American Insurance Association

NAMIC®
NATIONAL ASSOCIATION OF MUTUAL INSURANCE COMPANIES

PCI **Property Casualty Insurers Association of America**
Shaping the Future of American Insurance

August 16, 2004

Federal Trade Commission
Office of the Secretary
Room H-159 (Annex N)
600 Pennsylvania Avenue, N.W.
Washington, D.C. 20580

RE: FACT Act Scores Study, Matter No. P044804

Dear Sir or Madam:

The following comments are submitted by the undersigned on behalf of the following property-casualty insurance industry trade associations (collectively "Insurance Trades")[1] and their member companies: American Insurance Association ("AIA"), National Association of Mutual Insurance Companies ("NAMIC"), and Property Casualty Insurers Association of America ("PCI"). They respond to a notice and request for public comment regarding methodology and research design for conducting a study of the effects of credit scores and credit-based insurance scores which was mandated under Section 215 of the Fair and Accurate Credit Transactions Act of 2003 ("the Study").

The Insurance Trades urge the Federal Trade Commission ("Commission") and the Board of Governors of the Federal Reserve System ("Board") to proceed cautiously with the Study and to design it in a way that will ensure that it is as accurate as possible. This responsible approach is necessary given that the issues surrounding the impact of credit information could be politically charged and socially sensitive. Ideally, the Study would use correct demographic information for each individual policyholder, in order to give maximum integrity to the results. Since it appears that the Commission and the Board will not likely be able to identify many of the ECOA-protected classes on an individual basis with aggregate data, the Study risks being oversimplified and inaccurate. This underscores the need at the outset to understand the methodology to be used and the standards to be applied.

These comments respond to the specific questions posed with respect to credit-based insurance scores used by property and casualty insurers.

(1) How should the effects of credit scores and credit based insurance scores on the price and availability of mortgages, auto loans, other credit products, and property and casualty insurance be studied? What is a reasonable methodology for measuring the price and availability of mortgages, auto loans, credit cards, other credit products, and property and casualty

---

[1] Together the Insurance Trades represent more than 97% of property-casualty insurance premiums written in the United States.

insurance, and the impact of credit scores and credit based insurance scores on those prices and availability?

The best approach for statistically measuring the impact of credit-based insurance scores on the price of insurance might be to develop a generalized linear model incorporating known risk factors. This is discussed in greater detail in Question 3.

Price – A properly structured sample of individual policies, including risk factors and premium or loss experience for each policy, can be used to calculate two indicated premiums or pure premiums.[2] The first calculation would be the indicated pure premium before the use of the credit-based insurance score and the second calculation would be the indicated pure premium after the introduction of the insurance score. The difference between the actuarially indicated pure premiums (i.e., with and without credit-based insurance scores used as a risk factor) is the effect on the pure premium of introducing credit-based insurance scores. The percentage of insureds that have higher or lower pure premiums due to the use of credit-based insurance scores can be quantified by comparing the two indicated pure premiums for each policy record in the Study's database.

The effect on the pure premium will need to be "tempered" to determine the indicated impact on price. Not all the costs included in an insurance premium are variable with losses (i.e., pure premiums). The impact on price, or insurance premiums, will be somewhat less than the impact on the pure premiums because of these non-variable costs. The non-variable cost component of insurance premiums can be estimated from publicly available insurance expense data, such as the expense exhibits from the annual statements.

Availability – One indicator of availability of coverage is the change over time in proportion of the entire market which is being served through a "residual market". For example, an increase or decrease in the market share of the auto assigned risk plans, or an increase or decrease in the property insurance FAIR plans, is an indicator of the degree of the insurance industry's ability to provide insurance at

---

[2] There are two plausible choices for the model's dependent variable. The first is the pure premium, or loss costs, which models the loss cost as a function of the independent variables. Then, some simplifying assumptions are made about how the indicated loss cost ultimately becomes a rate. The second approach models price directly as a function of the independent variables.

One benefit of the loss cost approach is data availability. Also, modeling loss cost (via frequency and severity) is a well accepted actuarial technique. Should pure premium be used, it is with the caveat that it is not perfect since the ultimate price may also have been affected by other items such as business concerns and regulatory environments. On the other hand, the strength of a premium approach is that it may more directly answer the question of how credit impacts the price of insurance. The premium and pure premium approaches may differ as to whether the risk is systemic or non-systemic. The premium approach may allow the researcher to control for how and the degrees to which companies employ credit in their rating algorithms.

For the ease of reading, only the modeling of pure premium has been included in the body of these comments. At this time, the Insurance Trades have not reached a consensus on the optimal approach and we would like an opportunity to provide additional input on this issue.

an adequate price. While there are a multitude of factors which could be simultaneously impacting the size of the residual markets, the use of credit-based insurance scores to more accurately assess risk is likely one such factor.

Coverage availability is primarily a function of adequate rates. If insurers are allowed to accurately price for the risk, then even those with poor insurance scores will have greater access to insurance coverage.

Besides relying on the changes in size of the residual markets, a study of availability of coverage should also rely on evidence solicited directly from insurers. An insurer that has introduced credit-based insurance scores into its rating or underwriting program may be able to supply information showing how insurance scores enabled it to serve a broader insurance market.

One insurer reports that without the use of credit-based insurance scoring, it would not have expanded its market reach. In fact, it reports that it now offers quotes to 96% of all applicants. Prior to the use of scoring, when applications were evaluated one on one, it restricted its writing to the preferred auto insurance market and therefore would reach far fewer applicants (it estimates that earlier quote rate nearer 65 to 70%).

While some may suggest that changes in the uninsured motorist figures may be a useful indicator of availability, as a preliminary matter we anticipate problems with this approach - not only because the data are generally unreliable and without a good source, but also because the underinsured motorist[3] numbers are often not separated from the uninsured numbers.

Other – One possible reading of this question is whether an inquiry by an insurer into an individual's consumer report will be counted against the consumer either by another insurer or by an institution that issues credit. It is our understanding that Fair Isaac represents that insurance inquiries are not counted in their scoring models used by mortgage companies or by retail credit evaluators. Furthermore, many states prohibit an insurer from counting other insurers' inquiries against a consumer.

(2) An effect can often only be measured relative to a counterfactual (that is, relative to some hypothetical alternative situation). To determine the effects of credit scores on the price and availability of credit products, what is a reasonable counterfactual to the current use of credit scores? To determine the effects of credit-based insurance scores on the price and availability of property and casualty insurance, what is a reasonable counterfactual to the current use of credit-based insurance scores?

When measuring the impact on price, it seems an appropriate counterfactual to using credit-based insurance scores may be the indicated pure premium prior to

---

[3] Underinsured motorist coverage provides additional compensation for insured motorists involved in accidents when the other at fault driver involved in the accident has inadequate policy limits to cover accident or injury costs.

the introduction of insurance score as a rate factor. The difference between the actuarially indicated pure premiums with and without credit-based insurance scores used as a risk factor, is the impact on the pure premium of credit-based insurance scores. These data can be used to calculate the percentage of insureds that have higher or lower pure premiums, and the distribution of insureds by the size of the premium change.

The counterfactual for measuring the impact on coverage availability will be the residual market populations and the insurers' underwriting practices just prior to using credit-based insurance scores as part of the rating/underwriting process.

(3) Paragraph (a)(2) of Section 215 requires a study of "the statistical relationship, utilizing a multivariate analysis that controls for prohibited factors under the (ECOA) and other known risk factors, between credit scores and credit-based insurance scores and the quantifiable risks and actual losses experienced by business." (The ECOA "prohibited factors" are race, color, religion, national origin, sex or marital status, and age.) What is an appropriate multivariate technique for studying this relationship? What data would be required to under take such an analysis? What data are available to undertake such an analysis?

Multivariate analysis generally refers to the simultaneous analysis of two or more factors. To properly control for the overlaps that exist between risk factors, it is important that the Study be conducted so that risk factors be simultaneously analyzed. We suggest the first run of the model be done with all the factors, except credit-based insurance scores. Credit-based insurance scores would be added to the second run of the model.

The model should be a multivariate regression model with loss costs as the dependent variable (i.e., pure premium).[4] The independent variables would be credit-based insurance scores, other known risk factors, and the ECOA prohibited factors.

Insurance-specific generalized linear model (GLM) software products – such as EMB's Emblem or Watson Wyatt's Pretium – are available that can efficiently handle the multivariate analysis. However, other statistical packages, such as SAS, could also be used to perform the statistical calculations. The SAS/Stat module has a procedure, "Proc Genmod," that can perform Generalized Linear Modeling (as opposed to General Linear Models). It is commonly accepted that the best statistical models for insurance data use a Poisson error structure for claim frequencies and a Gamma error structure for average claim severities. Further, the use of a "log-link" function generally provides a better fit than other approaches. The modeled pure premiums are a combination of the modeled claim frequencies and claim severities.

Data required to conduct the multivariate analysis include individual policy records containing known risk factors and actual loss experience; a common

---

[4] See discussion of price under Question 1

credit-based insurance score which can be attached to each policy record; and information concerning the ECOA-prohibited factors.

Information concerning race, color, religion, natural origin, and income are not contained in the insurer's policy records. Therefore, if studied, these data would need to be obtained from publicly available data files and be attached to each individual policyholder record in a way that is accurate for each specific policyholder.

We recommend the use of a common insurance score. Significant differences among companies in the construction of insurance scores make it impossible to conduct a study based on aggregated data from multiple insurers, unless a common insurance score is used.

(4) What is an appropriate methodology to determine whether the use of credit scores or credit based insurance scores results in "negative or differential treatment" of ECOA-protected classes?

Importantly, a working definition of "negative or differential treatment" needs to be adopted before the Study is conducted. Without such a definition, the arguments over the Study's conclusions are likely to be endless.

An effective measure of "differential treatment" is to test whether credit-based insurance scores are statistically significant predictors of insurance loss propensity for the ECOA-protected classes. For example, if for every ECOA-protected class it were found that credit-based insurance scores added value to the risk assessment process as a statistically significant predictor of losses, then all classes would be receiving substantially the same treatment.

By its very nature, any risk factor used in pricing insurance will cause prices to be higher for some insureds and lower for others. Increases and decreases in premiums will exist within any grouping of insureds.

It is essential that, in evaluating whether the use of credit based insurance scores results in "negative or differential treatment" of ECOA protected classes, the methodology that is used recognizes that "negative or differential treatment" in insurance must be analyzed within the context of a consumer's risk of future loss. Under well established statutory and case law, treatment of a protected class is negative or differential only if it unintentionally results in members of that class as a whole paying more for insurance than is actuarially justified.

(5) What is an appropriate methodology to determine whether the use of specific factors in credit scores or credit based insurance scores results in "negative or differential treatment" of ECOA protected classes?

It is unclear that an evaluation is necessary, if the insurance score itself produces no negative or differential treatment on insurance price and availability.

The incremental effort needed to assess this question would be sizable. A determination of whether any component factor of a credit-based insurance score results in negative or differential treatment of ECOA-protected classes would require a significant expansion of the Study's database. In addition to attaching the credit-based insurance score to each policy record in the database, it would be necessary to attach a data field for each separate component factor of the score.

Further, the central question being researched is the insurance score, not the components of the insurance score, being used to price and underwrite insurance.

(6) What is an appropriate methodology to determine whether there are factors that are not considered by credit scores or credit based insurance scores that result in "negative or differential treatment" of ECOA protected classes?

If a risk factor is not being used, then by definition it could not be impacting the price and availability of insurance.

If the intent of the question is to identify factors not being used, but which might improve upon the current insurance scores, we encourage the Commission to focus on the predictive accuracy of one of the most commonly used insurance scores, rather than broadening the Study to invent new variables.

(7) In order to address paragraphs (a)(2) and (a)(3) of Section 215, data are needed on the geography, income, ethnicity, race, color, religion, national origin, age, sex, marital status, or creed of borrowers, potential borrowers, insurance customers, or potential insurance customers. Are these data available, and if so, where?

Insurers do not collect information pertaining to income, ethnicity, race, color, and religion, so it is not available as part of the policy record for any specific insured. Such data would need to be obtained from a public source and be attached to each policy record. The ideal situation would be to attach information to each policy record which is accurate for the individual policyholder. Unfortunately, as far as we understand, this information is not available. While some vendors sell estimates of such data, we are concerned about its accuracy, especially as it applies to race because it is based on different outside sources of varied quality.

If data for individuals are not available, then the next best alternate may be census-block data. Census-block may be the smallest unit of data for which demographic data are available from the U.S. government. It is preferable to use the smallest unit of demographic data available so as to be as accurate as possible about the demographic characteristics of each individual insurance policyholder.

Data concerning the geographical location of the insured, and the age, gender and marital status of the driver which is being rated on the policy are included on each policyholder record.

The geographical information is one of the most important data elements and also one of the most challenging for a multi-state study. Geographical risk is highly important to proper and accurate risk assessment for both auto and homeowners insurance. An analysis of impact on any of the ECOA-prohibited factors cannot be accurate unless the analysis properly controls for geographic risk.

For a variety of reasons the geographical indicators (i.e., rate territory and zip code) commonly found in all policyholder records cannot easily be used in a multi-state, multi-insurer study. We are currently researching data elements related to geographic risk, such as population density, that might be used in the multivariate analysis to control for geographic risk.

(8) If the data discussed in question 7 are not available, what proxies are available for the geography, income, ethnicity, race, color, religion, national origin, age, sex, marital status, or creed of borrowers, potential borrowers, insurance customers, or potential insurance customers?

As previously discussed in the response to Question 7, racial and income data for individual households are available from private vendors. We question the accuracy of the data, especially the information for some racial groups. The next best alternate to individual information may be census-block demographic data.

When used for rating purposes, age, gender and marital status of the rated driver for each insured car are in the policyholder records. We are currently researching a suitable indicator for geographic risk to use in lieu of the rate territory and the zip code which are in each policyholder record.

(9) If there are proxies for the geography, income, ethnicity, race, color, religion, national origin, age, sex, marital status, or creed of borrowers, potential borrowers, insurance customers, or potential insurance customers, what type of analysis would allow inferences to be drawn using the proxies instead of actual data on individual characteristics? What limitations are there to the inferences that can be drawn using proxies in place of data on individual characteristics?

One of the dangers of assessing a proxy is the tendency to look to the extremes.

We presume this question seeks input on the issue of aggregation bias. Several analytical approaches have been suggested by some groups, such as the "neighborhood model," Goodman's regression and "ecological inference." However, each of these methods has strong critics. The methods appear to be sensitive to simplifying assumptions used as input, thereby potentially resulting in inaccurate and misleading results. Care must be taken to frame these assumptions. If other known risk characteristics are not properly accounted for, then any such analysis will be so flawed as to be useless.

The best approach to avoid these problems would be to use as little aggregated data as possible. As mentioned above, risk-level data would be most desirable, but it is not (reliably) available for some protected classes, such as race, income, religion, etc. Such data are, however, available for risk characteristics commonly in use for insurance rating purposes.

(10) One potential proxy for individual characteristics may be Census data about the location where a borrower or insurance customer resides. What type of analysis would allow inferences to be drawn using data about the characteristics of the location where a borrower or insurance customer resides instead of data on individual characteristics? What limitations are there to the inferences that can be drawn using data about the characteristics of the location where a borrower or insurance customer resides in place of data on individual characteristics?

At this time we view census-block data as the next best possible alternative to individual-specific data, especially for some racial groups.[5] However, geographic risk differences are a concern in this sort of analysis. Territorial rating is used nearly universally in property casualty insurance. It is an accepted, legal, and valuable tool for predicting future loss. Yet, no single set of territorial grouping exists. Insurers generally use different boundaries from one another, and may have differences among their various lines of insurance.

Some areas have significantly higher loss propensity than others. Since some protected classes tend to be heavily concentrated in urban areas, a means to control for these geographically-related risk differences is critical for results to be meaningful.

One possible approach to address this issue of properly controlling for geographic risk in multivariate analysis may be to gather aggregate data on insurance claims at the zip code level for each line/coverage to be studied. That is, accumulate earned exposures, claim counts, and incurred loss amounts for each zip code in the United States from a number of insurers. Then a zip code's claim frequency and average cost per claim could be used as an indicator variable in the sample data used for the multivariate analysis described above. This aggregated data would seemingly be available much more quickly, and at a much lower cost than attempting to build a fully credible database containing all risk characteristics, including geographic and insurance score related data.

---

[5] Regardless of census-block data being the next best possible alternative to actual information on the demographics of policyholders, using census-block data is an imperfect approach with downsides. Not only is it merely a snapshot in time that disregards the fact that people are mobile, but it may ignore the fact that many neighborhoods and homes are racially diverse. Using census-block data is imprecise in that the distribution of race or of other demographic characteristics may differ at the extremes – conclusions drawn based on the extremes, but attributed more to a group broadly, may be flawed. It may therefore be dangerous to make an inference that a small unique subset is a fair representation of an entire demographic group. To reiterate a crucial point, an attempt to study impact without information matched to an individual's race, income, ethnicity, etc., will not produce a result based on actual data and, if used for assessing public policy, it should be regarded with caution.

Ideally, the aggregate data would pre-date the time period from which the sample data are drawn.

* * * * *

The Insurance Trades appreciate your attention to these comments and offer to supply additional information as questions may arise during the process of developing the Study methodology. We would like to reserve the right to supplement or amend these comments as you move forward through the process of studying this important issue. Finally, we would like to reiterate the importance of careful consideration of mitigating problems – given the limitations on the data – associated with a likely imperfect study methodology.

Respectfully submitted,

/s/

David F. Snyder
Assistant General Counsel, AIA

/s/

Neil Alldredge
Director of State Affairs, NAMIC

/s/

Diana Lee
Assistant Vice President – Research, PCI