in accord with a fee schedule is *not a service benefit* and the collection of reliable data on the additional charges made by physicians would be well nigh impossible. It is for these reasons that the existing statistical and claim processing set-up does not provide for the collection of this information.

I concur with Mr. Bevan that a further analysis by diagnosis of the Master Medical category "All Other" is desirable. This will be possible because the statistical card contains a three digit international classification of disease code. It is hoped that such information will be made available in the near future.

There is no doubt that a further study is needed to explain the "freakish" phenomenon of the distribution by size particularly for Master Medical. The short time at my disposal did not permit the desired analysis but the following facts may throw some additional light on the situation:

(a) For all years (1956 to 1959) the average cost of a coronary attack was $284. While this is an arithmetical average, subsequent studies may confirm that a large number of such cases are close to this average. From Table VI it is seen that this diagnosis accounts for approximately 9% of all Master Medical cases.

(b) The size of loss analysis for the year 1959 is based on incurred experience as of March 31, 1960. From Table II we find that the payments are only 42% of the ultimate incurred cost. A very large number of out-standing claims carry a flat reserve of $300. (The over-all average cost indicated by the 1959 experience as shown in Table VII for family contracts is $354.94.)

I am grateful to Mr. Bevan for taking the time to prepare the discussion of my paper and to call attention to the various elements where detailed information is of essence in rate making.

EXPERIENCE RATING REASSESSED

BY

ROBERT A. BAILEY

Volume XLVIII, Page 60

DISCUSSION BY JOHN W. CARLETON

Mr. Bailey has made a timely contribution to the technical literature. His examination of the structure of experience rating is well organized, and the conscientiousness with which he sorts out and labels the areas of judgment is to be commended.

The paper should be reviewed by someone qualified to deal with the mathematical developments. Absent such a qualification, I find it necessary to comment on the criteria which Mr. Bailey uses to define a good experience rating plan.

He starts with four fundamental criteria, the first relating to the measurement of something called information and the other related to practical aspects of the insurance transaction. If I understand him, he says the actuary

should first find the formula that uses risk loss experience so as to maximize predictive accuracy. Then he should see if the formula so obtained will be acceptable to underwriters and buyers. If so, proceed; if not, compromise prudently.

There is something troublesome about the concept of a correct design that may have to be compromised in order to make it work. Such a concept of correctness is suspect. Perhaps it would be worthwhile to examine the statistical model that seems to underlie Mr. Bailey's symbols.

Statistical, mathematical or economic models can be used either to advance the frontiers of abstract thought or to strip a complex practical problem to its manageable essentials. It is only in the latter use that it is proper to be concerned about the degree of correspondence between the simplified model and the more complex reality. The correspondence can never be perfect, but if it is too thin a different approach to model design may be suggested.

In his paper Mr. Bailey does not refer to a model. The model I'm going to talk about is one I infer from what he says. This is dangerous—dangerous for me. He may have a way of looking at the problem that requires no model, or he may have a much more sophisticated one than I would be capable of sensing.

His model seems to be the one frequently used to maximize accuracy in predicting "inherent hazards." The expression "inherent hazard" is used to label a quantitative attribute of individual risks. Within the model, risks are assumed to be heterogeneous in this attribute, but they lend themselves to classification in such a way that the dispersion of inherent hazards within classes is less than for the total risk population. Certainly, common sense and all known data fully support this obvious assumption, and perhaps that support is sufficient to give the concept of "inherent hazard" working reality.

Most actuaries know what is meant when one of their brethren talks about "inherent hazard," but many of them have difficulty defining the concept tightly and without circuity, particularly if they are to stay inside their statistical model and assume that the risk population, both within and among classes, is heterogeneous. For definition purposes, many have been forced to set up a second model of homogeneous risks, describe the behavior of such a model, and then define the inherent hazard of an individual risk in terms of its membership in such a group.

"Inherent hazard," as we use the expression in our models, is not part of the working vocabulary of many underwriters. It is safe to assume that it is in the working vocabularies of even fewer insurance buyers. This should not concern us if we were confident that we have a precise, if esoteric, way of expressing what the buyers and sellers actually want to accomplish, but it should concern us if there is any possibility we are letting the model dictate the pricing problem rather than the other way around. We can examine the model further with this in mind.

In the simpler statistical models used in experience rating analysis, the inherent hazards are assumed not to change with the passage of time. Perhaps it would be fairer to say that they are assumed to have a degree of continuity such that changes can be ignored within time spans that encompass the sum of the experience period, the rating lag and the period to which the revised rates are to be applied. I know of no objective support for this assumption of continuity. It seems reasonable—perhaps more reasonable than an under-

writer's annual renewal review or a safety engineer's optimism—but reasonable or not the assumption is inconsistent with some operating practices in the business.

The model is operated in such a way that the accuracy of the inherent hazard prediction is enhanced by adding information from the risk's own loss experience to information already used in setting up the class rate. Mr. Bailey uses the word "information" in a sense similar to that in which it is used by information theorists. Many of the papers on Credibility in the Proceedings were published after the mathematical foundations of information theory were laid down but before they were translated into lay English that nonprofessionals could hope to understand. There now appears to be a striking parallelism. Information is measured by the extent to which it reduces the uncertainty of the receiver. Thus, the same message will carry more information to one receiver who has a lot of uncertainty than it will to a second receiver whose uncertainty has been cut down by other sources.

Similarly, individual loss experience may add very little to what the underwriter already knows about the hazard of a risk if a great deal of good information has been poured into a refined classification system. On the other hand, the same risk's losses may be entitled to quite high credibility if the ratemaker's only prior information is represented by some classification rates set up with guesswork.

When this relationship is worked out mathematically in the statistical model, the uncertainty implicit in (or the lack of information conveyed by) the class rates is usually expressed as the dispersion of inherent hazards within classes. Mr. Bailey has done this, and followed through consistently. There is a troublesome consequence:

As the valid information upon which class rates were established approaches negligibility, the theoretical recognition that should be given risk losses, regardless of risk size, approaches one hundred per cent. When Mr. Bailey collides with this result he modifies the criteria to emphasize fluctuation control at the expense of predictive efficiency.

For this weighing of the information carried in the class rate against the information carried by the risk's losses, the model itself seems deficient because of a conspicuous noncorrespondence with reality. In the model we have no information about the validity of the class rate other than the dispersion of inherent hazards within classes and we have no information about peculiarities of the individual risk other than incurred losses, claim by claim, during the experience period. In practice we seldom have any knowledge of the dispersion of inherent hazards within classes but usually do have considerable background information about the source of the class rates which should contribute to informed judgment on their application to an individual risk. Moreover, we quite commonly have a good deal of information about the individual risk that should enable us to match it against the source of the class rate. Of course, this is only to say this statistical model uses only the kinds of information that it was designed to use.

To some, predictive accuracy is an end in itself; to others it is only a contribution to a larger problem of finding the points where buyers and sellers can meet and agree. To examine this difference we can look for the economic or market model in which the best predictive accuracy would be mandatory for survival.

If all competing carriers were to use the same experience rating plan year after year, if the plan's predictive accuracy were lower than it need be, and if all underwriters were equally ignorant of how low that accuracy was; then nothing adverse would happen that an off-balance factor couldn't cope with. The three requisites might be difficult to maintain indefinitely, but while they were maintained the plan could be looked upon as one in which the policyholder pays some portion of last year's losses next year and thus is rewarded for being a good risk and punished for being a poor one. Only enough predictive accuracy is required to keep the third requisite in effect.

At the other extreme, there is a market model in which each carrier uses a plan different from that of every other carrier. More importantly, each policyholder is a perfect price buyer who considers each year's insurance as a separate transaction and annually shops the entire market for the lowest renewal quotation which he buys. Under such conditions, I believe all carriers having experience rating plans with less than the best predictive accuracy would be in financial difficulty. I'm not certain what would happen to the carrier whose plan had the best predictive accuracy. It might do satisfactorily or it might only be in less financial difficulty than the companies with less gifted actuaries. But in either event the overpowering demand for predictive accuracy would call for statistical models capable of using more and different kinds of information than the model we have been talking about.

Of course, such a concept of the market doesn't correspond with reality either. Other things being not too unequal, most buyers prefer to maintain a continuing relationship with the same carrier or producer. Service effectiveness and service satisfaction typically improve with time. Even price buyers tend to have more confidence in longer period comparisons than shorter period ones. A company that rates its business in such a way as to make its better customers feel at home should expect them to respond by staying there.

Is it possible to imagine a statistical model that has a closer correspondence to the pricing problem? In such a model predictive accuracy probably would not be controlling, but certainly conspicuous deficiencies of it would limit the inertia of the business. That inertia-like characteristic would be recognized quantitatively, together with the factors that contribute to it. The model should permit the buyer to dissociate himself from insurance pooling to the extent that he is willing to absorb his own losses, either directly or in rating. If a simpler definition of rateable losses brings about an easier meeting of the minds, the model should balance that gain against the loss of predictive efficiency. If the buyer wishes both to minimize pooling and to contain fluctuations, the model should permit him to extend himself in time.

It is easier to point to the elements that an existing statistical model does not contain than it is to design a better one. I am not at all certain that a better one can be designed or that one containing the elements I have mentioned would be a foundation for constructive mathematical inquiry. I do feel that Mr. Bailey could do it if anyone could.

DISCUSSION BY LEWIS H. ROBERTS

This paper is distinguished by two virtues which are unfortunately not often found in combination: on the one hand, incisive theoretical analysis, and on the other, thorough practicality. The first deserves mention because the