

AMERICAN INSURANCE ASSOCIATION
STATEMENT ON
PROPOSED RISK CLASSIFICATION DATA CALL

*** * * * ***

NAIC PROPERTY & CASUALTY (C) COMMITTEE HEARING
SEPTEMBER 30, 2010

Catherine I. Paolino
Senior Counsel

As the NAIC Property and Casualty (C) Committee weighs a proposed risk classification data call, the American Insurance Association (AIA) respectfully asks you to carefully consider and balance all of the fundamental concerns outlined in these comments and urges you not to move forward with a similar data call. This testimony divides questions and problems into those that are data related and those that are administrative in nature. It also highlights some steps that some states already employ to learn more about credit-based insurance scoring.

DATA CONCERNS

In reviewing the proposed data call, insurers identified serious technical concerns in both the gathering stage at an insurance company and the analysis stage at an Insurance Department.

Importantly, if the information submitted in response to the proposed data call is inconsistent across the industry, it will be of questionable value for a regulator looking to better understand practices in his/her insurance marketplace. This would be particularly disappointing given insurers' cost and burden to provide the data. Below are several big picture issues.

Assumptions / Comparability

Where the proposed data call asks for all other factors to remain equal,¹ it seems likely that insurers could make assumptions about "the other things being equal" and arrive at different answers, especially if an insurer's rating plan contains rating factors that are interrelated.

For example, a non-member provides this illustration: assume an insurer's rating plan uses gender and age, working together so that there are six different factors.

		Gender	
		Male	Female
Age			
16-21	2.80	2.25	
22-25	1.50	1.35	
26+	1.00	1.00	

If by assuming "other things being equal" means that every one is age 45, then the range in the factors is (1.00/1.00) or 1.00, but if one assumes that everyone is age 18, then the range would be (2.80/2.25) or 1.24.

Without at least some common assumptions provided, the inconsistency between insurer interpretations would jeopardize the validity of conclusions a regulator may draw from this data.

Storage & Reprocessing

The kinds of information stored and the manner in which rating factors work together are important topics for dialog between regulators and insurers.

¹ Consider the request in **Table III** for highest and lowest factors while holding all other factors equal.

One insurer explains that the particular information kept by an insurer and the places where the information may be stored may vary by company. For many insurers, the data and factors used to calculate the initial premium will not be stored in the system. As data and factors change (e.g. adding a youthful), the fields populated by the initial data are overlaid with the new data and the new factors. Factors, premium, insurance scores, scoring bands, territories, driver information – all these things may change throughout the tenure of the policy.

To go back to find the closest approximation for correct relativities may demand significant manual review and analysis; for some insurers this may mean recalculation by hand. The multivariate nature of their rating further complicates the process.

Another insurer explains that the computer programs that calculate the policy premiums generally are designed to calculate the total premium (rather than to isolate the impact of each individual rating variable). Thus, in order to provide the information requested in the proposed data call, many insurers will need to revise their policy rating programs to isolate and store individual rating variable information. These rating programs are extremely complex and they are expensive to modify. After program modifications are accomplished, these insurers would have to process millions of policies through the revised rating programs to collect that individual rating variable information and conduct extensive calculations to put the data into the requested format.

For many insurers, due to interaction at the coverage level among multiple risk factors, to calculate the distribution just for one risk factor, these insurers would be required to go through complex and tedious efforts. Yet another insurer reports that while such exercises are theoretically possible, they would require thousands of IT hours.

Even after rebuilding the rating system and reprocessing the policyholders, the information provided to the regulators would be of questionable value because the proposed data call does not reflect the actual sophistication of most personal auto insurers.

Similarly, decline-to-quote ratios for the years immediately preceding and following implementation of credit-based insurance scoring would not be available for many insurers.² Such information is generally maintained for only several years; many insurers have been using scoring for more than a decade. It is our understanding that in many cases, even if insurers do have that data available, the process of collection would be manual.

Isolating Impact & Interactions

Portions of the proposed data call do not reflect the current level of sophistication of most personal auto insurers. While they seem to assume that for each rating variable there is a single factor that applies to the entire premium, this is not the case. An algorithm that could isolate each criterion's effect on the policyholder's rate would be extremely difficult to put together.³ Here are several examples to illustrate the problem:

Most auto class plans have factors that vary by age, gender, marital status, and operator status (e.g. occasional vs. principal operator). If a 17 year old single male, principal operator has a factor of 5.00 in the class plan, it would be nearly impossible to determine

² This comment applies to **Additional Information Question 16**. Consider AIPSO data on broader information on shrinking residual market premium.

³ This would be required by **Table IV** and **Table V**, which ask for a detailed distribution of policyholders receiving a credit/surcharge by risk classification criteria.

how much of the 5.00 factor is due to his age being 17, vs. his gender being male, vs. his being single, vs. him being the principal operator. The factor for that company was determined as the combination of the interaction of the four attributes.

Many risk classification variables vary by vehicle within the policy, for example annual mileage. Multi-car policies will have multiple annual mileage factors applied. Yet, the proposed data call requests information by policy. It is unclear how insurers are to report vehicle level information on the requested policy level basis.

Most insurers use a separate set of rating factors and a separate rating formula for each of the major coverages (Bodily Injury, Property Damage, Uninsured Motorist, Personal Injury Protection, Comprehensive, and Collision). While an insurer generally will assign a credit-based insurance score for a policy, such single score may generate separate factors for each of the coverages. It is unclear how insurers are to report these multiple credit factors on the requested policy level basis.

Categorizing the ranges of rate relativities is a complex process due to the interactivity of factors.⁴ While companies could provide the lowest possible rate relativity and the highest possible rate relativity for some risk classifications, they could not do so for all classifications. Some classifications do not have a rate relativity independent of another risk factor. Many risk characteristics are co-dependent on others. Even if an insurer indicated a lowest possible rate relativity, in a multivariate environment, the risk classifications interact with each other and a single relativity is not accurate.⁵ This would not reflect customers actually charged that relativity.⁶ Some risk classifications are not independent. Others are codependent with so many others that the interactions between them make it impossible to determine the distribution of policyholders receiving a discount or surcharge off the median or average.

Said another way, the proposed data call seems to assume one score/factor applies to the entire premium. Today, for many insurers, many rating factors do not act in isolation or independently of one another. Rather, sophisticated rating models are built to leverage interactions among factors.⁷

... In general, the premium that an insured is charged is determined by many considerations... The relative weights for a given factor often vary with other factors and vary uniquely and considerably among insurers. How the factors combine to determine the premium also varies among insurers. Because rating algorithms are general multiplicative in nature, it is confusing to assign a credit or weight to any individual factor without first agreeing upon a set of rules or procedures that would be used to calculate or organize all the rating variables used in the algorithm. Further, while it might be possible to go through several different mathematical exercises and assign various weights to each of several components, such calculations could be misleading and could suggest incorrect conclusions. *[Followed by an example in Appendix A.]*⁸

At best it would take significant resources to satisfy the requirements, which would require all of an insurer's policies for a given state to be individually run through the algorithm. At worst,

⁴ This is particularly evident in **Table III**.

⁵ See **Table III**.

⁶ See also **Table IV** and **Table V**.

⁷ Much of **Table III** may be unanswerable given current systems because of interaction among variables.

⁸ See the May 15, 2009 Supplemental comments submitted by the American Academy of Actuaries.

results from different sources would not be comparable because insurers may interpret the request differently.

Definitions / Comparability

Without outlined standard approaches or definitions, the inconsistency between insurer interpretations would jeopardize the validity of conclusions a regulator may draw from the data. Consider this problem in the context of the definitions discussed below.

Median Value Definition⁹

While the definition of “median” relates to the middle number in a series of numbers, the proposed data call is unclear as to which figures to include and how to approach binary situations.

For example, in determining the median for a rating variable like territory it is unclear whether: (1) the high level values should be considered (assuming five values with rating factors of 0.77, 0.82, 0.86, 0.95, and 1.00 – the median value is 0.82); or the customer legal values should be considered (assuming 400 customers in the lowest rated territory, 250 in the second lowest, 200 in the third lowest, 400 in the fourth lowest, and 901 in the highest rated territory – the median value is .95).

For example, in determining the median for factors where there are only two results (like home ownership), it is unclear whether: (1) there is no median value because there is no middle value in a series of only two values; or (2) the distribution of insureds is considered and the median value would be assigned to the factor of the group that had the most insureds.

Renewal Definition¹⁰

State requirements differ with respect to the handling of “renewal” business, when it comes to credit-based insurance scoring. It is unclear when a “yes”/“no” may be appropriate:

For example, where a state requires credit information to be run every three years it is unclear whether: (a) all renewals would be considered, or (2) only those that have been rescored on a particular renewal would be considered. The different approaches may give the regulator insight into different aspects of the impact.

For example, where a state limits renewal use of credit to situations where a score improves, it is unclear whether the insurer would be expected to provide information on the score that was returned even if it was not actually used for the customer (because it was not an improvement).

Note, that other criteria – like age, accident history, etc. – may be updated automatically at every renewal.

Credit-Based Insurance Scoring¹¹

⁹ Consider where “median” is used, including **Tables III** and its importance in **Tables IV** and **Table V**.

¹⁰ This comment is particularly relevant for **Additional Information Question 17**.

¹¹ This definition applies through the entire proposed data call.

There are a wide variety of scores – some from vendors and others are proprietary to an insurer. Not only do the models differ, but the scores may have different meanings.

For example, it is our understanding that FICO-style scores have ratings between 250ish and 850ish with a higher score meaning a lower loss frequency/severity. Other scores may be designed with a higher score meaning a higher likelihood of filing claims. Some proprietary models may not even derive a score.

Other

There may be additional terms that require clarification, including:

Vehicle value,¹² in connection with risk classification.

Vehicle safety features.¹³

Driver primacy,¹⁴ as it relates to lowest possible rate versus highest possible rates (driver primacy is the rate for the primary driver).

Risk classification,¹⁵ should not be so broadly defined as to include marketing items.

Tier/company placement,¹⁶ which may be defined differently in the data call than in a particular jurisdiction.

Extraordinary life circumstances,¹⁷ should allow for an explanation beyond a yes/no.

Foreclosure,¹⁸ which is vague as to whether it is considered alone or in a credit-based insurance score.

ADMINISTRATIVE DECISIONS

Insurance Departments feeling they need to take additional steps to learn more about risk classifications will be faced with a number of important administrative decisions, whether they move ahead with something like the proposed data call or with another alternative. AIA encourages your early consideration of these decisions and of the important issues they raise.

Data Call Approach – Open Questions

Applicability

¹² See **Table III**.

¹³ See **Table III**.

¹⁴ See **Table III**.

¹⁵ See **Table I** and **Table III**.

¹⁶ In particular, note **Table I** and **Additional Information Question 8**.

¹⁷ See **Additional Information Question 3**.

¹⁸ See **Additional Information Question 4**.

The extent to which the proposed data call would be applicable is unclear. For example:

Domicile – Whether the request would be limited to those insurers domiciled in a state is unclear. AIA asks regulators to consider conversations with domiciliary insurers to better understand risk classification factors, before turning to an extensive data call.

Size – If moving forward, AIA suggests that the proposed data call not be distributed to all those doing business in a state. Rather, it should be limited to the top writers on an individual company basis. While the impact on all insurers would be meaningful, the staffing and costs would particularly impact smaller companies.

Date¹⁹

Given that rating factors and models are updated from time to time, for any data call it would be important to provide a common date on which to report the factors used and how. The proposed data call refers to the company's complete "current" rating plan, but it is possible that the plan could change while the response is being assembled. Rather than worrying about a moving target where updates may be required during the process of preparing the submission, AIA asks that you specify a common date.

Vehicle for Issuing Data Call

As far as AIA is aware, no statement has been made about the process states that elect to issue the data call will follow. Consider what vehicles offer appropriate discretion to the regulator and protections to the insurer. AIA urges careful advance review of applicable state laws.

Confidentiality & Proprietary Information

The proposed data call presents serious questions about confidentiality and protecting proprietary information.²⁰ Areas where insurers differ, particularly in terms of other risk classification factors, may highlight their competitive advantage. These very open ended inquiries ask for proprietary and confidential information. Insurers should be provided with clear indications of how this information will be protected.

Data Analysis

To the extent there may be major administrative costs for analyzing the results, AIA asks Insurance Departments to consider how they will be addressed in advance of issuing the data call to insurers.

Whether states plan on coordinating and or aggregating the information has not been part of the proposed data call discussion. The mechanics of these plans – and how they address the confidentiality concerns – will interest industry. AIA asks that these procedural and administrative aspects of analysis be added to the discussion soon.

¹⁹ This concern would apply across the entire proposed data call.

²⁰ While this is an overall concern, it is particularly critical on **Tables II** and **Table V** where insurers are asked to reveal additional and possibly less common factors.

Timeline

As far as AIA is aware, no mention has been made of the timeframe an insurer would have to prepare a response to a data call. AIA strongly encourages adding this to the discussion as soon as possible. Given the scope and other problems, insurers have serious concerns about the amount of time and flexibility that would be provided.

Scope

The breadth of the data call and the elements considered is vast. AIA strongly urges you to narrow the scope of the request and perhaps reassessing whether further information is needed after reviewing the results of a more targeted request. If the primary purpose of the data call is to evaluate how insurers use credit-based insurance scores, please limit the request. Importantly, please evaluate whether there are more tailored approaches to getting at the questions concerning regulators. Possible alternatives are discussed in another section below.

The predictive nature of a risk factor should be the standard against which it is judged. For credit-based insurance scoring,²¹ some are concerned that focus around the testing periods, data year, calibration, and model age may mean that regulatory emphasis is moving away from whether the model remains a predictive tool. Beyond this issue, insurer practices are a matter of competitive advantage.

Alternative Approaches - Existing or Available Tools

States have used various lower cost tools that seem to address some of the objectives articulated as the (C) Committee discussed the proposed data call. Below are some examples:

To address the question of policyholder-level impact of changes in credit-based insurance scores, perhaps something could be loosely modeled after the consumer materials in **Pennsylvania**²² or **Massachusetts**.²³ Looking at hypothetical policyholders may be an easier, more direct, and less burdensome and costly way for Insurance Commissioners to supplement their understanding of cost variance resulting from certain discrete rating factors. By asking insurers to rate hypothetical policyholders with particular risk profiles (holding constant key variables), the regulator may get some of the answers sought in the more expansive proposed data call. Because it is based on a rating approach insurers go through daily (without systems or operational changes) the results may better reflect the marketplace and may avoid the actuarial and other problems discussed above. If the (C) Committee moves in the direction of hypothetical policyholders, AIA asks to be involved in the process and AIA would urge that the scope be reasonable and that the focus be narrowly defined. Factors to hold constant (with particular risk characteristics provided) would likely include commonly used variables like: credit-based insurance score, marital status, loss history, lapse in coverage, prior limits, gender, age, territory, liability limits, tickets or traffic violations or points, at-fault accidents, not-at-fault accidents, ownership of a home, vehicle year/make/model, number of cars in household. For any additional factors used, insurers would be required to submit all assumptions. While AIA

²¹ See **Additional Information Questions 8d, 18 and 19**.

²² See the assumptions provided in the Rate Comparison Chart materials by types of coverage on the Pennsylvania Insurance Department website: http://www.portal.state.pa.us/portal/server.pt/community/auto_insurance/9187

²³ See Auto Insurance Premium Comparisons on the Massachusetts Office of Consumer Affairs & Business Regulation website: http://www.mass.gov/?pageID=ocaterminal&L=4&L0=Home&L1=Consumer&L2=Insurance&L3=Automobile+Insurance&sid=Eoca&b=terminalcontent&f=doi_AutorateCompare_autoratecompare&csid=Eoca

does not think this is a perfect solution, it is listed for additional brainstorming around a possible way to assist regulators interested in looking at the data.

To address the question of the general overall impact of credit-based insurance scores, the **Arkansas** legislatively-mandated study²⁴ is a simple and effective tool for quantifying: (1) the number of customers who have had their rates increase due to credit; (2) the number of people who have seen their rates decrease due to credit; and (3) the number of people who have had no premium impact. Given that Arkansas has been conducting the study for several years, their detailed instructions allow for rather consistent data across companies.²⁵

To consider the question of consumer complaints in a jurisdiction, regulators could tap information housed by the **NAIC** and/or by **individual states**. Complaint data is especially reliable in the context of the insurance scoring notice/disclosure system, which requires consumers be provided “adverse action notices.” A regulator may keep an eye on what is happening in a state with respect to a particular company by monitoring complaints, considering trends and engaging in market surveillance/conduct activities. Good complaint data will give a clear indication of the magnitude of a problem. If current complaint data systems do not identify complaints arising out of scoring, then the complaint data systems could be changed to do so. These complaint numbers should then be put into context, by looking at them as a percentage of personal lines policies issued and renewed. To the extent regulators want more specific information about credit-related complaints it seems possible to consider additional coding to capture this specific information.

To consider overlap in what risk factors measure, some states, including **Washington**,²⁶ may require additional actuarial support and multivariate analysis in credit-related filings. While it is more work for insurers to provide this information, as far as AIA understands, insurers have responded to this request.

* * * * *

Given of the many technical issues presented by this proposed data call, some of which are outlined above, AIA respectfully urges the (C) Committee not to move forward with the proposed data call. To the extent the Committee feels it needs to do something and it is considering modifications and alternatives, it may not only be more efficient to target a search more finely and to tailor a request more narrowly, but it may lead to more effective and useful results.

AIA appreciates the opportunity to provide written testimony as well as to participate in the upcoming (C) Committee hearing on the proposed data call. AIA is still collecting information and reactions from members, so we ask to be permitted to modify and supplement this statement. AIA looks forward to working with you as members continue to provide input to facilitate improved regulatory understanding of risk classification factors.

²⁴ See the Annual Reports section of the Arkansas Insurance Department website:
<http://www.insurance.arkansas.gov/Administration/informationssystem/annualreport/annualreport.html>

²⁵ If the Committee moves forward with a template similar to the proposed data call, the clarifying Arkansas Instructions outline the definitions of “increase” and “decrease” may be instructive as this Committee considers consistent approaches to **Additional Information Questions 21 and 22**.

²⁶ See Wash. Admin. Code 284-24A-001 et seq.