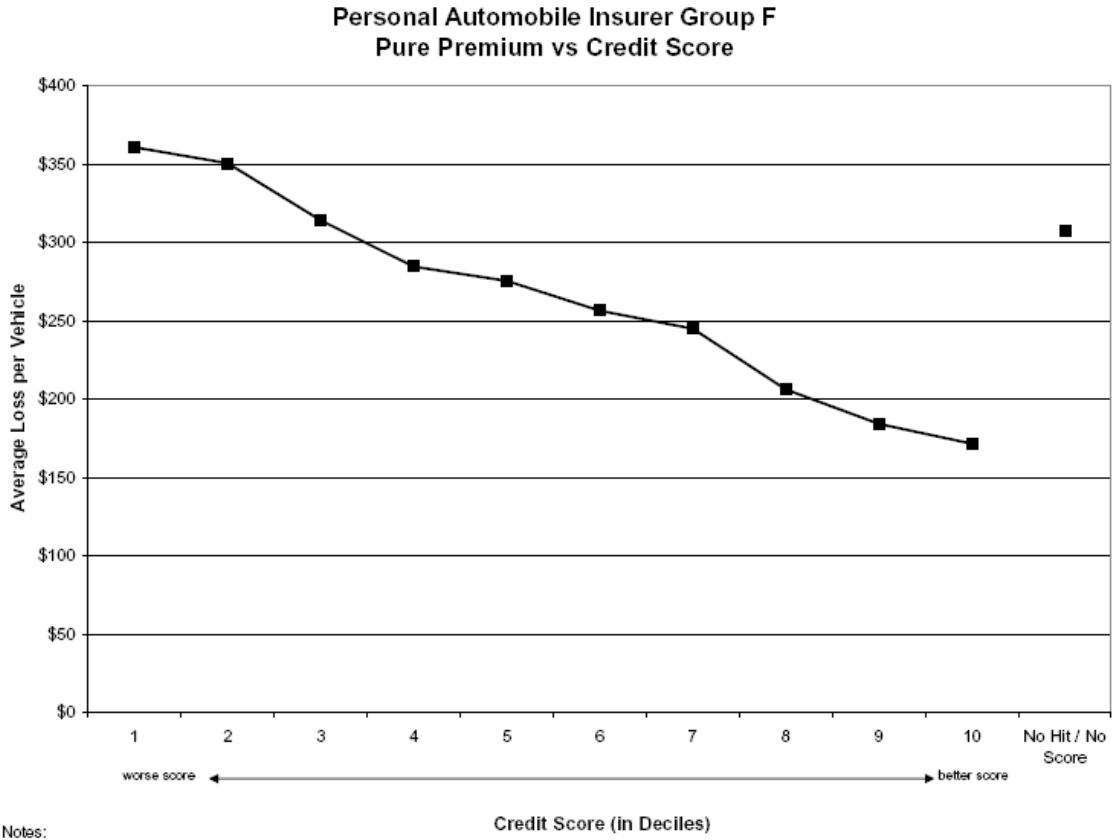
More than a dozen studies related to insurance scoring have appeared in the public domain in the last decade. To improve the exposition of information, I present evidence from various studies in order of increasing complexity. This does not exactly match the temporal order in which studies were released. Furthermore, many of studies produce very similar evidence and reach nearly identical conclusions. I make an effort to report from the most recent and clear studies.

The most basic result is the simple correlation between insurance scores and losses. A study conducted by the Texas Department of Insurance in 2004 (TDI, 2004), firmly establishes the simple correlation between insurance scores and losses. Using data representing approximately 2 million insurance policies, the authors group exposure units by deciles of credit scores and graph the coinciding average loss frequency and loss amount.

Figures 1 and 2 appear in TDI (2004) as Charts 7 and 9, respectively. Figure 1 shows that average loss per vehicle declines steadily across deciles of credit score. Those with the lowest scores average approximately $360 per vehicle, while those with the highest scores average approximately $175 per vehicle. Similarly, Figure 2 shows number of claims per 1,000 exposures decreasing from approximately 110 for those with the lowest scores to just over 60 for those with the highest scores. These results are qualitatively similar across all of the companies reporting automobile insurance data for the study.
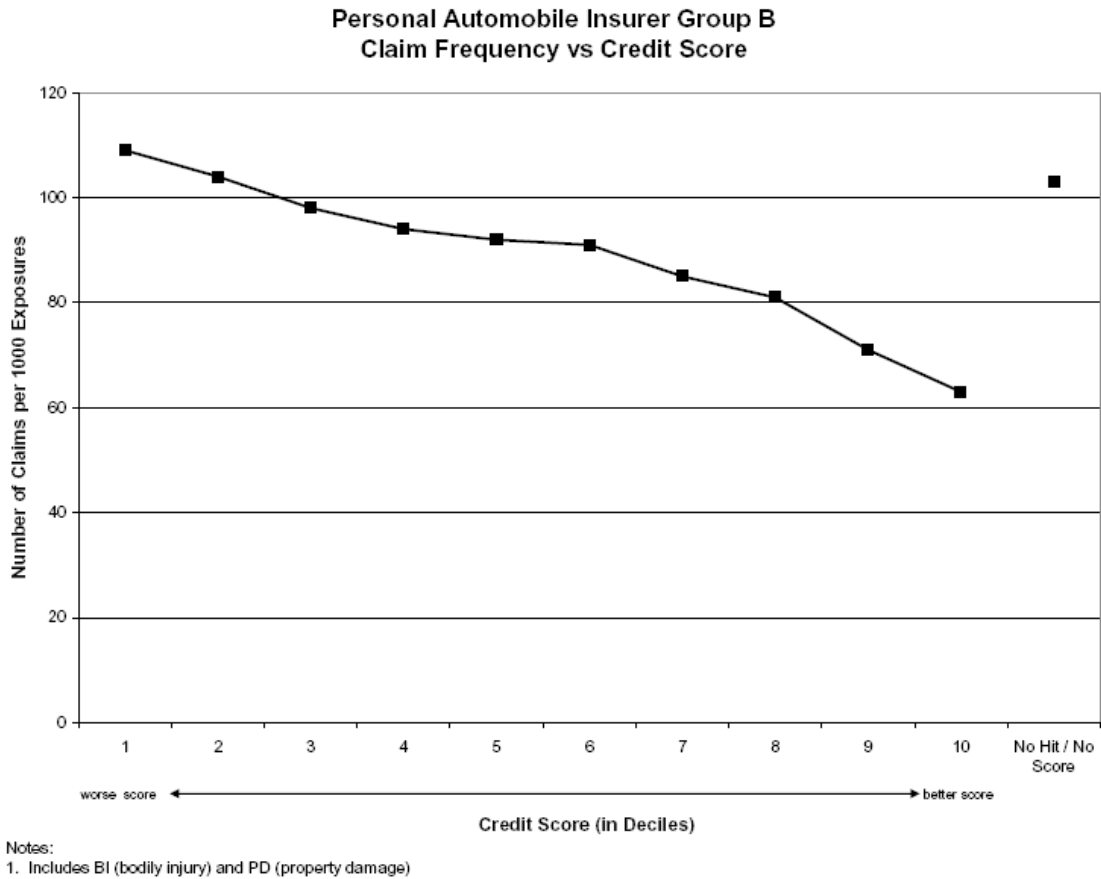
Several other studies reach similar conclusions using data from nationally representative samples (EPIC, 2003 and FTC, 2007), rather than the single state sample used by TDI.

Figure1: Credit Score and Average Loss per Vehicle

Personal Automobile Insurer Group F
Pure Premium vs Credit Score

Notes:
1. Includes BI (bodily injury) and PD (property damage)
2. Losses are capped at basic limits ($20,000/$40,000/$15,000)

Source: TDI (2004)

Figure2: Credit Score and Number of Claims per 1000 vehicles

Personal Automobile Insurer Group B
Claim Frequency vs Credit Score

Notes:
1. Includes BI (bodily injury) and PD (property damage)

Source: TDI (2004)

Critics of TDI (2004), including the Texas Department of Insurance itself, point out that simple correlation between a rating variable and losses is neither necessary nor sufficient to establish its validity as a predictor of losses. This is true because no variable can produce a more accurate prediction of losses alone than when combined with other accurate predictors of losses. Therefore, in addition to simple linear correlation between predictors and losses, one must also consider the interactions among a group of predictor variables. To do so requires multivariate analysis.

Multivariate analysis, as the name implies, involves analysis of two or more predictor variables at the same time. EPIC (2003), FCT (2007) and a second study by the Texas Department of Insurance (TDI, 2005) employ multivariate analysis to determine if

insurance scores are risk related. I summarize the analysis and primary findings of these studies below.

TDI (2005) examines a large database of personal automobile and homeowners insurance policies in Texas. The authors performed multivariate analysis considering the interaction of insurance scores and several other common predictors of insurance losses. They find that the strong correlation between insurance scores and losses persists even when controlling for other underwriting factors. TDI (2005) concluded that, "credit scoring 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." The authors also find that "use [of insurance scoring] is justified actuarially and it adds value to the insurance transaction."

EPIC (2003) examines a nationally representative sample of insurance scores, underwriting data, and policy outcomes (losses). The study produces four primary findings: First, insurance scores are correlated with risk of loss, even after controlling for relationships with other variables. The correlation is due primarily to loss frequency rather than loss severity. Second, insurance scores are correlated with some other common risk factors; however, even after controlling for other factors, insurance scores significantly increase the accuracy of the risk assessment process. Third, insurance scores are very powerful predictors of loss relative to other common risk factors. Finally, results from the study apply generally to all states and regions.

FTC (2007) also examines a large, nationally representative database to determine the relation between insurance scores and losses. The study finds that "even when non-credit variables are included in the analysis, credit-based insurance scores continue to predict the amount that insurance companies are likely to pay out in claims to consumers." More specifically, they find insurance scores are effective predictors of risk under automobile policies. They are predictive of the number of claims consumers file and the total cost of those claims. The use of scores is therefore likely to make the price of insurance better match the risk of loss posed by the consumer. Thus, on average, higher-risk consumers will pay higher premiums and lower-risk consumers will pay lower premiums.

These recent studies envelop a spectrum of backgrounds and data sources. Private groups and government agencies conduct them. They represent a single state and national samples. They employ different measures and methodologies. Nonetheless, they all reach the same general conclusion: that insurance scores are highly predictive of losses, even when controlling for other factors. As noted at the outset, insurers are unique in the U.S. economy as they do not know the ultimate cost of their product when they sell it so having a tool to more effectively predict losses helps insurers more fairly, for all consumers' benefit, price their products.

Section 3: Appropriateness of Insurance Scores

Regulators require insurance rates to meet three criteria. They must not be inadequate, excessive, or unfairly discriminatory.[2] A rating criteria is unfairly discriminatory if it is does not bear a reasonable relationship to the expected loss and expense experience among insured exposures. Given the evidence presented in Section 2, insurance scores clearly meet the third criterion. However, some people remain uncomfortable with application of credit information in insurance rating. In this section, I describe the individual and societal benefits of insurance scoring. Finally, I present evidence that competition in insurance markets prevents discrimination based on any factor other than expected losses.

Insurance scoring benefits society in several ways. All of the benefits accrue from improved efficiency and accuracy of risk estimates. The first benefit is that insurance scores provide a very high level of accuracy for a relatively small cost. Using insurance scores reduces cost for insurance companies. Because the market for insurance is competitive, this savings is passed through to consumers as lower premiums. Data from a recent report by the Arkansas Insurance Department indicates that if insurance scoring were eliminated as a rating factor, nearly 91 percent of automobile and homeowners insurance consumers would incur a rate increase. Using a slightly different method, the FTC (2007) study estimates that insurance scoring results in a decrease in insurance premiums for 59% of drivers.

---

[2] Almost every state also imposes additional restrictions on the use of insurance scores in the ratemaking process.

The next benefit of insurance scoring is that improved accuracy may make insurers more willing to offer insurance to high-risk consumers for whom they would otherwise not be able to determine an appropriate premium (FTC, 2007). For example, insurance scoring information can allow an insurer to offer coverage to drivers living in a geographic area with high traffic density at a price the driver can afford. Without information from insurance scores, insurers would not be able to differentiate sufficiently among these drivers. Therefore, they would not be able to offer the coverage at a lower price for the lower-risk drivers living in the area. Consistent with this assertion, FTC (2007) finds limited evidence that the advent of credit scoring in automobile insurance coincided with substantial decreases in residual market mechanisms. This suggests insurers, with the benefit of credit information, are more willing to offer coverage to high-risk drivers (at a risk-based price) than they were before the introduction of insurance scores.

Another advantage of using insurance scores is it improves accuracy of information used to classify drivers. In addition to calculating more accurate loss predictions, the scores, themselves, are less likely to contain material factual errors than are several of the driving history variables used to underwrite insurance. Studies by Associated Credit Bureaus (ACB, 1992) and Trans Union report material errors in credit information in only 0.2% of credit records. In striking contrast, a study by the Insurance Research Council (IRC, 1991) found public information available on only 40% of a sample of known automobile losses. Underreporting of traffic citations also appears problematic. IRC (1991) indicates less than a third of all traffic citations are accurately reported in state driving records. Furthermore, consumers have a strong incentive to correct inaccurate credit information; whereas the opposite incentive exists for driving records. This is true because recorded driving events can only be adverse events. Data describing instances in which drivers avoid collision by defensive driving and alertness are not collected.

The final benefit of insurance scoring I would like to address is that, because scoring produces more accurate loss estimates, it results in outcomes that are more equitable for individuals and society as a whole. As noted in Section 2, insurance scoring is likely to make the price of insurance better match the risk of loss posed by the

consumer. Thus, on average, higher-risk consumers will pay higher premiums and lower-risk consumers will pay lower premiums (FTC, 2007). This addresses a very common problem in the insurance mechanism called cross-subsidization.

When insurers cannot accurately classify applicants for insurance, they must either decline applications, or charge the same premium to high-risk and low-risk drivers. The latter case obviously leads to cross-subsidization – when low-risk drivers must over-pay to make up for underpaying high-risk drivers. However, the former case, declining applications for insurance, ultimately leads to the same outcome. This is type of cross-subsidization is facilitated by residual markets for insurance.

Each state has a residual market mechanism to make insurance available to drivers whom the voluntary market will not cover. Residual market mechanisms effectively set a maximum price that insurers may charge for insurance. If insurers are not willing to offer coverage at this price, consumers may purchase coverage at this price from the residual market. However, if the premium is not enough to cover losses and expenses, insurers in the voluntary market must make up the deficit in proportion to their market shares.

FTC (2007) shows that as insurance scoring has become more common in ratemaking models, the populations of states' residual markets have decreased. This suggests insurance scoring results in more equitable or fair outcomes compared to less accurate rating models that do not use insurance scores.

Perhaps the most controversial result appearing in FTC (2007) is the study's assertion that insurance scores exhibit a proxy effect for race. Objective consideration of this result leads me to doubt its validity. The econometric test used to support the existence of a proxy effect is flawed such that it would not withstand the scrutiny of a legitimate academic peer-review process. Clearly, the lack of objective confidence in the result suggests that public policy should not be altered to address this weak finding.

Another way to address the appropriateness of insurance scoring is to consider the level of competition occurring in insurance markets. If insurance markets are competitive, insurers will not be able to charge excessive or unfair prices. If an insurer tries to set prices based on anything other than expected losses and costs, it will either,

suffer substantial losses if the price is too low, or, if the price is too high, it will lose market share as its competitors offer a lower price to the same consumers.

Effective competition is a fundamental characteristic observed in U.S. insurance markets. Competition prevents insurers from charging excessive or unfair prices. In 2005, NAIC data show an average of 157 insurance companies underwriting the private passenger automobile cover in each state. It is, therefore, reasonable to believe that an insurer cannot systematically over-charge a group of drivers because one of the other 156 existing companies, or perhaps a new company, has an opportunity to cover that group of drivers at an equilibrium price. Compare such competition to other "required" services such as phone, gas, electric, etc. where consumers have at best the choice between two companies.

We are not in this hearing because everyone likes insurance scoring. I have heard critics of insurance scoring describe potential or anecdotal unfair outcomes associated with its use. I do not dispute the fact that some consumers have encountered individual rating scenarios that seem to lack intuition. For example, I know of a consumer in Arkansas who received an increase in his premium because his wife cancelled a credit card they were not using. However, he called a few competing insurance companies and found one that offered him the same coverage at a significant discount from what he was paying before the change in his credit. This is an example of competitive markets reaching an optimal outcome.

While competitive markets are very effective at making the goods and services consumers want available to them, critics have voiced concerns that when a drop in credit is unrelated to insurance risk some individuals could be mistreated by insurance scoring. In response to such concerns, almost every state has regulations in place to recognize the benefits of insurance scoring, while limiting its use in certain scenarios. I think it is worth noting that many insurers offered the same protections as these regulations require before the laws were enacted. This is another example of competitive markets creating an optimal outcome.

Conclusion

Setting reasonably accurate prices for insurance is a difficult task because insurers must establish prices without the benefit of knowing all of the costs involved. To offset this hardship, actuaries have developed complex pricing models using applied economic and statistical tools. While this complexity is necessary, it unfortunately leads to a lack of understanding among people who have not developed such specific expertise.

Insurance scoring is an example of a beneficial tool used in ratemaking that is often misunderstood. Insurance scores are relatively powerful and accurate predictors of losses, even when controlling for other factors known to be correlated with losses. When insurers use insurance scores to improve the accuracy of predicted losses, it benefits individuals and society. It increases the equity or fairness in insurance pricing outcomes because, on average, premiums are closely related to consumers' risk of loss. Insurance scoring also adds value to insurance transactions. It reduces the overall cost of providing insurance because insurance scores are accurate and inexpensive rating variables.

Finally, the vigorous competition exhibited by the property and casualty insurance industry suggests that pricing of insurance based on anything other than expected losses is nearly impossible. Insurance markets show strong signs of effective competition including a large number of suppliers and low barriers to entry.