

4A

Fuzzy Trends in Property-Liability Insurance Claim Costs

J. David Cummins
Richard A. Derrig

ABSTRACT

Accuracy in forecasting expected loss costs may well be the most important determinant of the ultimate profitability of a cohort of property-liability insurance policies. The existing literature on claim cost forecasting focuses on the selection of the "best" forecasting model or method, discarding information provided by closely ranked alternatives. In this article, we emphasize the selection of a "good" forecast rather than a forecasting model, where goodness is defined using multiple criteria that may be vague or fuzzy. Fuzzy set theory is proposed as a mechanism for combining forecasts from alternative models using multiple fuzzy criteria. The fuzzy approach is illustrated using forecasts of automobile bodily injury liability pure premiums. We conclude that fuzzy set theory provides an effective method for combining statistical and judgmental criteria in actuarial decision making.

Introduction

The problem of forecasting expected loss costs for a prospective policy period is one of the fundamental responsibilities of the property-casualty actuary. Forecasting accuracy may well be the most important ingredient in determining the ultimate profitability of the policy cohort (Derrig, 1987). Despite its critical nature, the actuarial literature on forecasting is rudimentary (Casualty Actuarial Society, 1990), and the more advanced research has focused primarily on forecasting severity (Cummins and Powell, 1980; Cummins and Harrington, 1985; and Cummins and Griepentrog, 1985).

The actuarial trending literature devotes little attention to the evaluation of alternative methods. Actuaries typically choose a model from a limited set of alternatives that has been assembled judgmentally. The model selection process usually does not have a formal feedback loop; that is, the selected model or models are almost never tested ex post to determine whether they performed well. The insurance economics literature (articles cited above) has focused on the forecasting accuracy of alternative methods, including time trend and

J. David Cummins is Harry J. Loman Professor of Insurance and Risk Management at the Wharton School, University of Pennsylvania. Richard A. Derrig is Senior Vice President of the Automobile Insurers Bureau of Massachusetts.

429

econometric models. However, these studies have been oriented toward the identification of a "best" forecasting method, whereas the true objective of the process should be the selection of a forecast and not a method. Stated differently, it is not necessarily true that a method that was in some sense "best" in the past will provide the most accurate forecast of claim costs in the future. The best-method approach not only focuses on the wrong objective but also disregards the information provided by alternative trending methods. This is inappropriate in view of the growing body of literature demonstrating that forecast accuracy can be improved by combining forecasts (Zarnowitz and Braun, 1992; Clemen, 1989) and that no one method can be expected to perform best over time for different data series (McNees, 1992).

In our view, the presumption that there exists a single "best" model is fundamentally flawed. Not only do alternative models provide additional information, but models with more or less equal historical performance rankings can provide vastly different *ex ante* forecasts. Moreover, many of the criteria used in selecting a forecast are inherently vague; for example, actuaries may seek forecasts that are "accurate," "adequate," or "reasonable." A general setting or framework is needed within which the complexity and inherent vagueness of the model/forecast selection problem can be highlighted and accommodated.

The framework proposed here is *fuzzy set theory* (FST), a relatively new branch of mathematics, which provides a mechanism for dealing with vague or "fuzzy" data and decision criteria. FST can be used to model linguistic variables such as "a tall person," "a good risk," and "accurate forecast," etc., which are mostly beyond the realm of conventional mathematics. FST also provides a way to systematize a decision process in the presence of multiple criteria, some or all of which are fuzzy. It offers a framework for retaining useful information provided by models that may not be "best" in the conventional sense. It also facilitates the integration of judgmental information with data that are more or less objective. Much more than a purely theoretical construct, fuzzy logic has been used in a wide variety of practical applications ranging from the design of video cameras and subway cars to medical diagnoses and computer network design. FST has been applied to insurance by Lemaire (1990).

The fundamental concept of FST is the alternative formalization of membership in a set to include the degree or strength of membership. In conventional set theory, membership in sets is dichotomous: an element either is or is not a member of a given set. In FST, on the other hand, elements can be definite members or non-members of a set, but they also can be more-or-less members of the set. The black and white membership of conventional set theory is replaced by black, white, and all shades of gray.

The attractiveness of FST for the forecasting problem comes from an exposition of what is meant by a "good" forecast or forecasting method. FST can be used to address practical forecasting problems by viewing individual forecasting methods as candidates for membership in the fuzzy set of "good" methods and/or by viewing individual forecasts as candidates for membership

in the fuzzy set of "good" or "accurate" forecasts. FST provides a natural framework for studying forecasting problems that combine statistics and judgment to satisfy multiple decision criteria, some or all of which are expressed verbally rather than quantitatively.

This article demonstrates the use of fuzzy set theory as a methodology for the selection of insurance claim cost forecasts. The intent is to develop a method for integrating judgmental and statistical information to evaluate or grade a set of forecasts according to an appropriate set of criteria. The method provides a natural framework for capturing the information from forecasts produced by a wide range of methods in proportion to their degrees of membership in the fuzzy set of "good" forecasts.¹ Contrary to the conventional approach, the goal here is not to choose the "best" forecasting method. The answer to the forecasting problem posed in this article is the fuzzy set of good forecasts rather than any given forecast or forecast method. However, because insurers must choose a single trend factor to include in the rates, we also propose a method for producing a trend factor taking into account the information provided by the fuzzy set.

We conduct forecast experiments similar to those presented in Cummins and Powell (1980), Cummins and Harrington (1985), and Cummins and Griepentrog (1985). But while they tested forecasts of severity, we focus on pure premiums, because the ultimate effect of forecasting on profits is determined by the accuracy of forecasts of pure premiums, not by severity alone. Massachusetts private passenger automobile insurance data are used in the experiments. Fuzzy set theory is then used to rank the forecasting methods in terms of their membership values in the fuzzy set of "good" forecasts, and the membership values are used to select a trend factor.

The section below outlines the claim cost forecasting problem, and the next section discusses the results of the experiments. The results point to the difficulty of choosing the best forecasting method because of multiple "goodness" criteria and the diversity of forecasts generated by alternative methods. Following an overview of fuzzy set theory and a discussion of its use in decision problems, we illustrate the use of fuzzy set theory by evaluating the alternative forecasts in the Massachusetts experiments.

The Forecasting Problem in Property-Liability Insurance

Importance of Claim Cost Forecasting

The important role of claim cost forecasting in property-liability insurance ratemaking is apparent from a simple discounted cash flow ratemaking model,

¹ The objectives of this article are consistent with Actuarial Standards of Practice No. 13 (ASP 13), which deals with trending (Actuarial Standards Board, 1990). ASP 13 states that "the purpose of trending procedures is to estimate future expected values by analyzing historical data and other relevant information." That is, the objective is the forecast and not the method. ASP 13 also explicitly states that judgmental and statistical information should be recognized in selecting a forecast and that a range of methods should be considered rather than a single method.

which gives the premium for a policy issued at time t with a single loss payment at time $t+j$:

$$P_t = L_t \frac{1 + \Delta_{t+j}}{(1+r)^j}, \quad (1)$$

where P_t = the premium,

L_t = expected loss payments at price levels prevailing at time t ,

Δ_{t+j} = the predicted change in the price level between the issue date (time t) and period $t+j$,

where current price level = 1.0,

r = the risk-adjusted discount rate for losses = $(1+r_f)(1-\pi)-1$,

r_f = the risk-free rate, and

π = the risk premium (profit loading).

Anticipated changes in claim costs are reflected in the *trend factor*, $(1+\Delta_{t+j})$. If insurance claim cost inflation were identical to general (economy-wide) inflation, it would not be necessary to devote much attention to estimating the trend factor. Interest rates are highly correlated with anticipated economy-wide inflation, and there is considerable evidence in support of the Fisher hypothesis, which states that: $(1+r_t) = (1+i)(1+r_f)$, where i = the expected rate of inflation, and r_f = the real rate of interest. If expected insurance and general inflation were identical and the Fisher hypothesis were correct, no explicit trend factor would be needed and claim costs would be discounted at $(1+r_t)(1-\pi)$.

Insurance inflation usually is not equal to general inflation. For example, during much of the 1980s, inflation in private passenger auto insurance on a countrywide basis exceeded general inflation by a margin approaching three to one (Cummins and Tennyson, 1992). Thus, it is important for the trend factor to reflect anticipated insurance inflation. Claim cost forecasting is the method used to estimate insurance (claim cost) inflation.

Errors in the trend factor have nearly a one-for-one effect on realized profit. To illustrate, consider a policy with a single loss payment one period from the policy issue date (t). Assuming that insurer assets are invested at r_f , the insurer's net income (Y) at the end of the period is

$$Y = P_t(1+r_f) - L_t(1+\Delta_a) = L_t \frac{1+\Delta}{1-\pi} - L_t(1+\Delta_a), \quad (2)$$

where Δ = the predicted change in the price level, and

Δ_a = the actual change in the price level during the period.

Suppose net income is normed to the correct (perfect foresight) premium accumulated at interest, that is, to $P_a(1+r_f) = L_t(1+\Delta_a)/(1-\pi)$. Then we have

$$\frac{Y}{P_a(1+r_f)} = \frac{L_t(1+\Delta) - L_t(1+\Delta_a)(1-\pi)}{L_t(1+\Delta_a)} = \left[\frac{\Delta - \Delta_a}{1+\Delta_a} \right] + \pi, \quad (3)$$

where $P_a = L_t(1+\Delta_a)/(1+r)$ = the perfect foresight premium.

The bracketed expression in equation (3) is the proportionate error in predicting loss growth, which we call the *total predicted change error* (TPCE). The TPCE is a direct offset against expected profit. Because π is usually small

(less than 0.05), trending errors can drastically alter the insurer's profit position. For example, if $\pi = 0.025$, $\Delta = 0.05$, and $\Delta_a = 0.08$, the insurer will realize a profit of -0.0028 rather than the expected profit of 0.025.

One of the principal reasons that insurers expect to make a profit by underwriting insurance policies is that the ultimate cost associated with those policies is uncertain at the time of sale. Doherty and Garven (1986) and Cummins (1988) have shown through options pricing models that even unsystematic loss cost uncertainty directly affects the fair price for the policy. Thus, the risk of claim cost trending error will be reflected in insurance prices. Insurers that are relatively unsuccessful at forecasting loss costs will not earn an adequate return on equity, reinforcing the need for the adoption of the best available forecasting techniques.

The Actuarial Approach to Trending

Expected losses are determined by the following relationship:*

$$E(L) = E(F) E(S), \quad (4)$$

where $E(L)$ = expected losses per exposure unit (the pure premium),
 $E(F)$ = expected claim frequency per exposure unit, and
 $E(S)$ = expected claim severity (monetary loss per claim).

Thus, ratemaking could proceed by either developing separate trend factors for frequency and severity and multiplying the results to obtain the pure premium trend or developing a single trend factor for the pure premium. Under some circumstances, disaggregation can create a pooling effect due to offsetting errors, leading to improved forecasts (Armstrong, 1978, pp. 455-458). However, this effect is not always present, and the value of disaggregation is ultimately an empirical issue.

The standard actuarial approach to forecasting claim costs is to develop separate trend factors for frequency and severity using simple linear or exponential time trend models. The models are estimated using ordinary least squares (OLS). The following equations and estimation period are typical:

$$\ln(S_{t-i}) = \alpha_s + \beta_s(t-i) + \varepsilon_{s,t-i} \quad (5)$$

$$\ln(F_{t-i}) = \alpha_f + \beta_f(t-i) + \varepsilon_{f,t-i}, \quad (6)$$

where $t-i$ = estimation period index, $i = 0, 1, \dots, 11$, and time t = the present,

α_j and β_j = coefficients to be estimated,

ε = is a random error term,

S_{t-i} = the average paid claim cost in period $t-i$, and

F_{t-i} = the average claim frequency per exposure unit in period $t-i$.

* The Relationship expressed in equation (4) makes the usual assumption that the claim occurrence and the size of the claim are independent (CAS, 1990). That assumption may not be valid if, for example, organized claim fraud is present (Weisberg and Derrig, 1991). See also Lemaire, 1985.

The trend factor is obtained as follows:

$$\Delta_{t+j} = (1 + b_s)^j (1 + b_F)^j, \quad (7)$$

where b_s and b_F = estimated values of the slope coefficients in equations (5) and (6), and

j = the length of the forecast period (quarters).

The forecast period is the time between the midpoint of the data period used in calculating the indicated rate level change and the average accident date for the policies that will be written under the new rates. A seven- or eight-quarter forecast period is typical.

Prior Research

The actuarial approach to forecasting is rudimentary, and it would not seem difficult to obtain better forecasts using more advanced methods. This line of investigation has been pursued by Cummins and Powell (1980), Cummins and Harrington (1985), and Cummins and Griepentrog (1985), who compare the actuarial approach with econometric models and more sophisticated time trending methods. The studies used severity data—that is, countrywide average paid claim costs for private passenger automobile insurance—and produced the following general findings:

1. Econometric models generally work better than the simple time trend models.
 2. Broad-based indices forecast claim costs better than specific sectorial indices. The private sector wage rate performed well in predicting aggregate claim costs.
 3. Single-variable econometric models are better than multiple-variable models. Adding another price or wage index as an explanatory variable does not improve predictive accuracy.
 4. Simple time trend models (i.e., linear or exponential trending) usually perform better than complicated (ARIMA-type) trend models.
 5. Forecasts can often be improved by estimating the forecasting equation using a method that corrects for first-order autocorrelation of the residuals.
 6. Complicated econometric estimation adjustments generally do not improve forecast results.
 7. A priori reasoning is a valuable guide for specifying models and selecting variables, but the best methods can be identified only by testing.
- A comprehensive survey of forecasting by Armstrong (1978) suggests that most of these conclusions apply in general to the forecasting of many types of economic and social science phenomena.

How to Identify a Good Forecast

The objective of this study is to explore the decision-making process for identifying a good forecast. Conventionally, the forecast selection process is based on theoretical arguments regarding the statistical properties of estimators

and forecast errors or on empirical tests of predictive accuracy. The conventional approach aims to choose a single best forecasting method based on the results of the empirical tests. There are two problems with this approach. First, isolating a single method based on somewhat subjectively selected accuracy criteria disregards the information contained in alternative forecasting methods that perform nearly as well as the conventionally-selected best method. The "nearby" methods could have performed "best" based on different accuracy criteria or under different economic conditions. Second, the conventional approach focuses on the choice of a method, even though the true objective of the process is the selection of the trend factor and not the method itself.² By way of analogy, consider the selection of a stock for investment purposes. The investor's objective is to choose a stock that will perform well in the future, not to develop methods that explain past performance in some optimal way.

This article proposes an alternative approach to the trend factor decision problem, the use of fuzzy set theory. FST is the mathematical approach that explicitly models the vagueness of linguistic variables such as "accurate," "better," or "reliable" (Lemaire, 1990). Because FST provides a mechanism for simultaneously considering several forecasting methods and distinguishing among them by means of an explicit grading or ranking process, it allows the decision-maker to retain more of the information contained in the set of alternative forecasts. Unlike conventional methods, our FST approach allows for explicit consideration of the output of the forecasting methods—the trend factor—as part of the decision-making process. Thus, it potentially provides a better model of the actual decision process than the conventional approach.

Actual forecasting processes typically combine judgmental and statistical criteria and often involve a range of forecasting methods. Extensive support for the multiple forecast approach has been provided in the macroeconometrics literature (Clemen, 1989; Zarnowitz and Braun). Support for using judgmental criteria also is provided in the literature. For example, McNees (1990) observes that the best forecasts are made not by abandoning models or judgment but by blending both sources of information. He concludes that the forecast of a model, adjusted judgmentally to take into account information not incorporated in the model, tends to be more accurate than one generated mechanically. The forecasting problem in property-liability insurance is to find the appropriate blend of forecast technique and judgment. Below, technique and judgment are blended within the context of fuzzy set theory.

² A recent Massachusetts private passenger automobile insurance rate decision posed as the fundamental forecasting problem "to choose a methodology which will accurately predict loss costs for the next rate year" (Opinion, Findings and Decision on 1990 Automobile Insurance Rates, Massachusetts Division of Insurance, December 15, 1989, p. 37). This view of the forecasting problem—that the decision-maker chooses the trend factor, not the method—was upheld by the Massachusetts Supreme Judicial Court in *Aetna Casualty and Surety Company vs. Commissioner of Insurance*, 408 MASS 363 (1990). This judgmental criterion or linguistic variable, as Lemaire (1990) has dubbed such things, enlarges the problem beyond pure statistical techniques.

Our study conducts and then evaluates a set of 72 benchmark forecasting methods using private passenger auto insurance data from Massachusetts. The methods are compared using conventional criteria to search for a good forecast. Then, FST is used to consolidate the methods in terms of fuzzy criteria. The output of the FST process is a fuzzy set of "good" trend factors for use in setting prices beyond the sample period. Finally, for those who need a single forecast, we relate the fuzzy set of forecasts to the conventional use of weighted averages.

Claim Cost Forecasting Experiments

This section presents the results of claim cost forecasting experiments using Massachusetts private passenger automobile insurance data to develop "accurate" forecasts of pure premiums. Experiments were conducted using various time-trend and econometric techniques. The tests focus on both the predictive accuracy and the *bias* of the methods, where bias is defined as the tendency for a method to yield forecasts that are consistently higher or lower than realized claim costs.

Measurement of Forecast Accuracy

Forecast accuracy is gauged by estimating and then applying the trend factors over seven-quarter forecast periods to obtain predicted claim cost inflation and then comparing predicted inflation with actual claim cost inflation over the same forecast periods.³ The principal measure of accuracy used here is the total predicted change error (TPCE), where

$$\text{Trend Factor} = 1 + \Delta, \quad (8)$$

$$\text{Actual claim cost index} = 1 + \Delta_a = \frac{x_{t+7}}{x_t}, \quad \text{and} \quad (9)$$

$$\text{TPCE} = \frac{1 + \Delta}{1 + \Delta_a} - 1 = \frac{\Delta - \Delta_a}{1 + \Delta_a}, \quad (10)$$

where x_t = the value of the forecasted variable in quarter t . The TPCE provides a direct indication of the impact of trending errors on profit (see above).

The forecasting experiments are designed to replicate as closely as possible the results that would have been obtained under each forecasting method if that method had been used consistently in the past. Thus, the experiments assume that the only available data for each forecast period are those prior to the

³ A seven-quarter forecast period most closely approximates the period typically used in Massachusetts automobile insurance forecasting. The forecast period varies somewhat by state and over time. However, the median is about seven or eight quarters.

period. For example, forecasts for the seven-quarter forecast period 1981.1 through 1982.3 were based exclusively on data available prior to 1981.⁴

Forecasting experiments were conducted for automobile bodily injury liability (BIL) coverage. The maximum number of forecast experiments was conducted; that is, the number of experiments was limited only by the starting point of the pure premium or economic series used in the forecasts.⁵

The Forecasting Methods

The forecasting methods chosen for analysis represent a synthesis of current actuarial time trending techniques and the econometric methods proposed by, for example, Cummins and Griepentrog (1985). The goal is to span the set of reasonably simple methods identified as "good" in terms of current actuarial practice and prior performance. We have chosen not to follow the conventional approach of judgmentally discarding methods from this set a priori. Instead, we bring judgment into the process through the application of fuzzy methods to the entire set.

The time trend models rely exclusively on claim cost data and are very similar to the prevailing actuarial methods. Linear and exponential time trending equations are used. The exponential equations are presented as equations (5) and (6). The linear equation for frequency is:

$$F_{t-i} = \alpha_F + \beta_F(t-i) + \epsilon_{F,t-i}, \quad (11)$$

where $i = 0, 1, \dots, I-1$, and I is the number of periods used in estimating the equation. The severity equation is analogous.

The frequency and severity trend factor calculations are illustrated below for the case of exponential trending:⁶

$$\text{Frequency: } 1 + \Delta_F = \frac{\exp(a_F + (I+7)b_F)}{F_t} \quad (12)$$

$$\text{Severity: } 1 + \Delta_S = \frac{\exp(a_S + (I+7)b_S)}{S_t} \quad (13)$$

$$\text{Overall Pure Premium Trend: } 1 + \Delta = (1 + \Delta_F) * (1 + \Delta_S). \quad (14)$$

⁴ The notation 19xx.q, where xx = 77, 79, ..., and q = 1, 2, 3, 4 stands for quarter q of year 19xx.

⁵ Since the objective is to compare forecast methods, the forecasting begins at the earliest point where all methods can be compared. Thus, the first forecast experiment begins in 1981.1. This allows the comparison of forecasting equations estimated over eight- and twelve-quarter forecast periods. Beginning the forecasts earlier would not have allowed for the use of twelve-quarter estimation periods in the earlier experiments, because the Massachusetts quarterly claim cost data begin in 1978.1.

⁶ The trend factor formulas (12) through (14) differ somewhat from the actuarial trend factor formulas of equations (5) through (7) above. Equations (12) through (14) are used here because they represent the time trending approach employed in Massachusetts.

F_t and S_t represent frequency and severity in the last quarter of the estimation period.

The econometric forecasts differ from the time trend methods in using the external (economic) data series as the independent variable in the severity equation. Linear and log-linear specifications are used, consistent with the prior literature:

$$\text{Linear Severity: } S_{t-i} = \alpha_s + \beta_s C_{t-i} + \varepsilon_{s,t-i} \quad (15)$$

$$\text{Log-linear Severity: } \log(S_{t-i}) = \alpha_s + \beta_s \log(C_{t-i}) + \varepsilon_{s,t-i}, \quad (16)$$

where C_{t-i} = the value of an economic variable in period $t-i$, and $i = 0, 1, \dots, I-1$. The calculation of the severity trend factor for the log-linear econometric case is illustrated below:

$$1 + \Delta_s = \frac{\exp(a_s + b_s \log(\hat{C}_{t+7}))}{S_t}, \quad (17)$$

where $\hat{C}_{t+7}$ = forecasted value of C_{t+7} . The severity forecast factor is used in equation (14), along with the frequency time trend factor, to obtain the overall trend factor.⁷ Econometric models were not used in frequency forecasting because preliminary work suggested that the use of economic indices would not improve forecast performance.

The severity forecast for the econometric case requires an estimate of the external (economic) index for the last quarter of the forecast period (i.e., an estimate of C_{t+7}). As a benchmark, the forecasts were conducted using the actual values of C_{t+7} . Forecasted values for the external economic index were also obtained using time trending. Denoting the last quarter of the estimation period for equations (15) and (16) as period t , the forecasting equations for C_{t+7} were based on estimation periods ending in $t+3$, reflecting the more timely availability of the economic data series.⁸

Only one economic index was used in the experiments for each coverage—the Massachusetts composite auto insurance price index for the coverage under consideration. These indices have been developed by the Massachusetts State Rating Bureau and are supposed to represent the costs of goods and

⁷ Examination of equations (5) through (7) along with (12) through (14) reveal that the TPCE depends upon the last actual observation x_t only through its effect on the regression coefficients. This is equivalent to assuming that the difference between actual and fitted x_t is a random error which conveys no information that is useful in forecasting. This may not always be true, and there are various ways of bringing the regression errors into the forecast. One way is to use a first-order autoregressive adjustment as we do below. Other possibilities could be explored in future research.

⁸ Due to data collection and processing lags, Massachusetts claim cost data for any given quarter are available three time periods later than the economic data.

services covered by automobile insurance policies.⁹ The forecasts of the economic series span periods $t+4$ through $t+7$.

A forecasting method is characterized by an estimation period, an estimation technique, a frequency model, and a severity model. The forecast methods are summarized in Table 1. Section A of Table 1 summarizes the combinations of estimation periods, estimation techniques, and frequency models. Three estimation periods were used to obtain the forecasting equations: $I =$ eight quarters, twelve quarters, and all available data prior to the forecast period. Two estimation techniques were used: ordinary least squares (OLS) and least squares adjusted for first-order autocorrelation of the residuals (AR1).¹⁰ Linear and exponential time trends (equations [11] and [6]) were used as frequency models.

For each combination of estimation period, estimation technique, and frequency model, eight severity forecasts were used. These are summarized in section B of Table 1. Thus, a total of 96 forecasts (twelve combinations from section A of Table 1 times eight severity forecasts per combination) were conducted for each forecast period. Twenty-four of these were benchmark forecasts using actual rather than predicted values of the economic indices. The forecast comparisons are based on the 72 experiments that do not assume knowledge of any information during the forecast period.

Comparison of Trending Methods Using Conventional Criteria

The annualized inflation rate in Massachusetts automobile (BIL) pure premiums from 1978 through 1988 was 12.3 percent, whereas the inflation rate in the Consumer Price Index (CPI) during this period was only 6.2 percent. Thus, it would be erroneous to assume that Massachusetts BIL inflation is equal to general inflation. Trend factors are needed to project claim costs.¹¹

To test the accuracy of alternative claim cost forecasting methodologies, experiments were conducted using Massachusetts BIL claim cost data provided by the Automobile Insurers Bureau of Massachusetts.¹² The results of the experiments are reported below. Unless stated otherwise, all discussions of forecast accuracy refer to the pure premium series.¹³

⁹ The State Rating Bureau is a state agency responsible for representing the Commonwealth of Massachusetts in automobile insurance rate hearings. The insurance industry is represented in Massachusetts by the Automobile Insurers Bureau. The automobile insurance price indices are obtained as weighted averages of price and wage indices. A list of the component indices and the weights used in this study are available from the authors. The indices and weights are presented in Automobile Insurers Bureau (1989).

¹⁰ The Prais-Winsten AR1 transformation method is used (see Kmenta, 1986, p. 304).

¹¹ The Massachusetts composite index for BIL is closer to premium inflation than the CPI but still does not track claim cost inflation very well. Previous research has shown that the Masterson indices, published by the A. M. Best Company for many years (e.g., see Masterson, 1990) also are poor predictors of claim cost inflation (Cummins and Nye, 1981).

¹² The claim cost series were adjusted for seasonality using the X11 technique.

¹³ Past Massachusetts rate-making procedures incorporated modification factors applied to the data series. The modifications were designed to correct for structural changes in the transportation

Table 1
Summary of Forecasting Experiments

A. Pure Premium Experiments^a

Method Number	Estimation Period	Estimation Technique	Frequency Model
1	8 Quarters	OLS	Exponential
2	8 Quarters	OLS	Linear
3	12 Quarters	OLS	Exponential
4	12 Quarters	OLS	Linear
5	All Data	OLS	Exponential
6	All Data	OLS	Linear
7	8 Quarters	AR1	Exponential
8	8 Quarters	AR1	Linear
9	12 Quarters	AR1	Exponential
10	12 Quarters	AR1	Linear
11	All Data	AR1	Exponential
12	All Data	AR1	Linear

B. Severity Forecasts

Method Number	Severity Model	Economic Index Model	Column in Exhibit
1	$S(t) = A + B \cdot t$	None	(1)
2	$LN(S(t)) = A + B \cdot t$	None	(2)
3	$S(t) = A + B \cdot X(t)$	Actual $X(t)$	(3)
4	$LN(S(t)) = A + B \cdot LN(X(t))$	Actual $X(t)$	(4)
5	$S(t) = A + B \cdot X(t)$	$X(t) = a + b \cdot t$	(5)
6	$LN(S(t)) = A + B \cdot LN(X(t))$	$X(t) = a + b \cdot t$	(6)
7	$S(t) = A + B \cdot X(t)$	$LN(X(t)) = a + b \cdot t$	(7)
8	$LN(S(t)) = A + B \cdot LN(X(t))$	$LN(X(t)) = a + b \cdot t$	(8)

Note: $S(t)$ = severity in period t , t = time, $X(t)$ = value of economic index in period t . OLS = ordinary least squares, AR1 = least squares adjusted for first-order autocorrelation in the residuals.

^a Each forecast method is characterized by an estimation period (8 quarters, 12 quarters, or all data), an estimation technique (OLS or AR1), and a frequency model (exponential or linear). Each of the 12 methods is combined with 8 severity estimates (see section B) to obtain pure premium trend factors.

As indicated above, the total predicted change error is the primary measure of forecast performance used in this study. Both average absolute and average TPCEs are reported. Earlier studies used absolute TPCEs as the principal measure of overall forecast accuracy. In real-world applications, average errors are equally important. For example, insurers will experience problems if they use methods with low absolute errors that consistently under-predict claim

environment such as seatbelt use, brake lights, important legal decisions, and anti-drunk driving statutes. The modifications are in the form of proportionate multipliers, generally slightly higher or lower than 1.0. In rare instances (e.g., the Massachusetts Auto Insurance Reform Act of 1988), the modification factors can deviate sharply from 1.0. The minor modifications were highly subjective and rarely based on empirical evidence. Forecasting results incorporating minor modification factors are not reported here because of the subjective nature of the modifications and because preliminary work indicated that they do not consistently improve the forecasts.

costs. However, average errors must be interpreted with caution because a low average error can mask large offsetting positive and negative errors. As explained below, FST provides a mathematically consistent way to capture information provided by both performance measures.

The BIL forecast errors are summarized in Tables 2 and 3. Table 2 presents absolute average total predicted change errors, while Table 3 presents average TPCEs. The tables report the results of seven-quarter forecasts conducted beginning in 1981.1 and in every subsequent quarter up to 1987.2. The last forecast was for the last available data point, 1988.4. A total of 26 separate forecasts was conducted for every method reported in Tables 2 and 3.¹⁴ The number presented in each cell of Table 2 (3) is the average of the absolute (non-absolute) error for the 26 forecasts using the method combination reported in the cell. The rows in Tables 2 and 3 correspond to the twelve combinations of estimation techniques, estimation periods, and frequency models presented in section A of Table 1, while the columns correspond to the severity models specified in section B of Table 1. Methods often are referred to by their row and column numbers; for example, method 5(3) refers to the results reported in row 5, column 3 of Table 2 or 3.

The forecast errors reported in Table 2 for bodily injury liability pure premiums range from slightly less than 4 to about 8 percent. Most of the methods have errors in the range from about 6 to 8 percent. The methods which use all available data to fit the forecasting equations (the all-data methods) clearly stand out as the most accurate in terms of the absolute average TPCE criterion.¹⁵ The average TPCEs for BIL are reported in Table 3. Most of the errors are quite low (the largest in absolute terms is 4.3 percent), but the majority tend to be biased toward under-predicting claim costs. Considering both average (Table 3) and absolute average (Table 2) errors, the best methods seem to be 6(7), 6(8), 11(7), and 11(8). The average absolute TPCEs for these methods are among the lowest in Table 2, and their average TPCEs are virtually indistinguishable from zero.

Additional information on forecast accuracy is presented in Table 4, which provides average and absolute average errors for frequency, severity, and pure premiums for two of the conventionally chosen "best" methods, 6(7) and 6(8). The forecast results illustrate a phenomenon that might be termed "forecast pooling." The methods summarized in the exhibit obtain relatively low overall average errors even though they consistently over-predict severity and under-predict frequency, because the frequency and severity errors tend to be

¹⁴ The first estimation period was based on data prior to 1981.1. The first forecast was for the period 1981.1 through 1982.3, and the first TPCE was computed for the end of this forecast period, 1982.3. The next estimation period was based on data prior to 1981.2, and the next forecast covered the period 1981.2 through 1982.4. The TPCE for this experiment measured the accuracy in predicting the pure premium for 1982.4. The last forecast experiment used an estimation period based on data prior to 1987.2 and a forecast period of 1987.2 through 1988.4. The last TPCE was for 1988.4. This gives a total of 26 forecasts for each method.

¹⁵ These are methods 5, 6, 11, and 12. See Table 1 for definitions of the forecasting methods.

Table 2
Bodily Injury Liability Pure Premiums:
Absolute Average Total Predicted Change Error Forecast Errors (Percent)

Method	Time Trend		Economic Predictions		Linear-Predicted		Exponential Predicted	
	Linear	Exponential	Actual Linear	Actual LN-Linear	Linear	LN-Linear	Linear	LN-Linear
1 8 Q, OLS, EXP FREQ	6.92	7.60	6.90	7.02	7.36	7.49	7.65	7.78
2 8 Q, OLS, LIN FREQ	6.39	6.98	6.21	6.33	6.71	6.85	6.95	7.09
3 12 Q, OLS, EXP FREQ	5.66	6.45	6.38	6.37	6.32	6.37	6.79	6.84
4 12 Q, OLS, LIN FREQ	5.19	5.82	5.74	5.72	5.76	5.84	6.13	6.21
5 All Q, ^a OLS, EXP FREQ	4.31	4.24	4.73	4.76	4.65	4.73	4.34	4.36
6 All Q, ^a OLS, LIN FREQ	4.90	3.96	5.33	5.37	5.26	5.36	4.21	4.28
7 8 Q, AR1, EXP FREQ	6.87	7.61	6.91	7.02	7.36	7.49	7.65	7.79
8 8 Q, AR1, LIN FREQ	6.36	6.99	6.21	6.35	6.71	6.86	6.96	7.10
9 12 Q, AR1, EXP FREQ	5.54	6.17	6.31	6.30	6.15	6.21	6.59	6.64
10 12 Q, AR1, LIN FREQ	5.09	5.60	5.71	5.71	5.63	5.73	5.97	6.05
11 All Q, ^a AR1, EXP FREQ	4.19	4.45	4.53	4.53	4.61	4.67	4.45	4.50
12 All Q, ^a AR1, LIN FREQ	4.61	4.06	4.96	4.96	5.15	5.23	4.59	4.67

Note: Method numbers are defined in more detail in Exhibit 1. Method 1 uses an 8-quarter (8Q) estimation period and ordinary least squares (OLS) as the estimation method for frequency, severity, and economic index prediction. Method 1 uses exponential time trending (EXP FREQ) to predict frequency, while method 2 uses linear time trending (LIN FREQ) for frequency. Column labels refer to severity forecast methods. Time Trend = severity forecasted by time trend (Linear or Exponential). For the Economic Predictions, the columns labeled Actual use actual economic index values to predict severity in either a Linear or LN-Linear severity forecast equation. The columns labeled Linear-Predicted use a linear time trend to forecast the economic index, with this prediction used in either a Linear or LN-Linear severity forecast equation. The last two columns in the exhibit (Exponential-Predicted) use an exponential time trend to predict the economic index, with this forecast again used in a Linear or LN-Linear forecast equation for severity. Frequency and severity forecasts are multiplied to obtain the pure premium forecast.

^a All Q = all available data prior to forecast period used in estimating forecast equation.

Table 3
Bodily Injury Liability Pure Premiums:
Average Total Predicted Change Error Forecast Errors (Percent)

Method	Time Trend		Economic Predictions		Linear-Predicted		Exponential Predicted	
	Linear	Exponential	Actual	LN-Linear	Linear	LN-Linear	Linear	LN-Linear
1 8 Q, OLS, EXP FREQ	-0.25	1.02	-0.51	-0.59	-0.39	-0.46	0.21	0.14
2 8 Q, OLS, LIN FREQ	-1.23	0.04	-1.48	-1.56	-1.37	-1.43	-0.77	-0.83
3 12 Q, OLS, EXP FREQ	-0.28	1.50	-0.48	-0.67	-0.30	-0.47	0.73	0.55
4 12 Q, OLS, LIN FREQ	-1.45	0.31	-1.67	-1.85	-1.48	-1.64	-0.46	-0.64
5 All Q, OLS, EXP FREQ	-2.61	2.46	-3.33	-3.24	-2.76	-2.66	1.20	1.24
6 All Q, OLS, LIN FREQ	-3.58	1.44	-4.30	-4.21	-3.73	-3.62	0.19	0.23
7 8 Q, ARI, EXP FREQ	-0.32	0.96	-0.54	-0.61	-0.41	-0.47	0.19	0.13
8 8 Q, ARI, LIN FREQ	-1.28	-0.01	-1.50	-1.58	-1.38	-1.44	-0.78	-0.84
9 12 Q, ARI, EXP FREQ	-0.33	1.42	-0.53	-0.71	-0.34	-0.51	0.65	0.46
10 12 Q, ARI, LIN FREQ	-1.47	0.26	-1.68	-1.86	-1.49	-1.66	-0.51	-0.69
11 All Q, ARI, EXP FREQ	-2.33	2.92	-3.11	-2.98	-2.82	-2.68	-0.24	-0.12
12 All Q, ARI, LIN FREQ	-3.21	1.99	-3.99	-3.86	-3.70	-3.55	-1.14	-1.01

Note: Method numbers are defined in more detail in Exhibit 1. Method 1 uses an 8-quarter (8Q) estimation period and ordinary least squares (OLS) as the estimation method for frequency, severity, and economic index prediction. Method 1 uses exponential time trending (EXP FREQ) to predict frequency, while method 2 uses linear time trending (LIN FREQ) for frequency. Column labels refer to severity forecast methods. Time Trend = severity forecasted by time trend (Linear or Exponential). For the Economic Predictions, the columns labeled Actual use actual economic index values to predict severity in either a Linear or LN-Linear severity forecast equation. The columns labeled Linear-Predicted use a linear time trend to forecast the economic index, with this prediction used in either a Linear or LN-Linear severity forecast equation. The last two columns in the exhibit (Exponential-Predicted) use an exponential time trend to predict the economic index, with this forecast again used in a Linear or LN-Linear forecast equation for severity. Frequency and severity forecasts are multiplied to obtain the pure premium forecast.

Table 4
 Total Predicted Change Errors (Percent)
 "Best" Methods by Frequency, Severity, and Pure Premiums

<i>Method</i>	<i>Number of Forecasts</i>	<i>Frequency</i>	<i>Severity</i>	<i>Pure Premiums</i>
Best Methods 6(7) and 6(8):		Absolute Averages:		
With Linear Severity	26	7.90	8.75	4.21
With Log-Linear Severity	26	7.90	8.78	4.28
Best Methods 6(7) and 6(8):		Averages:		
With Linear Severity	26	-7.90	8.75	0.19
With Log-Linear Severity	26	-7.90	8.78	0.23

Note: "Best" method uses all available data, linear frequency, and an exponential time trend to predict the economic index.

offsetting. This suggests that relative forecast errors may be reduced by pooling across forecasts (see Armstrong, 1978, and Granger and Newbold, 1986), a technique that is worth exploring in future work. Table 4 provides additional evidence that it is advisable to consider both average and absolute average errors.

Trend Factors for 1989.1-1990.3

The forecasting models were used to estimate trend factors for the period 1989.1-1990.3 on the assumption that the last available data point is 1988.4. The trend factors are presented in the Appendix. The average pure premium factor for all benchmark methods is 13.7 percent. The 72 individual trend factors, however, show wide variations, even among the "best" methods. Obviously, the wide range of trend factors produced by the alternative forecasting methods creates uncertainty with regard to the choice of a forecast. The next step in the traditional analysis would be to apply statistical methods to eliminate inaccurate forecasts and combine the remaining forecasts to reduce forecast error variance. Such methods are available in the statistics and econometrics literature (e.g., Granger and Newbold, 1986), and their application represents a promising direction for future research. The following sections, however, illustrate the evaluation of the forecasting methods using fuzzy set theory.

Elements of Fuzzy Set Theory

In order to describe the fuzzy set theory context for the pure premium forecasting problem, some basic definitions are needed. We follow Lemaire (1990) in describing some fundamentals of FST. Examples and applications focus on the forecasting problem. See also Kandel (1986) and Zimmerman (1991).

Basics of FST

The fundamental concept of FST is the alternative formalization of membership in a set to include the degree or strength of membership. Let $X = \{x\}$ be an ordinary set of objects. For example, X might be the set of all

feasible trend factors, a subset of the real numbers. A fuzzy set A in X is a set of ordered pairs

$$A = \{ x, U_A(x) \}, \quad x \in X, \quad (18)$$

where U_A maps X to the interval $[0, 1]$. The value $U_A(x)$ is called the grade or value of membership of x in the fuzzy set A with 0 and 1 representing, respectively, the lowest and highest grades (values) of membership. When $\{U_A(x)\}$ only contains the two points 0 and 1, A is non-fuzzy.

EXAMPLE 1: A "GOOD" BI FORECASTING METHOD. Let X be the set of 72 BIL forecasting methods whose average absolute TPCEs are shown in Table 2, that is, X is the set of models considered in this study.¹⁶ Let us assume that a "good" trend methodology is one that predicted well enough in the past (according to the TPCE values shown in Table 2) to be "near" the "best" method according to the Table 2 test results; namely, the method that uses linear OLS time trends with all data for frequency and an exponential time trend for severity. That method, denoted method 6(2) for row 6 column 2 in Table 2, is chosen as the "best" method according to this criterion because it has the lowest average absolute TPCE.

We define the membership function for X by

$$U_{g_1}(X) = \frac{TPCE[6(2)]}{TPCE[X]} = \frac{3.96}{TPCE[X]}, \quad (19)$$

where $U_{g_1}(X)$ = the membership function for the fuzzy set of forecasting methods that are "good" in terms of historical forecast accuracy. Then, for example, $U_{g_1}[6(2)] = 1$. Hence 6(2) is definitely a "good" method. On the other hand, for two other methods we have

A. All Data, OLS, Exponential Frequency, Exponential Severity:

$$U_{g_1}[5(2)] = 3.96/4.24 = 0.934 \text{ and}$$

B. Eight-Quarter, OLS, Linear Frequency, Exponential Severity:

$$U_{g_1}[2(2)] = 3.96/6.98 = 0.567.$$

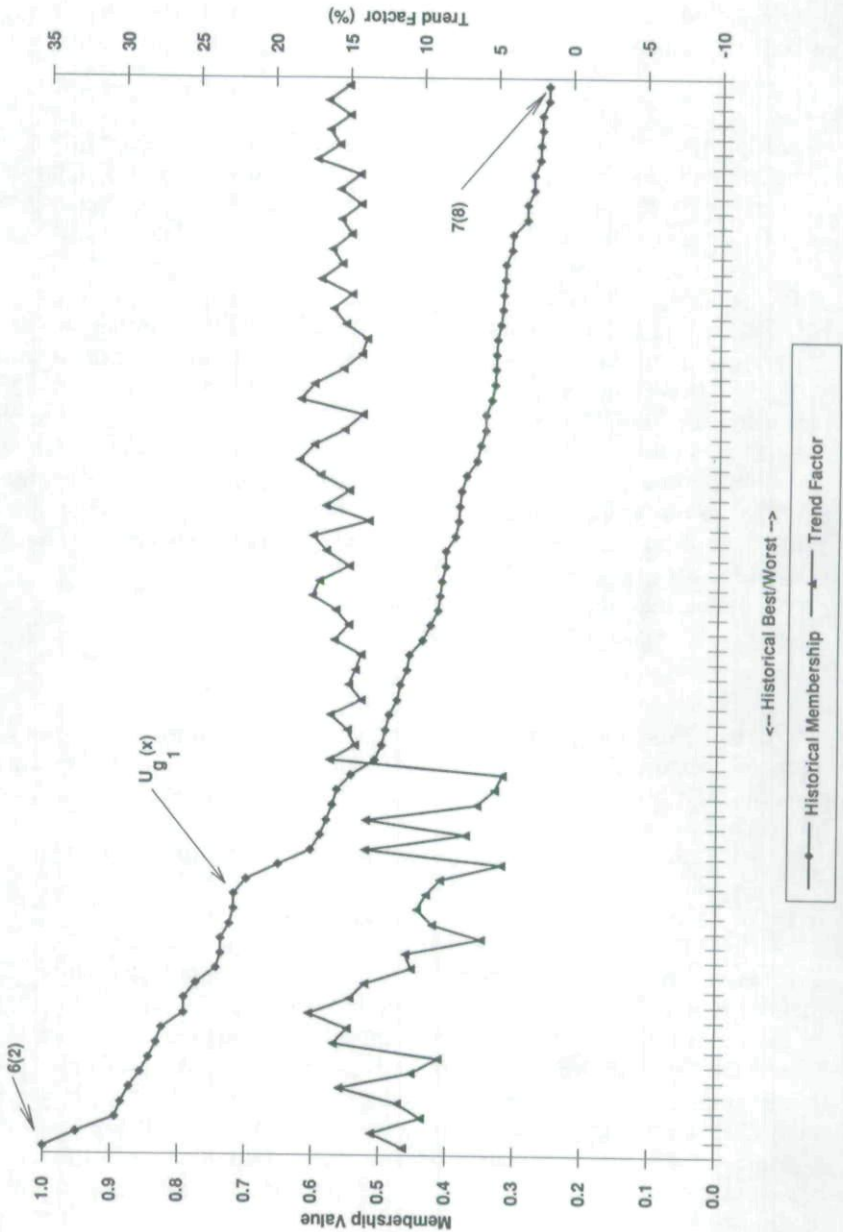
Liberalized interpreted, Method A is quite good, although not the best, while Method B is little more than half as good as the best. Figure 1 illustrates the membership function of the set of good BIL forecast methods from the best 6(2) to the "worst" 7(8), the latter with a membership value of only 0.508. In this example, the fuzzy set is the entire set of objects (the 72 benchmark methods). If we enlarged the set of objects considered to all forecast methods in some suitable space, the membership function above would "fuzzify" the benchmark forecasts by setting the membership function (19) for all other methods equal to zero.

A fuzzy set is said to be *normal* if $\sup_x U_A(x) = 1$. *Subnormal* fuzzy sets can be normalized by dividing each $U_A(x)$ by $\sup_x U_A(x)$.¹⁷

¹⁶ The forecast methods shown in columns 3 and 4 that use actual values of the independent variable are not used in subsequent illustrations.

¹⁷ Of course, the Sup operator is the supremum, or least upper bound, applied to the set of membership values $\{U_A(x)\}$.

Figure 1
1990.3 Trend Factors for Bodily Injury Liability Trend Methods



The membership function for a fuzzy set A is used to define the *complement* of A so that the grade of membership in A is formally related to the grade of membership in the complement. Specifically, A_c is said to be the complement of A if and only if

$$U_{A_c}(x) = 1 - U_A(x), \text{ for all } x. \quad (20)$$

Unlike the ordinary set theoretic definition, the fuzzy set complement allows objects x to be members, albeit gray-shaded members, of both A and its complement A_c . In our forecast example, the literal translation for membership in the complement of good methods might be: "This method is not so good but it's not so bad either."

In many cases, it is useful to consider the membership of a set of objects in more than one fuzzy set. Without loss of generality, consider the membership of the ordinary set of objects $X = \{x\}$ in the fuzzy sets A and B . Several fuzzy set concepts are needed to study membership in more than one set. For example, a fuzzy set A is *contained in* or is a *subset of* fuzzy set B ($A \subset B$) if and only if

$$U_A(x) \leq U_B(x), \text{ for all } x. \quad (21)$$

The *union* of A and B , denoted $A \cup B$, is defined as the smallest fuzzy set containing both A and B . Its membership function is given by

$$U_{A \cup B}(x) = \max [U_A(x), U_B(x)], x \in X. \quad (22)$$

The *intersection* of A and B , denoted $A \cap B$, is defined as the largest fuzzy set contained in both A and B . Its membership function is given by

$$U_{A \cap B}(x) = \min [U_A(x), U_B(x)], x \in X. \quad (23)$$

It can be shown that these definitions of fuzzy union and intersection are the only ones that naturally extend the corresponding standard set theory notions by satisfying all the usual requirements of associativity, commutativity, idempotency, and distributivity (Lemaire, 1990). They are also the only operators with other desirable properties such as that "membership in $A \cap B$ requires more, and membership in $A \cup B$ requires less, than the membership in one of A or B " (Dubois and Prade, 1980, p. 12).

Other Definitions of the Intersection

Other definitions of the intersection have been suggested; they correspond to more flexible interpretations of common membership in two fuzzy sets. The selection of a specific operator for intersection, other than the minimum operator, will depend on its possibilities to allow for *cumulative effects*, *interactions*, and *compensations* among criteria. These properties are desirable for "softer" versions of intersection operators.

Property 1 (cumulative effects): Two infringements are worse than one.

$$U_{A \cap B}(x) < \min [U_A(x), U_B(x)], \text{ if } U_A(x) < 1 \text{ and } U_B(x) < 1. \quad (24)$$

Property 2 (interactions between criteria): Assume $U_A(x) < U_B(x) < 1$. Then the effects of a decrease of $U_A(x)$ on $U_{A \cap B}(x)$ may depend on $U_B(x)$.

Property 3 (compensations between criteria): If $U_A(x)$ and $U_B(x) < 1$, the effect of a decrease of $U_A(x)$ on $U_{A \cap B}(x)$ can be reduced by an increase of $U_B(x)$ (unless, of course, $U_B(x)$ reaches 1).

The reason for seeking alternative definitions of the intersection is that the standard FST intersection operator (the minimum operator, equation [23]) does not allow for any of the three properties even though it is desirable to allow for them in some applications. Lemaire lists four alternative definitions and extensions of the standard fuzzy intersection definition (equation [23]). Each one provides for systematic ways to introduce linguistic decision variables into FST. These are the *algebraic product*, the *bounded difference*, the *Hamacher operator*, and the *Yager operator*.

The *algebraic product* of A and B is denoted AB and is defined by

$$U_{AB} = U_A(x) U_B(x). \quad (25)$$

The *bounded difference* of A and B is denoted $A \ominus B$ and is defined by

$$U_{A \ominus B}(x) = \max [0, U_A(x) + U_B(x) - 1]. \quad (26)$$

The *Hamacher* and *Yager operators*, H and Y, define the intersection of two fuzzy sets A and B by

$$U_H^p(X) = \frac{U_A(x) U_B(x)}{p + (1-p)[U_A(x) + U_B(x) - U_A(x)U_B(x)]}, \quad 0 \leq p \leq 1 \text{ and} \quad (27)$$

$$U_Y^p(x) = 1 - \min \left[1, [(1 - U_A(x))^p + (1 - U_B(x))^p]^{\frac{1}{p}} \right], \quad p \geq 1. \quad (28)$$

Fuzzy intersection operators differ in the degree to which they satisfy properties 1 through 3. The bounded difference satisfies only properties 1 and 3, thereby excluding interactions. The algebraic product satisfies all three and provides for maximum interactions. The Hamacher and Yager operators allow for all three properties but temper the effects somewhat in comparison with the algebraic product. The degree of tempering is reflected by the parameter p. A more complete discussion of operators and criteria for selecting among them is presented in Zimmerman (1991).

The α -Cut Concept

Finally, we will find useful the notion of the α -cut. If A is a fuzzy subset of X, its α -cut A_α is defined as the non-fuzzy subset such that

$$A_\alpha = \{x \mid U_A(x) \geq \alpha\}, \text{ for } 0 \leq \alpha \leq 1. \quad (29)$$

EXAMPLE 2: REASONABLE TREND FACTORS. Traditional choices of "best" forecasting methods have relied on choosing the method with the optimal value under some pre-defined test criterion or criteria. In the claim cost forecasting experiments discussed above, the minimum absolute TCPE criterion would produce method 6(2) as the "best." The judgmental or "subjective intersection" of the absolute average and average TPCE criteria leads to the selection of the four "good" methods 6(7), 6(8), 11(7), and 11(8), none of which was the "best" by any single-criterion.

Other methods that were almost as good as the conventionally-selected "best" choices would be discarded by optimizers, but not necessarily by actuaries or intervalists. FST membership functions allow for "good" or "reasonable" methods, as well as the traditional best method; α -cut models discard values that are extreme in some sense, that is, beyond the minimum grade of α for membership.

The Appendix shows the 72 benchmark trend factors for the trending period 1989.1-1990.3. Let us assume that a fuzzy subset R of X , the set of "reasonable" trend factors, is defined by the membership function:

$$U_R(x) = \begin{cases} 1 - \frac{|\Delta_x - \bar{\Delta}|}{2\sigma}, & \text{if } |\Delta_x - \bar{\Delta}| \leq 2\sigma \\ 0, & \text{otherwise} \end{cases} \quad (30)$$

where Δ_x = the trend factor produced by method x ,

$\bar{\Delta}$ = the average trend factor over all methods,

σ = the standard deviation of the trend factors, and

$U_R(x)$ = the membership function of the fuzzy set of "reasonable" trend factors.

We constructed equation (30) to reflect our lack of confidence in extreme forecasts, and our objective of choosing a "good" trend factor rather than the historically "best" method(s). Accordingly, equation (30) is independent of the historical accuracy of the methods under consideration. It has been our experience that real-world actuaries and regulators invariably give less credibility to extreme results when confronted with a range of alternatives. Empirical support for this approach is provided by the forecasting literature, which shows that average forecasts tend to outperform individual forecasts and that past performance is not an infallible guide to future accuracy (e.g., Zarnowitz and Braun, 1992; McNees, 1992).

The α -cuts defined by equation (30) correspond to discarding those trend factors that are "extreme" in the sense that the trend factors not in the α -cut are beyond the grand mean by some multiple of the standard deviation. For instance, based on equation (30), we have

α	α -cut
0.75	within $\frac{1}{2} \sigma$
0.50	within 1σ
0.25	within $1\frac{1}{2} \sigma$
0	within 2σ

We now use these FST concepts to reinforce Lemaire's expectation that FST has a wide range of application to actuarial problems by exploring FST as a decision-making method for forecasting.

Evaluation of Trends Using Fuzzy Set Theory

Fuzzy set theoretic concepts can be used to reformulate the forecasting decision process. For example, the decision-maker may specify one or more goals $\{g\}$ to be attained. Provided that fuzzy membership functions can be created to measure the degree to which each goal is met, the decision problem can be recast within FST.¹⁸

Straightforward intersection of the goal fuzzy sets will yield a fuzzy solution set. More importantly, however, the calculation of the fuzzy intersection ranks the alternatives with a degree of membership in the final fuzzy set of alternatives. Other fuzzy intersection operators also can be used to reformulate the problem using alternative criteria. Each intersection operation yields a fuzzy set of alternatives. The "answer" to the decision problem is the fuzzy intersection-set rather than a specific forecast or method. However, this fuzzy set can be used to select a forecast, as shown below.

The FST decision-making procedure can be illustrated using the BIL trend factors and historical measures of accuracy presented above. In order to keep the arithmetic simple, we initially use only the two "best" methods shown in Table 4 plus the method with the lowest absolute average TPCE. Our decision criteria are based on the following goals:

Goal 1: *Historical Accuracy*—The forecasting method should be reasonably good for past predictions as measured by the average absolute total percentage change error.

Goal 2: *Unbiasedness*—The forecasting method should be reasonably unbiased as measured by the average total percentage charge error in the historical forecasts.¹⁹

¹⁸ Fuzzy optimization can involve constraints as well as goals. The distinction is somewhat semantic because they can be treated identically in the decision process (see Lemaire, 1990, and Dubois and Prade, 1980, pp. 242-244).

¹⁹ Although actuaries and economists take as given that the forecast methods used in estimating trends in claim costs should be unbiased in the statistical sense, that is not always the case in a regulated environment. For example, Derrig's (1993) review of Massachusetts automobile profitability demonstrates that regulatory decisions in that state have consistently underestimated claim cost inflation.

Goal 3: *Reasonableness*—The trend factor selected should be reasonably close to the average benchmark forecasts; that is, extreme choices should be avoided.

The goal 1 concept of “reasonably good for past predictions” can be made operational using the Table 2 test results and the membership function $U_{g1}(X)$, defined in equation (19). Application of equation (19) yields the fuzzy set of historically good forecast methods illustrated in Figure 1.

The goal 2 concept can be made operational in an analogous manner by formulating a membership function. Strictly unbiased forecast methods as demonstrated by past performance are the “best” under goal 2. The following membership function operationalizes this goal:

$$U_{g2}(x) = e^{-|100 \text{ Average}(TPCE_x)|}, \quad (31)$$

where $\text{Average}(TPCE_x)$ = the average TPCE for method x , and $U_{g2}(x)$ = the membership function of the fuzzy set of forecasts that are “good” according to the historical unbiasedness criterion (goal 2). Equation (31) also would be an appropriate membership function if the average error were used as an alternative measure of accuracy rather than an indicator of bias.

Finally, the goal 3 concept of a reasonable (or “moderate,” “not too extreme,” “realistic,” etc.) method can be made operational by applying equation (30) using the mean and standard deviation of the 72 trend factors shown in the Appendix. The use of deviations from the mean forecast to define “reasonableness” is justified by the growing evidence that average forecasts outperform individual forecasts (Zarnowitz and Braun, 1992). An alternative would be to use deviations from the mean of some α -cut of the intersection-set of goals 1 and 2.²⁰

As an example of the FST decision-making process, we use two of the conventionally-chosen “best” methods and, for contrast, the method with the lowest TPCE in Table 2:

Method A: All data, OLS, linear frequency, exponential economic index, linear severity: 6(7).

Method B: All data, OLS, linear frequency, exponential economic index, log-linear severity: 6(8).

Method C: All data, OLS, linear frequency, exponential time trend severity: 6(2).

Applying the FST definitions associated with our three goals, we have the following decision array.

²⁰ An interesting avenue for future research would be to test the accuracy of trend factors derived using various definitions of reasonableness. This would be an advance over existing research on combining forecasts, which uses straight averages.

Fuzzy Set	Grade of Membership		
	Method A	Method B	Method C
Goal 1	0.941	0.925	1.000
Goal 2	0.827	0.795	0.237
Goal 3	0.634	0.503	0.577

An examination of the membership grades for all 72 methods led to a reevaluation of goal 2. This goal turned out to be much more rigorous than the other two decision criteria. For example, the average membership grade among the 72 methods according to goal 2 is only 0.43, whereas the average for goal 1 is 0.70. In addition, methods with relatively small average errors (e.g., less than 2 percent) were given very low scores by the goal 2 membership function (equation [31]). For example, method C had an average error of 1.44 percent and a membership function value of 0.237 for goal 2. Using goal 2 in its unmodified form would give much more weight to bias than to the other criteria.

Fortunately, fuzzy set theory provides methods for dealing with situations such as the assigning of unusually high or low weights by a membership function. One obvious approach would be to change the function. However, in some instances, the nature of the problem dictates the choice of a particular function. Although this is not the case in our application, we will take advantage of the relative stringency of goal 2 to illustrate two additional operations permitted by fuzzy set theory.

The operations are part of a class "that have no counterparts in ordinary set theory; they are uniquely fuzzy" (Lemaire, 1990, p. 44). One operation used here, *dilation*, is defined as

$$U^{DIL}(x) = [U(x)]^a, \quad a < 1. \quad (32)$$

Dilation stretches the fuzzy set by increasing the grade of membership of all elements in the set with membership function values greater than 0 and less than 1. The parameter a is called the *power* of the operation. A related operation is *concentration*, defined as

$$U^{CON}(x) = [U(x)]^a, \quad a > 1. \quad (33)$$

To align the stringency of goals 1 and 2 and illustrate dilation and concentration, the membership function for goal 1 was concentrated by raising equation (19) to the power +2, while goal 2 was dilated by raising equation (31) to the power 0.75. The concentration operation lowered the average membership function value for goal 1 from 0.696 to 0.504, while dilation increased the average goal 2 membership values for the 72 forecast methods from 0.432 to 0.506. The goal 1 membership values for methods A, B, and C change from 0.941, 0.925, and 1 to 0.885, 0.856, and 1; the goal 2 membership function values change from 0.827, 0.795, and 0.237 to 0.867, 0.842, and 0.340, respectively. More detailed comparisons reveal that the revised membership functions for goals 1 and 2 are much more consistent in terms of

stringency than the original membership functions. Consequently, the revised membership functions were adopted for the remainder of the analysis.

The selected membership functions illustrate the solution of the decision problem. First, consider the three methods A, B, and C. The problem is solved by taking the intersection of the three membership functions. Applying the ordinary intersection operator of equation (23) plus the other four intersection operators from the previous section yields the results presented in Table 5. Methods A and B are clearly better than method C, the best method under criterion 1. Conventional decision-makers might be inclined to decide that A is dominant and to discard the other two methods. In fuzzy decision making, however, the "answer" for any given intersection operator is the fuzzy set consisting of A, B, and C and their membership function values. We would like to retain the information provided by methods A, B, and C in proportion to their degrees of membership in the fuzzy set of "good" forecasts. A method for doing this is proposed below.

Table 5
Comparison of Forecast Methods Using Alternative Intersection Operators

Intersection	Grade of Membership		
	Method A: 6(7)	Method B: 6(8)	Method C: 6(2)
Ordinary	0.634	0.503	0.340
Algebraic Product	0.486	0.363	0.200
Bounded Difference	0.386	0.201	0.000
Hamacher ($p = 0.5$)	0.511	0.393	0.228
Yager ($p = 2$)	0.594	0.459	0.216

Returning to the set of 72 forecast methods, Figure 2 shows the fuzzy membership values for all 72 forecasts according to goals 1 and 2. In general, the historically best methods are among the least unbiased and, conversely, the most unbiased methods are the least accurate historically. This illustrates the importance of looking at multiple criteria. An insurer choosing a forecast method with the lowest TPCE will not be successful if this method consistently underestimates claim costs. In ordinary decision making, it is difficult to combine information on the "goodness" of forecasts when multiple criteria are used. The intersection operators of FST provide a natural way to solve this problem. This well-established procedure is known as fuzzy averaging (Zimmerman, 1991).

Figure 3 shows the ordinary intersection of goal 1 and goal 2 fuzzy sets for the entire sample of 72 forecast methods. Notice that the fourth, sixth, eleventh, and twelfth best criteria 1 methods are shown in Figure 3 to be the better methods using both criteria simultaneously. Those four methods are 6(7), 6(8), 11(7), and 11(8), respectively. These methods also would be identified by conventional methods as the best of the group. However, unlike FST, conventional methods would not provide information on how "good" these

Figure 2
Two Fuzzy Sets for Good Trend Methods

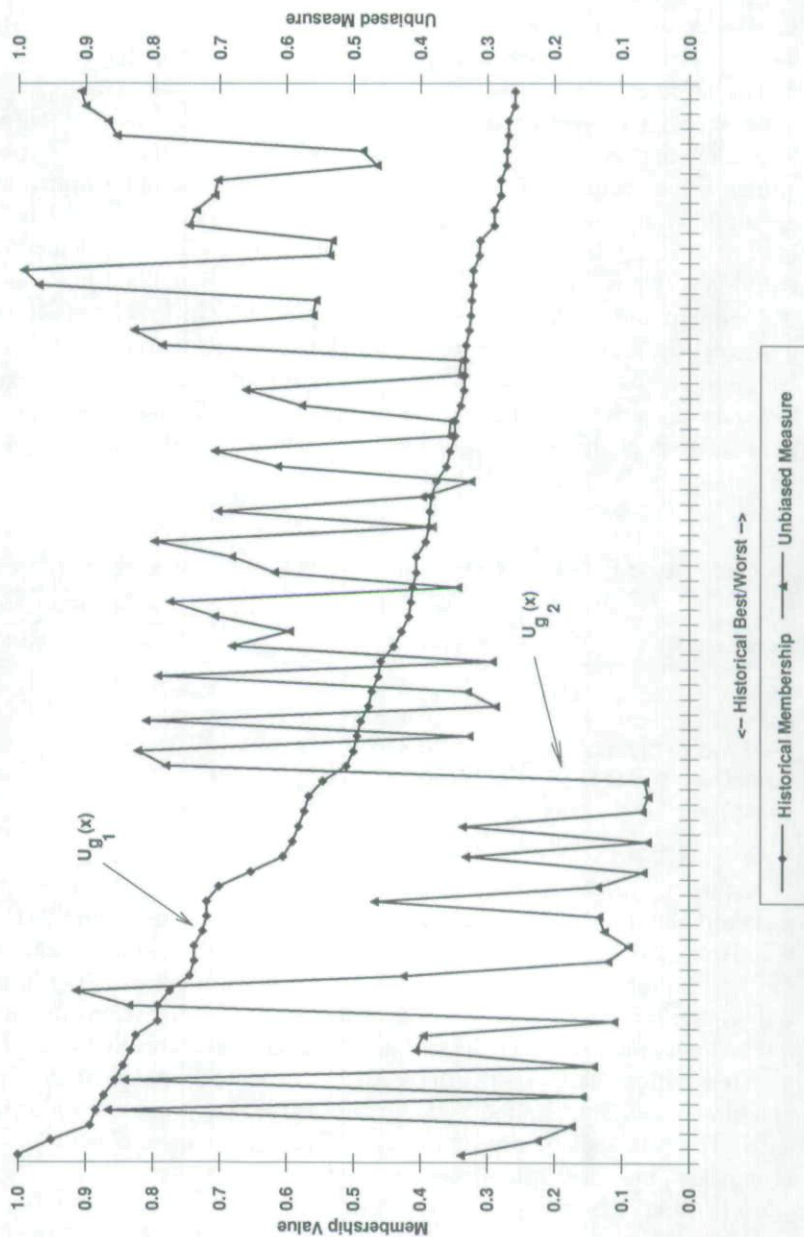
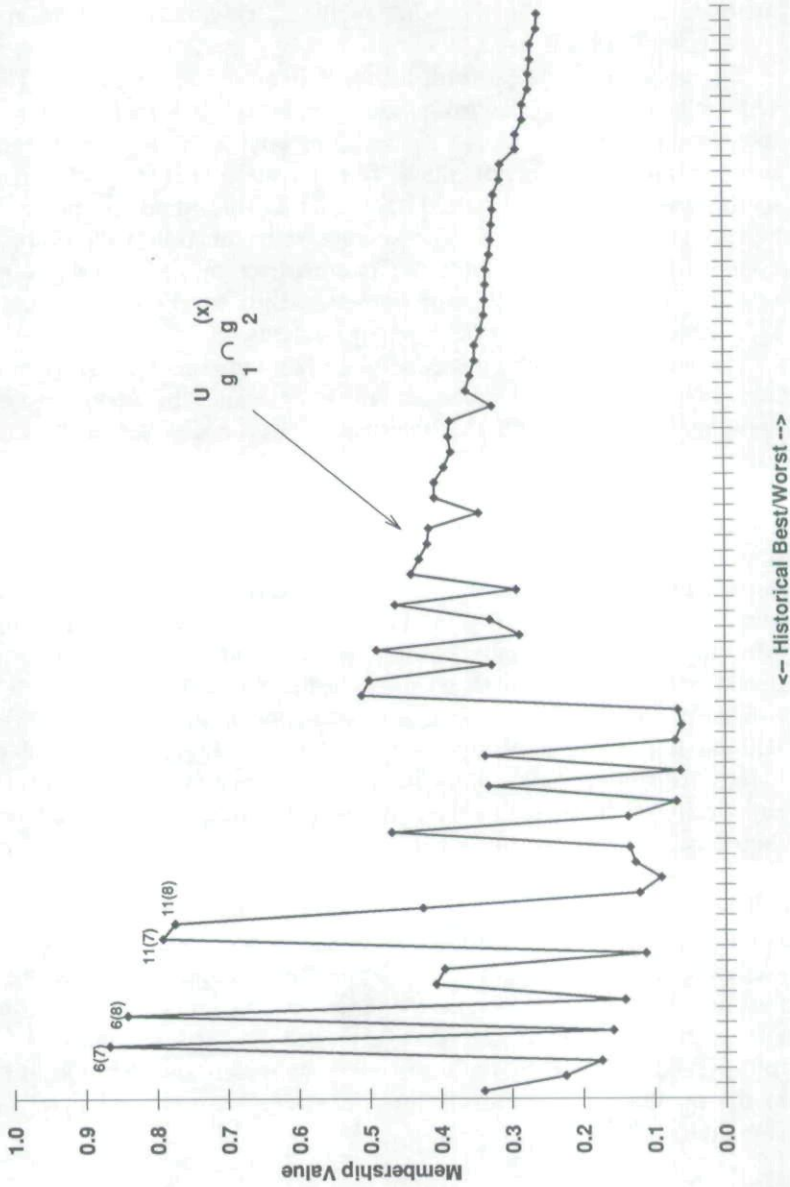


Figure 3
Intersection of Bodily Injury Liability Trend Fuzzy Sets



methods are relative to the other methods, especially when more than one goodness criterion is used.

The next step is to perform all three intersection operations. This is done sequentially, that is, the fuzzy intersection set for goals 1 and 2 is itself intersected with the fuzzy set produced by goal 3. We perform the intersection using membership function values from equation (19) for goal 1, concentrated to the power +2, equation (31) for goal 2, dilated to the power 0.75, and equation (30) for goal 3. The membership function values are shown in columns F, G, and H of Table 6. Columns I through M of Table 6 present the results obtained by applying the five intersection operators (equations [23] and [25] through [(28)]) to the 72 forecast methods.

The numbers in each intersection-operator column of Table 6 can be viewed as a fuzzy set. For example, column I represents the fuzzy set obtained by applying the ordinary intersection operator to the membership function values for goals 1, 2, and 3.

To choose a trend factor, one must first select an intersection operator. The choice of an operator depends upon how rigorously the decision-maker wants to assign membership function values and on how important it is to allow for cumulative effects, compensations, and interactions. The bounded difference operator is the most stringent (i.e., assigns the lowest membership values), while the ordinary operator is the least stringent. Since the choice is context dependent, we do not select an operator but proceed to consider all five.

One possibility for choosing a trend factor would be to adopt the method with the highest membership function value. The highest membership function values for all operators are obtained using methods 11(7) and 11(8). The superiority of these two methods in the fuzzy optimization problem is due to their having relatively high, but not necessarily the best, scores on all three rating criteria.

Choosing 11(7) or 11(8) as the solution to the decision problem, however, would disregard the information provided by closely ranked forecasting methods. Another solution is to use the membership function scores for each intersection operator as weights to compute weighted average trend factors over the entire set of methods. This approach has the advantage of retaining information provided by the set of methods under consideration in proportion to the membership function scores. The weighted average trend factors are computed as follows:

$$\text{Weighted Trend Factor: } 1 + \Delta = \sum_{i=1}^{72} \frac{m_i}{\sum_i m_i} (1 + \Delta_i), \quad (34)$$

where m_i = the membership function score for method i , and

Δ_i = the projected inflation factor produced by method i .

The weighted trend factors are presented in the last row of Table 6. They range from 13.3 percent for the bounded difference operator to 14.6 percent for the ordinary operator. These are quite close to the trend factors produced by methods 11(7) and 11(8), 14.5 and 13.5 percent, respectively. Except for the

Table 6
Bodily Injury Liability, Fuzzy Data

(A) Method	(B) Model	(C) Bodily Injury Liability Absolute TPCE	(D) Bodily Injury Liability Average Error	(E) Bodily Injury Liability Trend	(F) Historical Measure	(G) Unbiased Measure	(H) Reasonable Measure	(I) Intersection Ordinary	(J) Intersection Algebraic Product	(K) Intersection Bounded Difference	(L) Intersection Hamacher (1/2)	(M) Intersection Yager (2)
6	2	3.96%	1.44%	10.83%	1.000	0.3396	0.5766	0.3396	0.1958	0.0000	0.2276	0.2155
12	2	4.06%	1.99%	13.02%	0.9513	0.2248	0.8973	0.2248	0.1919	0.0734	0.2038	0.2165
11	1	4.19%	-2.33%	9.65%	0.8932	0.1742	0.4039	0.1742	0.0628	0.0000	0.0876	0.0000
6	7	4.21%	0.19%	11.22%	0.8848	0.8672	0.6337	0.6337	0.4862	0.3857	0.5122	0.5937
5	2	4.24%	2.46%	15.17%	0.8723	0.1580	0.7879	0.1580	0.1086	0.0000	0.1262	0.1224
6	8	4.28%	0.23%	10.33%	0.8561	0.8416	0.5034	0.5034	0.3627	0.2011	0.3934	0.4593
5	5	4.31%	-2.61%	8.51%	0.8442	0.1412	0.2370	0.1412	0.0283	0.0000	0.0454	0.0000
5	7	4.34%	1.20%	15.58%	0.8326	0.4066	0.7279	0.4066	0.2464	0.0821	0.2842	0.3260
5	8	4.36%	1.24%	14.66%	0.8249	0.3946	0.8626	0.3946	0.2808	0.0821	0.3105	0.3550
11	2	4.45%	2.92%	17.25%	0.7919	0.1119	0.4834	0.1119	0.0428	0.0000	0.0615	0.0000
11	7	4.45%	-0.24%	14.50%	0.8353	0.8860	0.8860	0.7919	0.5861	0.5131	0.6076	0.7112
11	8	4.50%	-0.12%	13.50%	0.9139	0.9139	0.9675	0.7744	0.6847	0.6558	0.6947	0.7563
12	7	4.59%	-1.14%	10.38%	0.7443	0.4253	0.5107	0.4253	0.1617	0.0000	0.2080	0.2031
11	5	4.61%	-2.82%	10.76%	0.7379	0.1206	0.5664	0.1206	0.0000	0.0000	0.0708	0.0000
12	1	4.61%	-3.21%	5.70%	0.7379	0.0900	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
5	5	4.65%	-2.76%	8.97%	0.7252	0.1262	0.3043	0.1262	0.0278	0.0000	0.0460	0.0000
11	6	4.67%	-2.68%	10.01%	0.7190	0.1340	0.4566	0.1340	0.0000	0.0000	0.0661	0.0000
12	8	4.67%	-1.01%	9.41%	0.7190	0.4688	0.3687	0.3687	0.1243	0.0000	0.1680	0.1284
5	6	4.73%	-2.66%	8.48%	0.7009	0.1360	0.2326	0.1360	0.0222	0.0000	0.0387	0.0000
6	1	4.90%	-3.58%	4.41%	0.6531	0.0682	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	1	5.09%	-1.47%	13.57%	0.6053	0.3320	0.9778	0.3320	0.1965	0.0000	0.2283	0.2238
12	5	5.15%	-3.70%	6.77%	0.5913	0.0623	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
4	1	5.19%	-1.45%	13.53%	0.5822	0.3371	0.9719	0.3371	0.1907	0.0000	0.2238	0.2159
6	6	5.23%	-3.55%	6.04%	0.5733	0.0698	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	5	5.26%	-3.73%	4.86%	0.5668	0.0610	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
6	6	5.36%	-3.62%	4.38%	0.5458	0.0662	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
9	1	5.44%	-0.33%	16.00%	0.5109	0.7808	0.6664	0.5109	0.2658	0.0000	0.3109	0.3687
10	2	5.60%	0.26%	14.33%	0.5001	0.8228	0.9109	0.5001	0.3748	0.2338	0.4024	0.4622
10	5	5.63%	-1.49%	14.81%	0.4947	0.3271	0.8407	0.3271	0.1360	0.0000	0.1752	0.1436
3	3	5.66%	-0.28%	15.98%	0.4895	0.8106	0.6694	0.4895	0.2656	0.0000	0.3088	0.3630
10	6	5.73%	-1.66%	13.91%	0.4776	0.2879	0.9724	0.2879	0.1337	0.0000	0.1661	0.1164
4	5	5.76%	-1.48%	14.75%	0.4727	0.3296	0.8494	0.3296	0.1323	0.0000	0.1712	0.1339
4	5	5.82%	0.31%	14.32%	0.4630	0.2925	0.9124	0.4630	0.3348	0.1679	0.3643	0.4177
4	4	5.84%	-1.64%	13.94%	0.4598	0.7923	0.9680	0.2923	0.1301	0.0000	0.1631	0.1091
4	6	4.84%	-0.51%	15.74%	0.6822	0.7045	0.7045	0.4400	0.2115	0.0000	0.2577	0.2915
10	7	5.97%	-0.69%	14.78%	0.4284	0.5960	0.8450	0.4284	0.2158	0.0000	0.2582	0.2830
10	8	6.05%	-0.69%	14.78%	0.4284	0.5960	0.8450	0.4284	0.2158	0.0000	0.2582	0.2830
4	7	6.13%	-9.46%	15.66%	0.4173	0.7082	0.7162	0.4173	0.2117	0.0000	0.2559	0.1856
4	5	6.15%	-0.34%	17.27%	0.4146	0.7749	0.4805	0.4146	0.1544	0.0000	0.1992	0.0122
9	9	6.17%	1.42%	16.78%	0.4119	0.3447	0.5522	0.3447	0.0784	0.0000	0.1191	0.0122
4	8	6.21%	-0.64%	14.80%	0.4066	0.6188	0.8421	0.4066	0.2119	0.0000	0.2532	0.2772
9	6	6.21%	-0.51%	16.35%	0.4066	0.6822	0.6152	0.4066	0.1706	0.0000	0.2175	0.2247

Table 6 continued

(A)	(B)	(C)	(D)	(E)	(F)	(G)	(H)	(I)	(J)	(K)	(L)	(M)
Method	Model	Bodily Injury Liability Absolute TPCE	Bodily Injury Liability Average Error	Bodily Injury Liability Trend	Historical Measure	Unbiased Measure	Reasonable Measure	Intersection Ordinary	Intersection Algebraic Product	Intersection Bounded Difference	Intersection Hamacher (1/2)	Intersection Yaeger (2)
3	5	6.32%	-0.30%	17.22%	0.3926	0.7985	0.4878	0.3926	0.1529	0.0000	0.1964	0.1803
8	1	6.36%	-1.28%	13.49%	0.3877	0.3829	0.9661	0.3829	0.1434	0.0000	0.1793	0.1300
3	6	6.37%	-0.47%	16.40%	0.3865	0.7029	0.6079	0.3865	0.1651	0.0000	0.2107	0.2136
2	1	6.39%	-1.23%	14.84%	0.3841	0.3975	0.8363	0.3841	0.1277	0.0000	0.1680	0.1230
3	2	6.45%	1.50%	16.79%	0.3769	0.3247	0.8363	0.3247	0.1066	0.0000	0.1054	0.0000
9	7	6.59%	0.65%	18.22%	0.3611	0.6142	0.5508	0.3247	0.0757	0.0000	0.1146	0.0047
9	8	6.64%	0.46%	17.24%	0.3557	0.7082	0.4849	0.3557	0.1221	0.0000	0.1656	0.1250
2	5	6.71%	-1.37%	15.21%	0.3483	0.3579	0.7821	0.3483	0.0975	0.0000	0.1357	0.0595
8	5	6.71%	-1.38%	14.01%	0.3483	0.3552	0.9578	0.3483	0.1185	0.0000	0.1527	0.0822
3	7	6.79%	0.73%	18.15%	0.3401	0.5784	0.3516	0.3401	0.1071	0.0000	0.1071	0.0000
3	8	6.84%	0.55%	17.27%	0.3352	0.6620	0.4805	0.3352	0.1066	0.0000	0.1492	0.0911
2	6	6.85%	-1.43%	15.31%	0.3342	0.3422	0.7675	0.3342	0.0878	0.0000	0.1247	0.0357
8	6	6.86%	-1.44%	14.07%	0.3332	0.3396	0.9490	0.3322	0.1074	0.0000	0.1408	0.0601
7	1	6.87%	-0.32%	13.74%	0.3323	0.7866	0.9973	0.3323	0.2607	0.1162	0.2809	0.2990
1	1	6.92%	-0.25%	15.18%	0.3275	0.8290	0.7865	0.3275	0.2135	0.0000	0.2452	0.2740
2	7	6.95%	-0.77%	16.03%	0.3247	0.5613	0.6620	0.3247	0.1207	0.0000	0.1633	0.1266
8	7	6.96%	-0.78%	14.79%	0.3237	0.5571	0.8436	0.3237	0.1521	0.0000	0.1907	0.1766
2	2	6.98%	0.04%	16.87%	0.3219	0.9704	0.5391	0.3219	0.1684	0.0000	0.2019	0.1796
8	2	6.99%	-0.01%	15.47%	0.3209	0.9925	0.7440	0.3209	0.2370	0.0574	0.2602	0.2743
2	8	7.09%	-0.83%	16.15%	0.3120	0.5366	0.6445	0.3120	0.1079	0.0000	0.1497	0.0975
8	8	7.10%	-0.84%	14.88%	0.3111	0.5326	0.8304	0.3111	0.1376	0.0000	0.1760	0.1504
1	5	7.36%	-0.39%	15.53%	0.2895	0.7464	0.2895	0.2895	0.1582	0.0000	0.1937	0.1995
7	5	7.36%	-0.41%	14.25%	0.2895	0.7353	0.9226	0.2895	0.1964	0.0000	0.2234	0.2379
1	6	7.49%	-0.46%	15.64%	0.2795	0.7082	0.7191	0.2795	0.1423	0.0000	0.1786	0.1735
7	6	7.49%	-0.47%	14.32%	0.2795	0.7029	0.9124	0.2795	0.1793	0.0000	0.2078	0.2157
1	2	7.60%	1.02%	17.21%	0.2715	0.4653	0.4893	0.2715	0.0618	0.0000	0.1308	0.0615
7	2	7.61%	0.96%	15.72%	0.2708	0.4868	0.7074	0.2708	0.0933	0.0000	0.1736	0.1589
1	7	7.65%	0.21%	16.37%	0.2680	0.8543	0.6121	0.2680	0.1402	0.0000	0.2315	0.2315
7	7	7.65%	0.19%	15.04%	0.2680	0.8672	0.8070	0.2680	0.1876	0.0000	0.1702	0.1496
1	8	7.78%	0.14%	16.49%	0.2591	0.9003	0.5947	0.2591	0.1387	0.0000	0.2091	0.2247
7	8	7.79%	0.13%	15.13%	0.2584	0.9071	0.7938	0.2584	0.1861	0.0000	0.1895	0.1742
	Mean	5.92%	-0.65%	13.72%	0.5042	0.5064	0.6362	0.3140	0.1590	0.0345	0.1895	0.1742
	Standard Deviation	1.13%	1.46%	3.42%	0.2076	0.2749	0.2732	0.1538	0.1234	0.1123	0.1259	0.1629
	Minimum	3.96%	-3.73%	4.38%	0.2584	0.0610	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	Weighted Trend			13.72%	12.96%	14.85%	14.59%	14.64%	14.44%	13.31%	14.50%	14.45%

Note: (F) = [minimum of Column (C)]/[(C)²], (G) = $\exp(-0.75 * |(D)|)$, (H) = maximum [0.1 - |(E)| - mean of column (E)]/2, (I) = minimum[minimum[(F)_i](G)_i], (J) = (F)²*(G)²*(H), (K) = maximum of [0, (minimum of [(F)_i+(G)_i]-1)+(H)-1], (L) = $f(f(F)_i, (G)_i, (H)_i)$ where $f[x, y] = (x*y)/(0.5+(x*y-x*y))$, (M) = $f(f(F)_i, (G)_i, (H)_i)$ where $[x, y] = 1 - \text{minimum}[1, ((1-x)^2) + ((1-y)^2)]/0.51$. TPCE = total percentage change error.

bounded difference operator, the trend factors produced by the fuzzy intersection operators cluster around 14.5 percent. Thus, a trend factor in the neighborhood of 14.5 percent would seem to be a good choice. It is noteworthy that this trend factor is about 0.75 percentage points higher than the average trend factor, reflecting the predictive information of the membership functions. Except for the bounded difference operator, the weighted trend factor is not very sensitive to the choice of an intersection operator.

The traditional actuarial approach would be to select a few (e.g., two or three) methods based on judgment and/or prior forecast performance and to obtain the trend factor as a simple average of the factors produced by the selected methods. In fuzzy terms, this would be analogous to taking an " α -cut" from the set of available methods based on implicit "membership function" scores and then averaging the trend factors produced by the resulting methods. In true fuzzy decision making, the membership values are assigned explicitly by functions rather than implicitly by the decision-maker; that is, the fuzzy decision-maker is required to express his or her judgment in mathematical terms. In our weighted average approach (equation [34]), the information provided by the membership function values is explicitly incorporated in the trend factor rather than giving equal weight to the methods in the traditionalist's choice set. While multiple criteria are applied subjectively by the traditional actuary, the fuzzy decision-maker uses mathematical operators with well-defined properties. Thus, the fuzzy approach imposes an important measure of rigor in areas where it is lacking in the traditional approach.²¹

Because of its subjective elements, FST is controversial among mathematicians and scientists. We view the subjective nature of FST to be an advantage rather than a disadvantage. The choice of a forecasting method and trend factor is inevitably a subjective decision involving multiple criteria. The value of FST in such a decision context is that it provides a way to organize and systematize the objective and subjective information at the decision-maker's disposal. It requires the decision-maker to quantify the judgmental criteria, thus ensuring consistency in applying the criteria across methods. The decision-maker's judgment is reflected in the choice of membership functions and an intersection operator. Thus, FST provides a set of mathematical rules through which judgment is introduced into the decision process. The involvement of decision-makers in the specification of goals, constraints, and membership functions is vital to the success of FST.

Although the selection of membership functions is of necessity somewhat arbitrary, a more systematic approach could be taken in practice. For example, an analysis of the relative costs to the insurer of various types of ratemaking errors could be used to formulate goals, specify membership functions, etc.

²¹ Of course, more rigorous "classical" approaches also are available, such as utility theory and statistical decision theory. Although these methods have not been used very much by practicing actuaries, comparisons with the FST approach offer an interesting avenue for future research.

In some respects, FST optimization is similar to utility theory, although utility theory always has been difficult to apply in practical decision-making situations because of the need to decide whose preference function is relevant and then estimate the functional form or degree of risk aversion. Fuzzy set theory provides a potentially superior method for quantifying preferences in a way that is easily understood by decision-makers and easily operationalized in a group decision-making environment. It is not surprising that Japanese manufacturers, with their emphasis on consensus management, have found FST to be a useful tool.

Conclusions

We have explored the use of fuzzy set theory as a decision-making technique in insurance loss cost forecasting. The decision process involves the evaluation of alternative forecasting methods using multiple criteria, where the degree to which a given element satisfies a criterion is vague or fuzzy. Although the historical performance of various methods provides valuable information, methods that performed well historically will not necessarily give the best forecast in the future. The ultimate objective is the choice of a "good" forecast rather than a forecasting method.

Massachusetts private passenger automobile quarterly pure premium data for bodily injury liability were used to test empirically 72 benchmark forecasting methods for the standard seven-quarter ahead ratemaking problem. The results illustrate the difficulty in selecting the best method using conventional methods when many alternatives are available and multiple "goodness" criteria are employed. The decision problem is especially acute when methods that perform equally well in historical tests give widely divergent forecasts for the future rating period.

The nature of the loss cost forecasting problem suggests the use of fuzzy set theory as a decision-making technique. Fuzzy set theory is appropriate for the following reasons: First, some or all of the criteria used in evaluating forecasts are inherently vague and most are expressed verbally rather than quantitatively. For example, the decision-maker seeks a forecast that is "accurate," "adequate," "reasonable," etc. Second, multiple criteria are present. For example, the forecaster is concerned not only about absolute accuracy but also about the bias of the methods under consideration. Of course, FST is not the only approach to multiple criteria decision making, but it is a particularly natural way to handle this type of problem in the presence of fuzzy information and fuzzy criteria. In such situations, conventional or "crisp" decision methods are likely to be less successful. Third, there is growing evidence that forecast accuracy can be improved by combining forecasts rather than relying on a single forecast or forecasting method. Conventionally, forecasters either choose a forecast based on a single method, apply judgment to an otherwise determined forecast, or combine forecasts using straight averages. All of these approaches are implicitly based on ordinary set theory, that is, a method is either a full member or not a member of the set of methods used in generating

a forecast. FST has the potential for combining forecasts in a more meaningful way using information provided by their degree of membership in the fuzzy set of "good" forecasts.

Fuzzy set theory provides an effective method for mixing statistical and judgmental criteria in the selection of forecasting methods and trend factors. Judgment is important in actuarial trending because past performance may not be an accurate guide to the future. In conclusion, we pose the following trend factor meta theorem:

1. The pursuit of a single forecasting method is futile.

2. All actuarial forecasting problems involve fuzzy sets and fuzzy operations. Solutions can be fuzzy or non-fuzzy sets of options.

Actuarial science has always involved the mixture of statistics and judgment, and FST provides a mechanism for doing this in a logical and systematic way. FST should become one tool in the standard repertoire of techniques in actuarial decision making.

APPENDIX

Massachusetts Automobile Insurance Pure Premium Trend Factors
 Bodily Injury Liability Coverage
 Forecast Period: 1989.1-1990.3

<i>Method</i>		<i>Frequency</i>		<i>Severity</i>	<i>Percent</i>
1(1)	OLS	8Q	TIME-EXP	TIME-LIN	15.18
1(2)	OLS	8Q	TIME-EXP	TIME-EXP	17.21
1(5)	OLS	8Q	TIME-EXP	ECON-LIN	15.55
1(6)	OLS	8Q	TIME-EXP	LOG(ECON-LIN)	15.64
1(7)	OLS	8Q	TIME-EXP	ECON-EXP	16.37
1(8)	OLS	8Q	TIME-EXP	LOG(ECON-EXP)	16.49
2(1)	OLS	8Q	TIME-LIN	TIME-LIN	14.84
2(2)	OLS	8Q	TIME-LIN	TIME-EXP	16.87
2(5)	OLS	8Q	TIME-LIN	ECON-LIN	15.21
2(6)	OLS	8Q	TIME-LIN	LOG(ECON-LIN)	15.31
2(7)	OLS	8Q	TIME-LIN	ECON-EXP	16.03
2(8)	OLS	8Q	TIME-LIN	LOG(ECON-EXP)	16.15
3(1)	OLS	12Q	TIME-EXP	TIME-LIN	15.98
3(2)	OLS	12Q	TIME-EXP	TIME-EXP	16.79
3(5)	OLS	12Q	TIME-EXP	ECON-LIN	17.22
3(6)	OLS	12Q	TIME-EXP	LOG(ECON-LIN)	16.40
3(7)	OLS	12Q	TIME-EXP	ECON-EXP	18.15
3(8)	OLS	12Q	TIME-EXP	LOG(ECON-EXP)	17.27
4(1)	OLS	12Q	TIME-LIN	TIME-LIN	13.53
4(2)	OLS	12Q	TIME-LIN	TIME-EXP	14.32
4(5)	OLS	12Q	TIME-LIN	ECON-LIN	14.75
4(6)	OLS	12Q	TIME-LIN	LOG(ECON-LIN)	13.94
4(7)	OLS	12Q	TIME-LIN	ECON-EXP	15.66
4(8)	OLS	12Q	TIME-LIN	LOG(ECON-EXP)	14.80
5(1)	OLS	LONG	TIME-EXP	TIME-LIN	8.51
5(2)	OLS	LONG	TIME-EXP	TIME-EXP	15.17
5(5)	OLS	LONG	TIME-EXP	ECON-LIN	8.97
5(6)	OLS	LONG	TIME-EXP	LOG(ECON-LIN)	8.48
5(7)	OLS	LONG	TIME-EXP	ECON-EXP	15.58
5(8)	OLS	LONG	TIME-EXP	LOG(ECON-EXP)	14.66
6(1)	OLS	LONG	TIME-LIN	TIME-LIN	4.41
6(2)	OLS	LONG	TIME-LIN	TIME-EXP	10.83
6(5)	OLS	LONG	TIME-LIN	ECON-LIN	4.86
6(6)	OLS	LONG	TIME-LIN	LOG(ECON-LIN)	4.38
6(7)	OLS	LONG	TIME-LIN	ECON-EXP	11.22
6(8)	OLS	LONG	TIME-LIN	LOG(ECON-EXP)	10.33

7(1)	AR1	8Q	TIME-EXP	TIME-LIN	13.74
7(2)	AR1	8Q	TIME-EXP	TIME-EXP	15.72
7(5)	AR1	8Q	TIME-EXP	ECON-LIN	14.25
7(6)	AR1	8Q	TIME-EXP	LOG(ECON-LIN)	14.32
7(7)	AR1	8Q	TIME-EXP	ECON-EXP	15.04
7(8)	AR1	8Q	TIME-EXP	LOG(ECON-EXP)	15.13
8(1)	AR1	8Q	TIME-LIN	TIME-LIN	13.49
8(2)	AR1	8Q	TIME-LIN	TIME-EXP	15.47
8(5)	AR1	8Q	TIME-LIN	ECON-LIN	14.01
8(6)	AR1	8Q	TIME-LIN	LOG(ECON-LIN)	14.07
8(7)	AR1	8Q	TIME-LIN	ECON-EXP	14.79
8(8)	AR1	8Q	TIME-LIN	LOG(ECON-EXP)	14.88
9(1)	AR1	12Q	TIME-EXP	TIME-LIN	16.00
9(2)	AR1	12Q	TIME-EXP	TIME-EXP	16.78
9(5)	AR1	12Q	TIME-EXP	ECON-LIN	17.27
9(6)	AR1	12Q	TIME-EXP	LOG(ECON-LIN)	16.35
9(7)	AR1	12Q	TIME-EXP	ECON-EXP	18.22
9(8)	AR1	12Q	TIME-EXP	LOG(ECON-EXP)	17.24
10(1)	AR1	12Q	TIME-LIN	TIME-LIN	13.57
10(2)	AR1	12Q	TIME-LIN	TIME-EXP	14.33
10(5)	AR1	12Q	TIME-LIN	ECON-LIN	14.81
10(6)	AR1	12Q	TIME-LIN	LOG(ECON-LIN)	13.91
10(7)	AR1	12Q	TIME-LIN	ECON-EXP	15.74
10(8)	AR1	12Q	TIME-LIN	LOG(ECON-EXP)	14.78
11(1)	AR1	LONG	TIME-EXP	TIME-LIN	9.65
11(2)	AR1	LONG	TIME-EXP	TIME-EXP	17.25
11(5)	AR1	LONG	TIME-EXP	ECON-LIN	10.76
11(6)	AR1	LONG	TIME-EXP	LOG(ECON-LIN)	10.01
11(7)	AR1	LONG	TIME-EXP	ECON-EXP	14.50
11(8)	AR1	LONG	TIME-EXP	LOG(ECON-EXP)	13.50
12(1)	AR1	LONG	TIME-LIN	TIME-LIN	5.70
12(2)	AR1	LONG	TIME-LIN	TIME-EXP	13.02
12(5)	AR1	LONG	TIME-LIN	ECON-LIN	6.77
12(6)	AR1	LONG	TIME-LIN	LOG(ECON-LIN)	6.04
12(7)	AR1	LONG	TIME-LIN	ECON-EXP	10.38
12(8)	AR1	LONG	TIME-LIN	LOG(ECON-EXP)	9.41
Mean					13.72
Standard Deviation					3.42

References

- Actuarial Standards Board, 1990, *Actuarial Standards of Practice No. 13: Trending Procedures in Property/Casualty Insurance Ratemaking* (Washington, D.C.: American Academy of Actuaries).
- Armstrong, J. Scott, 1978, *Long-Range Forecasting: From Crystal Ball to Computer* (New York: John Wiley).
- Automobile Insurers Bureau, 1989, *Test of the Decision Methodology: Section 100CD Historical Supplement*, Boston.
- Casualty Actuarial Society, 1990, *Foundations of Casualty Actuarial Science* (New York: Casualty Actuarial Society).
- Clemen, R. T., 1989, Combining Forecasts: A Review and Annotated Bibliography, *International Journal of Forecasting*, 5: 559-583.
- Cummins, J. David, 1988, Risk-Based Premiums for Insurance Guaranty Funds, *Journal of Finance*, 43: 823-839.
- Cummins, J. David and Gary Griepentrog, 1985, Forecasting Automobile Insurance Paid Claim Costs Using Econometric and ARIMA Models, *International Journal of Forecasting*, 1: 203-215.
- Cummins, J. David and Scott E. Harrington, 1985, Econometric Forecasting of Automobile Insurance Paid Claim Costs, in: J. David Cummins, ed., *Strategic Planning and Modeling in Property-Liability Insurance* (Norwell, Mass.: Kluwer Academic Publishers).
- Cummins, J. David and David J. Nye, 1981, Inflation and Property-Liability Insurance, in: John D. Long, ed., *Issues in Insurance*, Volume II, Second Edition (Malvern, Penn.: American Institute).
- Cummins, J. David and Alwyn Powell, 1980, The Performance of Alternative Models for Forecasting Automobile Insurance Paid Claim Costs, *Astin Bulletin*, 11: 91-106.
- Cummins, J. David and Sharon Tennyson, 1992, Controlling Automobile Insurance Costs, *Journal of Economic Perspectives*, 6: 95-115.
- Derrig, Richard A., 1987, The Use of Investment Income in Massachusetts Private Passenger Automobile and Workers' Compensation Ratemaking, in: J. David Cummins and Scott E. Harrington, eds., *Fair Rate of Return in Property-Liability Insurance* (Norwell, Mass.: Kluwer Academic Publishers).
- Derrig, Richard A., 1993, Price Regulation in U.S. Automobile Insurance: A Case Study of Massachusetts Private Passenger Automobile Insurance, 1978-1990, *Geneva Papers on Risk and Insurance*.
- Doherty, Neil A. and James Garven, 1986, Price Regulation in Property and Liability Insurance: A Contingent Claims Approach, *Journal of Finance*, 41: 1031-1050.
- Dubois, Didier and Henri Prade, 1980, *Fuzzy Sets and Systems: Theory and Applications* (New York: Academic Press).
- Granger, C. W. J. and Paul Newbold, 1986, *Forecasting Economic Time Series*, Second Edition (New York: Academic Press).
- Kandel, Abraham, 1986, *Fuzzy Mathematical Techniques with Applications* (Reading, Mass.: Addison-Wesley).

- Kmenta, Jan, 1986, *Elements of Econometrics*, Second Edition (New York: Macmillan).
- Lemaire, Jean, 1985, *Automobile Insurance Actuarial Models* (Norwell Mass.: Kluwer Academic Publishers)
- Lemaire, Jean, 1990, Fuzzy Insurance, *Astin Bulletin*, 20: 33-56.
- McNees, Stephen K., 1990, Man vs. Model: The Role of Judgment in Forecasting, *New England Economic Review*, July/August: 41-52.
- McNees, Stephen K., 1992, How Large Are Economic Forecast Errors? *New England Economic Review*, July/August: 25-42.
- Masterson, Norton E., 1990, Economic Factors in Property/Casualty Insurance Claim Costs, *Bests' Review*, May: 81-82.
- Weisberg, Herbert I. and Richard A. Derrig, 1991, Fraud and Automobile Insurance: A Report on the Baseline Study of Bodily Injury Claims in Massachusetts, *Journal of Insurance Regulation*, 9:427-541.
- Zarnowitz, Victor and Phillip Braun, 1992, Twenty-Two Years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and Comparisons of Forecasting Performance, NBER Working Paper 3965, National Bureau of Economic Research, Cambridge, Massachusetts.
- Zimmerman, H. J., 1991, *Fuzzy Set Theory and Its Applications* (Norwell, Mass.: Kluwer Academic Publishers).

Copyright of Journal of Risk & Insurance is the property of Blackwell Publishing Limited. The copyright in an individual article may be maintained by the author in certain cases. Content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.