Insurance: Mathematics and Economics 46 (2010) 554–567

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Detecting fuzzy relationships in regression models: The case of insurer solvency surveillance in Germany

Thomas R. Berry-Stölzle[a,*], Marie-Claire Koissi[b,1], Arnold F. Shapiro[c,2]

[a] Terry College of Business, University of Georgia, 206 Brooks Hall, Athens, GA 30602, United States

[b] Department of Mathematics, Western Illinois University, 1 University Circle, Macomb, IL 61455, United States

[c] Smeal College of Business, Penn State University, University Park, PA 16802, United States

ARTICLE INFO

Article history:
Received June 2009
Received in revised form
February 2010
Accepted 16 February 2010

JEL classification:
C12
C15
C16
C21
G22
G28

Keywords:
Test for fuzziness
Possibilistic fuzzy regression
Financial statement data
Insurance regulation

ABSTRACT

We develop a test for the fuzziness of regression coefficients based on the Tanaka et al. (1982) and He et al. (2007) possibilistic fuzzy regression models. We interpret the spread of the regression coefficients as a statistic measuring the fuzziness of the relationship between the corresponding independent variable and the dependent variable. We derive test distributions based on the null hypothesis that such spreads could have been obtained by estimating a possibilistic regression with data generated by a classical regression model with random errors. As an example, we show how our test detects a fuzzy regression coefficient in a solvency prediction model for German property–liability insurance companies.

1. Introduction

Standard econometric models assume a clear functional relationship between the dependent and the independent variables, and that this relationship is only distorted by random error. In many applications, however, the data used are subjective, incomplete, or vague, resulting in a rather vague functional relationship between the dependent and the independent variables.

To illustrate our point, we use an example based on accounting data from German insurance companies. We expect there to be vagueness in the data for the following reasons. The German accounting principles can be described as conservative. One prominent example of the conservative nature of the German accounting system is the way insurance companies have to display certain kinds of investments on their balance sheet. The "lower of cost or market principle" which applies, for example, to stocks[3] requires an insurer to use the minimum of the purchase price and the current market value of the stock holding on the balance sheet. An increase in the market value of this stock position, hence, results in hidden reserves not shown on the balance sheet. However, the management of an insurance company might, at any time, decide to sell this stock position and reinvest it in the same stock or some other stock. Such a portfolio rebalancing transaction makes the previously hidden asset values visible on the balance sheet. Thus, insurance managers in Germany have a certain degree of discretion in presenting the financial situation of their company. The source of imprecision in this example is the absence of a sharply defined classification criterion and not the presence of a random variable. A model framework explicitly capturing such vagueness is the fuzzy set theory originating from Zadeh's (1965) seminal work.[4]

* Corresponding author. Tel.: +1 706 542 5160; fax: +1 706 542 4295.
E-mail addresses: trbs@terry.uga.edu (T.R. Berry-Stölzle), mk122@wiu.edu (M.-C. Koissi), afs1@psu.edu (A.F. Shapiro).

1 Tel.: +1 309 298 2358.

2 Tel.: +1 814 865 3961; fax: +1 814 865 6284.

3 The lower of cost or market principle as stated in Section 253 (3) of the German code of commercial law (*Handelsgesetzbuch* or HGB) applies to stock investments which are short term in nature (Umlaufvermögen). Long term investments fulfilling certain requirements (Anlagevermögen) are subject to Section 253 (2) HGB instead.

4 Since Zadeh (1965) introduced fuzzy logic as a model for explicitly describing vagueness of set membership, his approach has gained recognition and inspired

0167-6687/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.insmatheco.2010.02.003

The goal of this paper is twofold. First, we develop a test procedure for explicitly examining whether an independent variable has a clear functional relationship with the dependent variable in a specific regression model, or whether the relationship between this independent variable and the dependent variable is fuzzy. Second, we provide an example of how our test procedure finds a significantly fuzzy relationship between an independent variable and the dependent variable in a regression model.

Our new test procedure builds on the Tanaka et al. (1982) fuzzy regression model. The starting point for Tanaka's approach is the assumption that the relationship between the dependent and the independent variables is vague or fuzzy. Therefore, Tanaka et al. (1982) model the regression coefficients as fuzzy numbers, and then fit their model to data by minimizing the overall fuzziness of the model. This approach is often referred to as possibilistic fuzzy regression (see, e.g., Chang and Ayyub, 2001; Shapiro, 2004a). The main advantage of a possibilistic regression is that we can interpret the spread of each of the regression coefficients as a measure for how fuzzy the relationship between the corresponding independent variable and the dependent variable is without having any further *a priori* assumptions about the independent variables (Chang and Ayyub, 2001; Shapiro, 2004a). We then derive an empirical test distribution for each of these spreads based on the null hypothesis that such a spread could have been obtained by estimating a possibilistic regression model with a dataset generated by a classical regression model with random errors, i.e.

$$Y = X\beta + \varepsilon. \tag{1.1}$$

If we can reject this null hypothesis then we can conclude that the functional relationship between the corresponding independent variable and the dependent variable is fuzzy. Hence, this test allows us to determine the regression coefficient for which the notion of "fuzziness" is actually meaningful. However, it is important to keep in mind that the detected fuzziness only refers to the relationship between an independent variable and the dependent variable; the test does not examine whether the independent variable itself is fuzzy. We would also like to point out that the results of our test for the fuzziness of regression coefficients are model specific. Adding or deleting variables from the regression model might change the test results as is the case with the significance of regression coefficients based on a standard t-test.

To illustrate our test procedure, we provide an example of a regression model predicting the financial strength of German property–liability insurance companies two years ahead. Applying our new test reveals that one of the ten regression coefficients in the model is significantly fuzzy, indicating that the relationship between the corresponding independent variable and the dependent variable is fuzzy.

This paper proceeds as follows. Section 2 introduces the fuzzy regression framework and our test for the fuzziness of regression coefficients. For deriving our baseline test distributions, we assume that the error term ε in Eq. (1.1) is a random vector with independent and normally distributed components ε_i, $i = 1, \ldots, N$. Section 3 describes solvency surveillance in Germany. This section discusses the regulatory situation in Germany, includes a brief overview of the insurer solvency surveillance literature, and explains the selection of our predictor variables. Section 4 describes our dataset. Section 5 presents the results from standard OLS and FGLS estimations as well as diagnosis tests for the residuals of these regressions. The tests indicate that for some model specifications, regression residuals are heteroscedastic and not normally distributed. To address the concern that our test for the fuzziness of regression coefficients only detects fuzziness because of misspecified test distributions, we derive two additional sets of test distributions in Section 6. One set of test distributions drops the normality assumption, and the other set of test distributions drops the independence assumption. Section 7 presents the results of these tests. The final section concludes.

2. Fuzzy regression analysis

2.1. The possibilistic fuzzy regression framework

In a fuzzy regression model the regression coefficients and/or the dependent variable are fuzzy, rather than crisp, numbers. In addition, the residuals between estimators and observations are not produced by measurement errors but rather by the parameter uncertainty in the model.

Fuzzy numbers are characterized by their membership functions which can be triangular, trapezoidal, Gaussian, generalized bell or a combination of these basic classes (extensive overviews and applications are provided in Shapiro, 2004b). For simplicity, fuzzy parameters in the form of triangular fuzzy numbers are used in this study:

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 - \dfrac{c-x}{l_A} & \text{if } c - l_A \le x \le c \\ 1 - \dfrac{x-c}{r_A} & \text{if } c \le x \le c + r_A \\ 0 & \text{otherwise} \end{cases} \tag{2.1}$$

where $\mu_{\tilde{A}}(x)$ is the membership function of the triangular fuzzy number $\tilde{A} = (c, l_A, r_A)$ with center $c \in \mathbf{R}$ and left and right spreads (l_A, r_A). The analysis can be extended to other membership types.

2.1.1. The standard possibilistic regression model

In the case where only the coefficients are fuzzy, the fuzzy regression equation is given by

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i1} + \tilde{A}_2 x_{i2} + \cdots + \tilde{A}_k x_{ik} \quad \text{for } i = 1, \ldots, N, \tag{2.2}$$

where $x_i = (1, x_{i,1}, \ldots, x_{i,k})$ is a real input vector of independent variables, and $(\tilde{A}_0, \tilde{A}_1, \ldots, \tilde{A}_k)$ are the fuzzy coefficients. Assuming symmetrical triangular fuzzy numbers, let c_j and s_j denote the center and spread of the fuzzy coefficient $\tilde{A}_j$, $j = 0, \ldots, k$, respectively. Then Eq. (2.2) can be rewritten as

$$\begin{aligned} \tilde{Y}_i = {} & (c_0, s_0) + (c_1, s_1)x_{i1} + (c_2, s_2)x_{i2} + \cdots \\ & + (c_k, s_k)x_{ik} \quad \text{for } i = 1, \ldots, N, \end{aligned} \tag{2.3}$$

which leads to

$$\begin{aligned} \tilde{Y}_i = {} & (c_0 + c_1 x_{i1} + c_2 x_{i2} + \cdots + c_k x_{ik}, \\ & s_0 + s_1|x_{i1}| + s_2|x_{i2}| + \cdots + s_k|x_{ik}|). \end{aligned} \tag{2.4}$$

The main characteristic of the possibilistic fuzzy regression is that Eq. (2.4) is fit to a dataset by minimizing the total spread of the fuzzy coefficients subject to the constraint that the observations for the dependent variable are included within a specified feasible

applications in mathematics and computer science (see, e.g., Dubois and Prade, 1980; Kandel, 1986; Zimmermann, 1996). Fuzzy logic has also entered the insurance literature with applications in underwriting (DeWitt, 1982), classification of insurance risks (Ebanks et al., 1992; Derrig and Ostaszewski, 1995), liability projections (Cummins and Derrig, 1993; Sánchez and Gómez, 2003), future and present value calculations (Buckey, 1987) and financial pricing (Lemaire, 1990). A comprehensive review of insurance applications of fuzzy logic is provided by Shapiro (2004b).

In the area of data modeling, Tanaka et al. (1982) and Diamond (1988) were the first to develop fuzzy regression models explicitly addressing the vagueness of data. Since then, fuzzy regression was used in various fields, including insurance (Sánchez and Gómez, 2003). Chang and Ayyub (2001) and Shapiro (2004a) provide a review of fuzzy regression models and applications.

data interval. This latter requirement can be formulated as $\mu(\tilde{Y}_{x,t} \subseteq \hat{\tilde{Y}}_{x,t}) \geq h$, where $h \in [0, 1]$ is the so called h-certain factor or "degree of belief", and is chosen arbitrarily. However, since the coefficients $\tilde{A}_j = (c_j, s_j)$ for a given $h \neq 0$ are proportional to the ones derived for $h = 0$ (Tanaka and Watada, 1988, Theorem 5), it is sufficient to analyze the case $h = 0$. For a given h, the possibilistic fuzzy regression model can be formulated as the following linear program:

$$\text{Minimize} \quad \sum_{i=1}^{N} \left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}| \right] \tag{2.5a}$$

subject to

$$\begin{cases} c_0 + \sum_{j=1}^{k} c_j x_{ij} + (1-h) \left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}| \right] \geq Y_i \\ c_0 + \sum_{j=1}^{k} c_j x_{ij} - (1-h) \left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}| \right] \leq Y_i, \\ s_j \geq 0 \quad \forall j = 0, 1, \ldots, k \end{cases} \quad i = 1, \ldots, N. \tag{2.5b}$$

This linear program can be solved with standard software tools like LINDO or Matlab (see Appendix A).

2.1.2. The possibilistic regression model with endogenous h-certain factor

He et al. (2007) developed a model which determines a degree of belief h_i for every observation i endogenously within the optimization. This model is expected to provide a better fit to the data. The changes that they proposed to the original fuzzy regression model are based on the constraints in Eq. (2.5b), which can be rewritten as (see Appendix B)

$$h \leq 1 - \frac{\left| y_i - \left(c_0 + \sum_{j=1}^{k} c_j x_{ij} \right) \right|}{s_0 + \sum_{j=1}^{k} s_j |x_{ij}|} \quad \text{for } i = 1, \ldots, N. \tag{2.6}$$

He et al. (2007) denote by h_i the right hand side of Eq. (2.6). For each observation i, h_i represents the grade of membership for an observed y_i belonging to the estimated $\hat{\tilde{Y}}_i$. Thus, the average of the h_i values $\bar{h} = \sum h_i / N$ can be interpreted as an overall measure of model fit, just as R^2 measures the model fit in OLS regressions. The revised fuzzy regression model is then obtained by choosing $h = 0$ in Eq. (2.6) and by adding

$$d_i = \left| y_i - \left(c_0 + \sum_{j=1}^{k} c_j x_{ij} \right) \right| \tag{2.7}$$

to the objective function, Eq. (2.5a), of the minimization problem. This additional term forces the optimization procedure to account for a better overall model fit measured by $\bar{h}$.

$$\text{Minimize} \quad \sum_{i=1}^{N} \left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}| \right] + \sum_{i=1}^{N} d_i \tag{2.8a}$$

subject to

$$0 \leq h_i = 1 - \frac{d_i}{s_0 + \sum_{j=1}^{k} s_j |x_{ij}|} \quad \forall i = 1, \ldots, N \tag{2.8b}$$

$s_j \geq 0, \quad \forall j = 0, 1, \ldots, k.$

This revised fuzzy regression model is still a linear program.

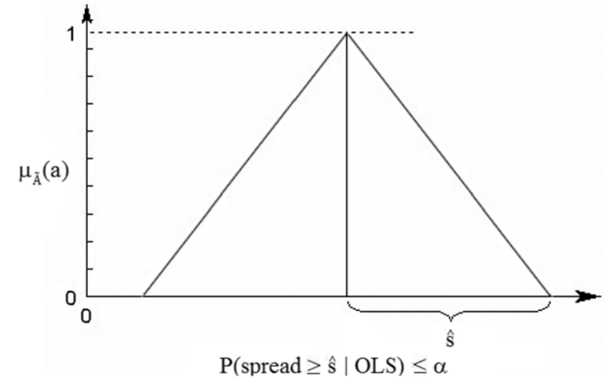

Fig. 1. Interpreting the regression coefficient spread as a test statistic for fuzziness.

2.2. A test for the fuzziness of regression coefficients

One big advantage of the possibilistic fuzzy regression model is that we can interpret the size of the spread of each of the fuzzy regression coefficients as a measure for how fuzzy the relationship between the corresponding independent variable and the dependent variable is without having any further *a priori* assumptions about the independent variables. If the spread s_j of the fuzzy coefficient $\tilde{A}_j$ in Eq. (2.2) is equal to zero then $\tilde{A}_j$ is a crisp number, and there is a direct linear relationship between the corresponding explanatory variable x_j and the dependent variable Y. If the spread s_j is greater than zero then $\tilde{A}_j$ is a fuzzy number. At first glance, we might conclude that the functional relationship between x_j and Y is fuzzy. But we have to keep in mind that the possibilistic fuzzy regression model does not have a random error term (see Eq. (2.2)). When fitting such a model to a dataset, some of the spreads have to be positive to capture the variation inherent in the data. This is conceptualized in Fig. 1, where $\mu_{\tilde{A}}(a)$ represents the membership function of a fuzzy regression coefficient, a denotes its domain, $\tilde{A}$ indicates a symmetrical triangular fuzzy number, and $\hat{s}$ is its empirical spread. The question becomes how likely it is to obtain the empirical spreads, given that the true data generating process is a classical regression model with a random error term, i.e. $Y = X\beta + \varepsilon$. To answer this question, we compare the spreads $\hat{s}_j, j = 1, \ldots, k$, estimated with the dataset under consideration, with the spread levels that can be obtained by estimating a possibilistic fuzzy regression model using hypothetical data simulated from a classical regression model with a random error term. Only if the former are significantly larger than the latter is the notion of fuzziness meaningful for an estimated regression coefficient, $\hat{\tilde{A}}_j$. On the basis of this concept, we derive k empirical test distributions, one for each spread $s_j, j = 1, \ldots, k$.

In addition, we derive an empirical test distribution for the goodness of fit measure $\bar{h}$ of the possibilistic regression model. We are interested in whether the possibilistic regression model is a better fit to our dataset than to data generated from a classical regression model with a random error term. If we can reject the null hypothesis that a goodness of fit measure $\bar{h}$ greater than or equal to the one obtained for the dataset under consideration could have been achieved for data generated by a classical regression relationship with a random error term, then we can conclude that our original dataset exhibits characteristics which fit better to a data generating process with fuzzy relationships than to one with a random error term. In the remainder of this paper, we will refer to such a dataset as "being generated by a process with fuzzy relationships", or as "exhibiting fuzzy relationships".

The test for the regression coefficients and the test for the goodness of fit measure $\bar{h}$ can be viewed as being complementary.

While the test for the fuzzy regression coefficients examines each coefficient separately, the test for the goodness of fit measure takes all fuzzy regression coefficients into account simultaneously.

Fig. 2 outlines our test procedure as well as the algorithm for computing the test distributions. The following section discusses the basic concepts underlying this algorithm. Note that we use the symbol $*$ as a superscript whenever we refer to simulated vectors and matrices. For example, Y^*, X^* and ε^* refer to simulated data whereas Y and X refer to subsets of the original data.

To obtain the empirical test distributions for $s_j, j = 1, \ldots, k$, and $\bar{h}$, we use a simulation approach similar to Deutsch (1992). In the first step, we estimate the OLS regression model $Y = X\beta + \varepsilon$ with the original dataset. This gives us the least squares estimator $\hat{\beta}$. Furthermore, we determine the empirical standard deviation $\hat{\sigma}$ of the OLS residuals, the empirical distribution of the exogenous variables in the column vectors $X_2, \ldots, X_{k+1}$ of the matrix X as well as their empirical correlation matrix Σ. In a second step, we generate 100,000 independent scenarios. Each scenario consists of an $N \times (k+1)$ matrix X^* and an $N \times 1$ vector Y^* derived from the linear model

$$Y^* = X^*\hat{\beta} + \varepsilon^* \tag{2.9}$$

where $\hat{\beta}$ is the $(k+1) \times 1$ vector of OLS estimates, ε^* is an $N \times 1$ vector of realizations from a Normal distribution with zero mean and standard deviation $\hat{\sigma}$ equal to the empirical standard deviation of the OLS residuals, and X^* is an $N \times (k+1)$ matrix with elements equal to 1 in the first column and elements in the other k columns bootstrapped from the empirical distributions of the corresponding variables while imposing the correlation structure Σ observed in the data.[5] In a third step, we estimate the possibilistic fuzzy regression model for each of the 100,000 scenarios consisting of X^* and Y^* by solving the linear program (2.5) or (2.8) and, hence, derive empirical distributions for the spread of each regression coefficient as well as for the $\bar{h}$ statistic.

Our use of correlated empirical distributions in the scenario generation process varies from Deutsch's (1992) original approach which assumes independently and uniformly distributed explanatory variables. Deutsch's (1992) robustness checks show that the distributional assumption for the explanatory variables does not affect the derived test distribution for Quandt's log likelihood ratio. However, the distributional assumption for the explanatory variables does indeed affect the derived test distributions for the spreads of the fuzzy regression coefficients. For example, comparing the 0.9-quantiles of the test distributions based on beta(0.25, 0.25) distributed explanatory variables with the 0.9-quantiles of the test distributions based on beta(4, 4) distributed explanatory variables reveals substantial differences of up to 225%. Similarly, comparing the 0.9-quantiles of the test distributions based on independently and uniformly distributed explanatory variables with the 0.9-quantiles based on correlated explanatory variables reflecting the correlation structure of the data reveals differences of up to 37%. To avoid any distributional misspecification of the explanatory variables, we employ the bootstrapping procedure outlined above.

In summary, the main idea underlying our test for the fuzziness of a regression coefficient is to show that the coefficient picks up more uncertainty than the level that can be explained by a random error. Thus, we generate test distributions by estimating a possibilistic fuzzy regression model with 100,000 datasets simulated with the classical regression model $Y^* = X^*\hat{\beta} + \varepsilon^*$ with an independent and normally distributed error term ε^*. Rejecting the null hypotheses for a fuzzy regression coefficient, hence, tells us that the regression coefficient captures more uncertainty than the levels that can be explained by an independent and normally distributed error term. This statement, however, is only meaningful if the original dataset can be well described with a linear regression model that has an independent and normally distributed error term. If on the other hand, an OLS regression with the original dataset exhibits heteroscedasticity or not normally distributed residuals, the test distributions are misspecified. The general approach is to generate the fuzzy test distributions on the basis of a regression model $Y^* = X^*\hat{\beta} + \varepsilon^*$ that fits well to the dataset. However, since there are various types of heteroscedasticity and different kinds of non-normality, there is no one-size-fits-all solution. In our case study on solvency surveillance in Germany, we estimate a classical OLS regression and examine the characteristics of its residuals first (see Section 5), and then derive two additional sets of test distributions reflecting these characteristics (see Section 6).

3. Solvency surveillance in Germany

In this section, we introduce the solvency surveillance example which we will use to demonstrate our test for the fuzziness of regression coefficients. Thus, Section 3.1 discusses the regulatory environment in Germany, Section 3.2 gives a brief overview of the insurer solvency surveillance literature, and Section 3.3 explains the variable selection for the financial strength prediction model based on the literature as well as the German environment.

3.1. The German solvency regulation

Solvency supervision in Germany primarily targets individual insurance companies. Since the German regulatory law does not allow one legal entity to write both life and health insurance and property–casualty business, holding companies contain separate, fully owned legal entities for the different forms of business. The German insurance authority (*Bundesanstalt für Finanzdienstleistungsaufsicht* or BaFin), thus, also monitors the solvency of insurer groups. However, the main focus of solvency surveillance remains on the company level since each insurance company can go bankrupt individually, and group members are not automatically obliged to back the distressed insurer.[6]

The German regulatory law in conjunction with the Solvency Ordinance (*Kapitalausstattungsverordnung*) explicitly specifies a level of equity capital that insurance companies are required to hold,[7] and this required capital level is a function of the insurer's

5 More precisely, to construct X, we first generate N random samples from a multivariate uniform distribution with correlation matrix Σ equal to the empirical correlation matrix of the data. We then use the inverse of each of the k empirical distribution functions to transform the correlated uniformly distributed variables over the range (0, 1) to correlated variables following the observed empirical distributions. Finally, we add the first column with all elements equal to 1 to the matrix X.

6 Usually holding companies help their distressed subsidiaries. However, the Mannheimer holding for example was not able to support its life insurance subsidiary whose capital base was depleted in 2003. The Mannheimer holding and its property–liability subsidiary are still in operation; the life insurance subsidiary is run off, now.

7 The description of the German regulatory environment throughout this paper assumes that the Directive 2002/13/EC of the European Parliament and of the Council of March 5, 2002, amending Council Directive 73/239/EEC as regards the solvency margin requirements for non-life insurance undertakings is in effect for all German property–liability insurance companies. Germany has incorporated this directive into the regulatory law together with a transition period. Therefore, for most insurance companies the new capital requirements are only binding as of January 1, 2007. However, all insurers have to report to the BaFin according to the new capital requirements; if an insurance company fails to meet the new requirements prior to 2007, it is sufficient to prove that the company meets the slightly laxer previously applied capital requirements (2004 yearbook of the BaFin).

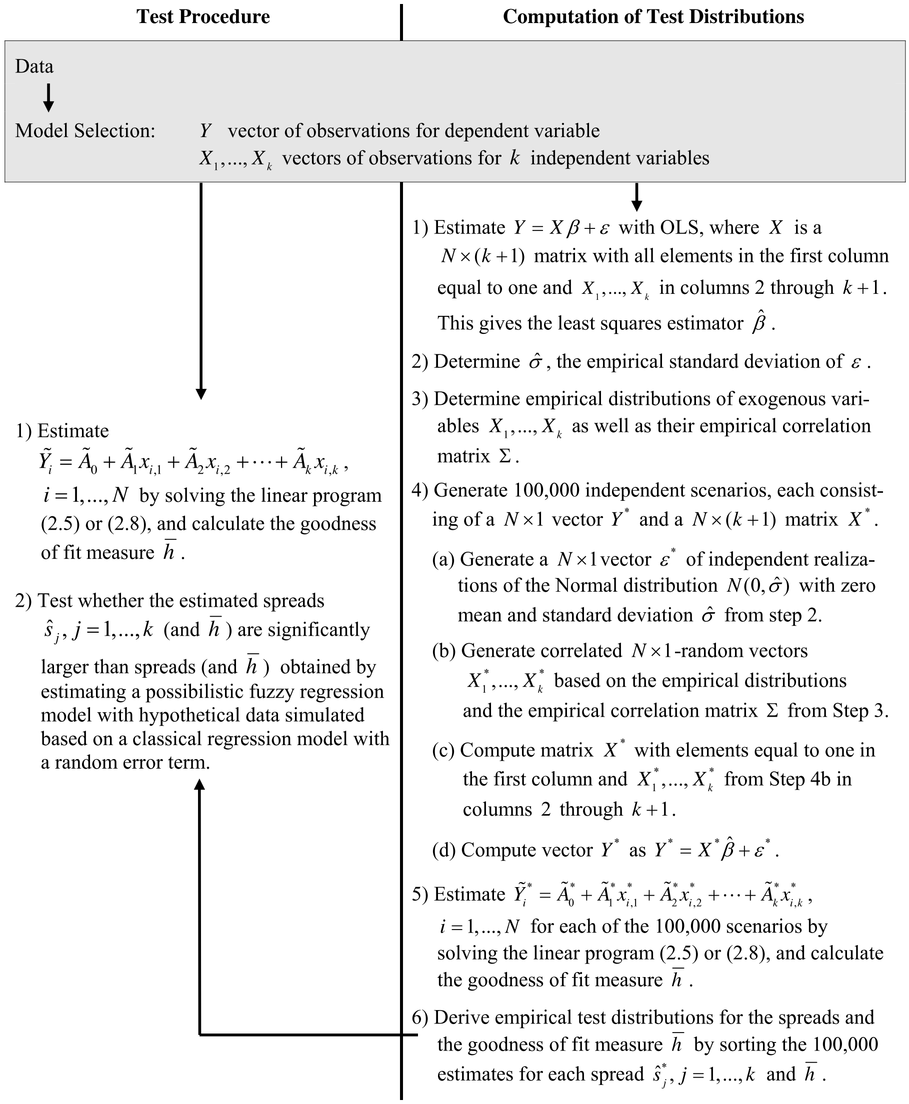

Fig. 2. Test procedure and computation of test distributions.

underwriting risk. The BaFin is responsible for monitoring whether the amount of equity capital actually shown on the balance sheet of an insurance company meets the required level.[8] Therefore, all insurance companies are required to report their equity capital to the BaFin on the basis of the German accounting principles. If an insurance company's equity capital is below the required level, the BaFin will ask the company to prepare a business plan for restoring the insurer's solvency. Whenever an insurance company fails to provide a convincing action plan to restore its solvency, the company will be declared bankrupt and will be run off.

Usually, German insurance companies hold much more capital than required. The average ratios of actual equity capital to required equity capital of all property–liability insurance companies supervised by the BaFin for the years 2002, 2003 and 2004 are 337%, 346% and 286% respectively.[9] Thus, solvency surveillance in Germany means ranking relatively financially healthy insurance companies for further regulatory scrutiny. Since German insurers have to provide the financial information for the previous accounting year to the BaFin by the end of July,[10] the BaFin would need a two-year-ahead prediction model to prioritize its on-site inspections for the following year. Thus, all independent variables in our study have a two-year lag.

A solvency or financial strength measure in line with the German regulatory law is the ratio of equity capital held by an insurance company to the amount of equity capital required for this

8 For the purpose of solvency monitoring, an insurer's equity capital consists of the sum of the paid-in capital stock, additional paid-in capital, retained earnings, profit-sharing rights outstanding and subordinate debt minus expenditure for the start-up or the expansion of business operations, goodwill of the company, and deferred taxes shown on the asset side of the balance sheet, and minus the net loss for the year if applicable. In addition, insurers can file an application with the BaFin to include 50% of the not paid-in capital stock as well as the hidden reserves in their investments in the calculation of equity capital for solvency purposes. However, it is only possible to include 50% of the not paid-in capital stock in the calculation if the paid-in part of the capital stock is greater than or equal to 25%.

9 Due to the high capital holdings, insolvencies are rare events in Germany. We are aware of only one insolvency of a property–liability insurance company since 1951. The failed company was a small insurer that specialized in the transportation insurance business line.

10 According to Section 15 of the Reporting Ordinance (*Verordnung über die Berichterstattung von Versicherungsunternehmen gegenüber der Bundesanstalt für Finanzdienstleistungsaufsicht* or BerVersV), property–liability insurance companies have to file financial information classified as *Nachweisung* 240, 241, 263 and 264 within five months after the end of the accounting year. Financial information classified as *Nachweisung* 242, 243, 244, 246 and 250 has to be filed within seven months after the end of the accounting year.

specific insurer. From a theoretical point of view, market data based financial strength measures should draw a more up-to-date picture. However, such measures are not available or not reliable for the following two reasons: First, not all insurers are stock companies. Second, solvency regulation focuses on the company level and most property–liability stock insurers are part of a holding company structure. Thus, the individual insurers are either 100% subsidiaries and not listed at a stock exchange, or they are listed but the free float is very small.

3.2. Previous literature

The literature on insurer insolvency prediction is extensive. While the early studies in this area focused on financial characteristics of insurance companies as insolvency predictors (BarNiv and McDonald, 1992), more recent research has either examined the predictive power of additional measures such as market characteristics (Browne and Hoyt, 1995; Browne et al., 1999), risk based capital (Cummins et al., 1995), NAIC FAST scores (Grace et al., 1998), and financial strength ratings (Pottier and Sommer, 2002) or applied more advanced modeling techniques like neural networks (Brockett et al., 1994) and cash flow simulations (Cummins et al., 1999). The vast majority of studies has focused on the US insurance industry with some exceptions: Kramer (1996) examines the financial solidity of Dutch property–liability insurance companies, Chen and Wong (2004) analyze both property–liability and life insurance companies in Malaysia, Singapore and Taiwan, and Sharpe and Stadnik (2007) explore factors associated with financial distress of Australian property–liability insurance companies. For our purpose, all of these previous studies have one key point in common: They use insurer's accounting data to predict their financial strength one or two years ahead.

3.3. A financial strength prediction model for German property–liability insurers

The empirical model used as an example to demonstrate our test for fuzziness of regression coefficients is based on previous insolvency literature and is similar to the one employed by Pottier and Sommer (2005) to study insolvencies of US property–liability insurance companies.[11,12] Our dependent variable measures the financial strength of an insurance company. We use the natural logarithm of the solvency ratio as defined by the German regulatory law: The equity capital held by an insurance company divided by the amount of equity capital required for the company.[13]

Changes in the financial strength of a company usually happen incrementally over time. Thus, we expect the financial strength of an insurance company today to be a strong predictor of the financial strength of that company two years in the future, and we include the two-year lagged log solvency ratio in our model.

An important indicator of an insurance company's solvency is its capital holding. Therefore, we include a measure of capitalization in our model and expect this variable to have a positive relationship with an insurer's solvency (Pottier and Sommer, 2002). Specifically, we use the ratio of equity capital as shown on the balance sheet to total assets.[14]

A measure of underwriting leverage is also included in the model. This variable is defined as net premiums written divided by equity capital as shown on the balance sheet. On the one hand a high underwriting leverage could make it too challenging to fulfill future claim obligations (Pinches and Trieschmann, 1977; Pottier and Sommer, 2005); on the other hand leverage can magnify the return on equity resulting in a higher surplus. Therefore, we do not have a clear expectation about the sign of this variable.

We also include a measure of insurer profitability in our model. It is defined as the ratio of net income before taxes to total assets. We expect insurer profitability to be positively associated with financial strength and solvency (Kahane et al., 1986; MacMinn and Witt, 1987; Sharpe and Stadnik, 2007).

To examine the effect of business diversification on financial strength, we include a line of business Herfindahl index variable in the model. This measure is calculated on the basis of gross premiums earned and takes on values between zero and 100% with higher values representing more concentration in the insurance business.[15] On the one hand, diversification provides a natural hedge against adverse developments in one line of business, leading to a lower probability of ending up in financial distress (Sommer, 1996). On the other hand, diversification exacerbates agency costs because managerial monitoring and bonding is more difficult in more complex firms (Hoyt and Trieschmann, 1991; Tombs and Hoyt, 1994; Elango et al., 2008; Liebenberg and Sommer, 2008). Increased agency costs reduce a firm's profitability and, thus, its financial strength. Therefore, we do not have a clear expectation about whether diversification has a positive or negative effect on an insurer's financial strength.

We also expect the characteristic of an insurance company's investment portfolio to affect its financial strength. On the one hand, the riskiness of a company's investments should increase its ruin probability; on the other hand the higher expected return associated with a riskier investment portfolio increases the profitability of the company and, hence its financial strength. Following the approach of Cummins and Nini (2002), we use the proportion of investments held in stock and real estate as a proxy for investment risk. However, we do not have a clear expectation about the sign of this variable.

A variable capturing the size of the company is also included in the model. All else being equal, bigger risk pools should produce

11 We use an established model framework to provide an example in line with the literature. We also experimented with Chen and Wong's (2004) model. However, the variables employed by Chen and Wong may capture the situation in emerging economies like Malaysia, Singapore and Taiwan well, but they have little explanatory power for the mature German insurance market. The Pottier and Sommer (2005) model on the other hand, has good explanatory power for our German dataset (see Section 5).

12 The following deviations from the Pottier and Sommer (2005) model are driven by institutional and regulatory differences between the German and the US insurance market: First, there are hardly any insurer insolvencies in Germany. Hence, we use a financial strength measure as dependent variable and not a dummy variable coded as 1 for insurer bankruptcies and 0 otherwise. Second, since we examine the financial strength of insurers and not "real" insolvencies, we are able to include the lagged financial strength measure in our model as a predictor variable. Third, because there is a third organizational form in Germany, namely public insurers, we include an additional dummy variable in our model identifying public insurers. Fourth, we do not include variables describing characteristics of the holding company or company group in our analysis. A financially strong group may support a distressed affiliated insurance company and, hence, financial characteristics of the group help explain ultimate insolvencies. The BaFin can ask distressed insurers for a business plan for restoring solvency, and such a plan may indeed include the infusion of capital from their parent company. But, for the purpose of detecting insurers which are about to get into financial troubles, measures capturing the characteristics of their parent company or group affiliation are not material.

13 We use the natural logarithm of the solvency ratio and not the solvency ratio itself because the solvency ratio is skewed to the right (Hair et al., 2006, p. 89 f.).

14 The balance sheet item A Equity Capital (Eigenkapital) consists of five sub-items and basically includes the capital stock, additional capital as well as various forms of retained earnings.

15 The Herfindahl index is defined as $\sum a_i^2/(\sum a_i)^2$, where a_i represents the gross premiums earned in business line i. The calculation uses premium data reported in the insurance companies' annual reports for the following 12 business lines: Personal Accident, Liability, Auto Liability, Other Auto, Fire, Homeowners Personal Property, Residential and Commercial Building Damage, Transportation and Aircraft, Legal Expenses, Credit and Collateral, Others, and Reinsurance.

Table 1
Summary statistics.

Variable	N	Mean	Std. dev.	Min	Median	Max
Solvency ratio 2004	114	360.04	415.23	69.93	246.65	3880.49
Log solvency ratio 2004	114	5.62	0.64	4.25	5.51	8.26
Log solvency ratio 2002	114	5.65	0.79	3.34	5.58	8.31
Capital/assets	114	21.16	9.71	3.87	18.88	66.55
Net premiums/capital	114	274.67	177.80	19.24	244.47	1526.28
Return on assets	114	0.92	6.54	−36.45	1.41	15.90
Line of business Herfindahl	114	40.72	29.75	11.86	30.17	100.00
(Stock + real estate)/assets	114	28.57	16.33	0.00	27.36	87.86
Natural log of assets	114	19.64	1.37	16.99	19.61	23.67
Mutual (%)	114	38.60				
Public (%)	114	14.04				

Note: The solvency ratio 2004 and the natural log of the solvency ratio 2004 are based on 2004 data; all other variables are based on the year 2002. Mutual is a dummy variable equal to 1 if the insurance company is a mutual insurer. Public is a dummy variable equal to 1 if the insurance company is a public insurer. Natural log of assets is based on the insurers' total assets in euros. All other variables are reported in per cent. All euro values are inflation adjusted with 2004 as the basis year.

less volatile claim payments. Consistently with this expectation, Cummins et al. (1995), Grace et al. (1998), Chen and Wong (2004), and Sharpe and Stadnik (2007) find a significant positive relationship between size and insurer solvency. Therefore, we include the natural log of total assets as size variable in our model.

Previous insolvency studies have found that mutual insurers have lower ruin probabilities than stock insurers (Cummins et al., 1995; Grace et al., 1998) because mutuals tend to focus on less risky business lines (Lamm-Tennant and Starks, 1993) and have less incentives to increase their risk exposure after policies are issued (Garven and Pottier, 1995). However, Liebenberg and Sommer (2008) find that mutual insurers are significantly less profitable than stock insurers, resulting in a comparative disadvantage regarding their financial strength. To capture any differences in financial strength between mutual and stock insurers, our model includes a dummy variable equal to 1 for mutual insurance companies and 0 for all others.

In addition to mutual insurance companies and stock insurance companies, there is a third organizational form operating in the German market, namely public insurance companies. The function of public insurance companies was originally to serve a specific region by providing reliable insurance coverage especially in the fire, the residential and commercial buildings, and the homeowners personal property business lines. Thus, being financially strong is in line with their focus on providing reliable coverage. However, since they are owned by local and state authorities which can bail them out with taxpayers' money, they might have an incentive to write more business per euro of capital.[16] Therefore, we do not have a clear expectation about whether public insurers are on average financially stronger or weaker than their private counterparts. To capture any effect of the public organizational form, our model includes a dummy variable equal to 1 for public insurance companies and 0 for all others.

4. Data

In our empirical example, we use company level data of property–liability insurance companies supervised by the German insurance authority (BaFin). We restrict our analysis to insurance companies which have gross premiums written of at least 50 million euros per year for the years 2002–2004. There are 114 such insurance companies, and these insurance companies account for 92.25% of the overall premium volume of the German property–liability insurance market. The dataset for the insurers in our sample is obtained from their annual reports, and includes the natural logarithm of the solvency ratios as defined by the German regulatory law for the year 2004 and the predictor variables described in the previous section for the year 2002.

Table 1 contains summary statistics for our dataset, and Table 2 provides the corresponding correlation coefficients. For the 114 insurance companies in our sample, the mean solvency ratio is 360% in 2004 indicating that these companies hold on average about 3.6 times the equity capital required by the regulatory law.

Table 3 presents differences between financially strong and weak insurance companies for all predictor variables. We classify the group of insurers with a solvency ratio below the median solvency ratio as financially "relatively weak" and all other insurers as financially "relatively strong". Table 3 provides means and medians for these two groups and for each variable. Asterisks indicate significant differences between financially weak and strong insurers based on t-tests for means and nonparametric k-sample tests for the equality of medians.[17] In addition to analyzing the complete sample of all 114 insurance companies, we also examine the two subsets of multi-line insurance companies and specialized insurance companies. To avoid classifying insurance companies with one major line of business and some negligibly small premium volume in other lines as a multi-line insurance company, we use a Herfindahl index of 50% as cut-off between the two groups. Insurers with a Herfindahl index of 50% or higher are classified as specialized insurers and insurers with a Herfindahl index below 50% as multi-line insurers.

Overall, the results presented in Table 3 indicate that financially weak insurers have on average less capital, write more business per capital, have a lower return on assets, are less diversified, invest less in risky assets like stocks and are smaller. The organizational form also seems to be able to explain differences in the solvency scores of insurance companies. A Pearson χ^2-test shows that many more public insurance companies belong to the group of financially strong insurers than to the group of financially weak insurers. Furthermore, we can see in Table 3 that the two subsets of multi-line insurers and specialized insurers have different characteristics. For multi-line insurers, for example, we can find a significant difference of the mean and median of the *return*

[16] This view is supported by the recent Fannie Mae and Freddie Mac disaster in the US and the following bailout. The downturn of the US housing market hit Freddie and Fanny's loan and loan guarantees business hard. On July 30, 2008, President George W. Bush signed a bill into law that gives the US Treasury Department unlimited power through 2009 to lend money to Fannie and Freddie or to buy their stock if needed.

[17] It tests the null hypothesis that the two samples of financially strong and weak insurance companies were drawn from populations with the same median. The test is performed by first computing the median score for all observations combined. Then all observations are compared to this overall median and classified as being either above or below the median. This classification is done for the strong and the weak insurers separately, resulting in a 2 × 2 contingency table. Finally, a Pearson chi-squared test is performed (Snedecor and Cochran, 1989, 124 ff.).

Table 2
Correlation coefficients.

Variable		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Log solvency ratio 2002	(1)								
Capital/assets	(2)	0.568***							
Net premiums/capital	(3)	−0.560***	−0.588***						
Return on assets	(4)	0.094	0.071	−0.103					
Line of business Herfindahl	(5)	−0.243***	0.046	0.000	0.134				
(Stock + real estate)/assets	(6)	0.197**	0.120	−0.044	−0.089	−0.054			
Natural log of assets	(7)	0.450***	−0.138	−0.106	0.087	−0.394***	0.315***		
Mutual dummy variable	(8)	0.039	0.103	−0.078	0.113	0.030	−0.214**	−0.183*	
Public dummy variable	(9)	0.275***	0.237**	−0.255***	−0.067	−0.187**	0.242***	0.107	−0.320***

* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

Table 3
Comparison of financially strong and weak insurance companies.

Variable	All insurers		Multi-line insurers		Specialized insurers	
	Strong	Weak	Strong	Weak	Strong	Weak
Log solvency ratio 2002						
Mean	6.12	5.17***	6.21	5.32***	5.79	4.81***
Median	6.16	5.22***	6.22	5.26***	5.63	4.96***
Capital/assets						
Mean	24.67	17.64***	24.73	17.00***	25.04	18.34
Median	22.85	16.90***	23.79	15.98***	19.57	16.38
Net premiums/capital						
Mean	201.96	347.38***	192.74	342.22***	229.89	367.21***
Median	188.79	324.58***	174.10	310.58***	217.42	331.91**
Return on assets						
Mean	0.94	0.90	2.24	−0.96***	0.56	2.72
Median	1.85	0.71	2.15	0.37***	2.04	1.37
Line of business Herfindahl						
Mean	33.37	48.07***	24.57	24.33	79.54	89.36
Median	25.36	36.31	20.36	20.82	83.90	100.00*
(Stock + real estate)/assets						
Mean	33.56	23.59***	32.22	24.92**	34.69	21.85*
Median	32.22	20.82***	31.96	21.73***	34.15	15.81***
Natural log of assets						
Mean	20.09	19.19***	20.24	19.59**	19.31	18.45***
Median	20.05	19.05***	20.26	19.61	19.31	18.38**
Mutual (%)	33.33	43.86	40.48	41.46	25.00	40.00
Public (%)	22.81	5.26***	30.95	4.88***	0.00	6.67
N	57	57	42	41	16	15

Note: The criterion for splitting the sample insurer into multi-line and specialized insurers is having a Herfindahl index smaller than 50% vs. having a Herfindahl index greater than or equal to 50%, respectively. Insurers with a solvency ratio below the median solvency ratio in 2004 are classified as "weak", and all others as "strong". Asterisks indicate significant differences between strong and weak insurers based on t-tests for means, a nonparametric k-sample test for medians and a χ^2-test for dummy variables.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

on assets variable, but not for specialized insurers. These results suggest that multi-line insurers and specialized insurers should be examined separately.

5. OLS and FGLS regression analysis

The goal of the following empirical analysis is to provide an example of how our test procedure finds a significantly fuzzy relationship between an independent variable and the dependent variable in a regression model. However, before we can test for fuzziness, we need to derive the test distributions, and the first step in deriving the test distributions is to estimate the OLS regression model $Y = X\beta + \varepsilon$ (see Fig. 2). Therefore, we start our empirical analysis with simple OLS regressions of the *log solvency ratio* variable on the insurers' characteristics outlined in Section 3.3. The residuals of these regressions are significantly heteroscedastic for the complete sample of all 114 property–liability insurance companies as well as for the sub-samples of multi-line insurance companies ($N = 83$) and specialized insurance companies ($N = 31$).[18] Therefore, Table 4 present OLS estimates with White-corrected standard errors as well as feasible generalized least squares (FGLS) estimates.

The FGLS estimates are based on the assumption that the variance of the error term ε_i is proportional to the squared *log solvency ratio* 2002 variable in the following way:

$$V(\varepsilon_i|x_i) = \sigma_i^2 = \sigma^2 z_i^2 \quad \text{for } i = 1, \ldots, N, \tag{5.1}$$

where z_i denotes the *log solvency ratio* 2002 for insurer *i*. This assumption is justified because the F statistic for the regressions of the squared OLS residual ε^2 on the squared log solvency ratio

18 We follow the Breusch–Pagan/Cook–Weisberg test procedure and regress the squared OLS residuals on the squared *log solvency ratio* 2002 variable (see Table 7 in Appendix C). The F statistic for the complete sample is 27.324 ($p = 0.000$) and, hence, significant at the 1% level. The F statistic for the sub-sample of multi-line insurance companies is significant at the 1% level as well ($F = 17.459$, $p = 0.000$). The F statistic for the sub-sample of specialized insurance companies is 7.210 ($p = 0.012$) and, hence, significant at the 5% level.

Table 4
OLS and FGLS regressions of log solvency ratios on insurer characteristics.

Variable	All insurers		Multi-line insurers		Specialized insurers	
	OLS	FGLS	OLS	FGLS	OLS	FGLS
Intercept	1.638**	1.919***	1.342*	1.588**	0.322	0.533
	(1.99)	(2.88)	(1.79)	(2.50)	(0.18)	(0.30)
Log solvency ratio 2002	0.468***	0.408***	0.400**	0.415***	0.284*	0.216
	(3.44)	(6.30)	(2.56)	(6.09)	(1.87)	(1.62)
Capital/assets	0.991	0.927*	2.710***	2.094***	−0.587	−0.113
	(1.10)	(1.66)	(2.73)	(3.05)	(−0.68)	(−0.15)
Net premiums/capital	−0.035	−0.045*	0.020	0.002	−0.112*	−0.126*
	(−1.11)	(−1.94)	(0.78)	(0.11)	(−1.83)	(−1.72)
Return on assets	−0.051	−0.113	1.792**	1.610**	−1.589***	−1.408**
	(−0.07)	(−0.22)	(2.02)	(2.42)	(−3.31)	(−2.34)
Line of business Herfindahl	−0.060	−0.120	0.376	0.180	−0.646**	−0.773**
	(−0.51)	(−0.99)	(0.91)	(0.47)	(−2.13)	(−2.74)
(Stock + real estate)/assets	0.357	0.343	0.626*	0.566**	0.763**	0.780**
	(1.30)	(1.58)	(1.92)	(2.17)	(2.35)	(2.15)
Natural log of assets	0.058*	0.065*	0.056	0.051	0.232**	0.242***
	(1.80)	(1.83)	(1.54)	(1.52)	(2.27)	(2.83)
Mutual dummy variable	−0.013	−0.032	0.035	0.061	0.009	−0.000
	(−0.18)	(−0.44)	(0.49)	(0.80)	(0.07)	(−0.00)
Public dummy variable	0.058	0.037	0.085	0.108		
	(0.51)	(0.31)	(0.86)	(0.93)		
R^2	0.690	0.678	0.772	0.776	0.822	0.819
Number of observations	114	114	83	83	31	31

Note: The dependent variable in the regressions is the natural logarithm of the solvency ratio for each individual insurer, based on 2004 data. All independent variables are based on the year 2002. The criterion for splitting the sample insurer into multi-line and specialized insurers is having a Herfindahl index smaller than 50% vs. having a Herfindahl index greater than or equal to 50%, respectively. T-statistics appear in parentheses. T-statistics of the OLS regressions are based on White-corrected standard errors.
* Significant at the 10% level.
** Significant at the 5% level.
*** Significant at the 1% level.

2002 is significant at the 1% level for the overall sample and the sub-sample of multi-line insurers, and significant at the 5% level for the sub-sample of specialized insurers (see Table 7 in Appendix C).[19] Thus, the FGLS estimator for the model $y = X\beta + \varepsilon$ can be obtained via the following steps:

1. Transform all observations to obtain
$$y_i/z_i = \beta_0/z_i + \beta_1(x_{i1}/z_i) + \cdots + \beta_k(x_{ik}/z_i) + \varepsilon_i/z_i. \qquad (5.2)$$
2. Running OLS on the transformed model yields the FGLS estimator $\hat{\beta}^{TM}$ for β.

We also test whether the regression residuals follow a Normal distribution. For the complete sample, both the Shapiro–Wilk and the Shapiro–Frankia tests reject the null hypothesis of normally distributed errors at the 1% level ($p < 0.001$ and $p < 0.001$). For the sub-sample of multi-line insurance companies, both the Shapiro–Wilk and the Shapiro–Frankia tests reject the null hypothesis of normally distributed errors at the 1% level ($p < 0.001$ and $p < 0.001$). However, for the sub-sample of specialized insurance companies, both the Shapiro–Wilk and the Shapiro–Frankia tests do *not* reject the null hypothesis of normally distributed errors ($p = 0.620$ and $p = 0.440$).

In summary, we find that the regression residuals are not normally distributed for some model specifications, and that these regression residuals are heteroscedastic. There are multiple kinds of heteroscedasticity. In our example, we specifically find that the variance of the residuals is proportional to the squared *log solvency ratio* 2002 variable. Since these findings violate the assumption of independent and normally distributed errors underlying our test (see Section 2.2), we develop two additional sets of test distributions in the following section, explicitly taking the characteristics of the observed residuals into account.

19 The F statistic tests the null hypothesis that all regression coefficients excluding the constant are zero.

6. Additional test distributions

As outlined in Section 2.2, our test for the fuzziness of regression coefficients is based on the null hypothesis that the data generating process is a classical regression relationship with an independent and normally distributed error term. In our example, however, the OLS regression residuals are heteroscedastic and deviate from the Normal distribution for some model specifications (see Section 5). To show that our test results are not just driven by misspecified test distributions, we generate two additional sets of test distributions for slightly different null hypotheses. The second set of test distributions drops the normality assumption and only assumes an independent and identically distributed error term. The third set of distributions still assumes normality but relaxes the independence assumption by introducing the kind of heteroscedasticity observed in our example dataset (see Section 5).

The computation for the second set of test distributions is quite similar to the one for the first set. The only difference is that when simulating the scenarios according to Eq. (2.9), the empirical distribution of the OLS residuals is used to draw realizations of ε_i^*, $i = 1, \ldots, N$, and not a Normal distribution.

The third set of test distributions is based on the assumption that the data generating process exhibits the kind of heteroscedasticity characterized in Eq. (5.1). Note that we use the superscript *TM* to identify variables from the transformed model in Eq. (5.2), and the symbol $*$ as superscript whenever we refer to simulated vectors and matrices rather than components of the original dataset. We first generate 100,000 independent scenarios of the transformed regression model in Eq. (5.2). Each scenario consists of the linear model

$$Y^{TM^*} = X^{TM^*}\hat{\beta}^{TM} + \varepsilon^{TM^*} \qquad (6.1)$$

where $\hat{\beta}^{TM}$ is the $(k+1) \times 1$ vector of FGLS estimates, ε^{TM^*} is an $N \times 1$ vector of realizations form a Normal distribution with zero mean and standard deviation equal to the empirical standard deviation of the FGLS residuals, and X^{TM^*} is an $N \times (k+1)$ matrix

Table 5
Standard possibilistic regression and test for fuzziness of explanatory variables.

Variable	All insurers					Multi-line insurers					Specialized insurers			
	Center	Spread	Test 1	Test 2	Test 3	Center	Spread	Test 1	Test 2	Test 3	Center	Spread	Test 1	Test 3
Intercept	1.629	0.000				1.815	0.000				0.623	0.000		
			(0.833)	(0.768)	(0.842)			(0.564)	(0.447)	(0.625)			(1.000)	(1.000)
Log solvency ratio 2002	0.064	0.023				0.121	0.098	∗∗∗	∗∗	∗∗∗	0.228	0.000		
			(0.237)	(0.234)	(0.462)			(0.002)	(0.011)	(0.002)			(1.000)	(1.000)
Capital/assets	3.250	1.861	∗∗∗	∗	∗∗	4.599	0.000				−1.137	0.000		
			(0.007)	(0.051)	(0.010)			(0.361)	(0.361)	(0.479)			(0.862)	(0.868)
Net premiums/capital	−0.054	0.036				0.044	0.021				−0.225	0.000		
			(0.293)	(0.367)	(0.113)			(0.386)	(0.401)	(0.229)			(1.000)	(1.000)
Return on assets	1.186	0.000				2.311	0.000				−1.499	0.000		
			(0.704)	(0.720)	(0.701)			(0.461)	(0.473)	(0.453)			(0.913)	(0.914)
Line of business Herfindahl	0.245	0.007				0.458	0.000				−0.778	0.283		∗
			(0.461)	(0.446)	(0.344)			(0.519)	(0.498)	(0.436)			(0.110)	(0.059)
(Stock + real estate)/assets	0.278	0.207				1.410	0.000				0.778	0.492	∗∗	∗∗
			(0.231)	(0.299)	(0.218)			(0.531)	(0.527)	(0.542)			(0.048)	(0.046)
Natural log of assets	0.154	0.000				0.079	0.000				0.258	0.000		
			(0.794)	(0.767)	(0.745)			(0.537)	(0.456)	(0.451)			(1.000)	(1.000)
Mutual dummy variable	−0.352	0.089				−0.039	0.000				0.076	0.000		
			(0.281)	(0.362)	(0.254)			(0.528)	(0.518)	(0.529)			(1.000)	(1.000)
Public dummy variable	−0.158	0.000				−0.156	0.000							
			(0.689)	(0.699)	(0.703)			(0.443)	(0.450)	(0.480)				
Model fit: $\bar{h}$		0.481					0.530	∗	∗	∗		0.409		
			(0.926)	(0.904)	(0.930)			(0.064)	(0.078)	(0.063)			(0.128)	(0.142)
Number of observations		114					83					31		

Note: The dependent variable is the log solvency ratio for each individual insurer, based on 2004 data. All independent variables are based on the year 2002. Test 1 for the fuzziness of regression coefficients as well as for the fuzziness of the overall sample is based on the null hypothesis that the data are generated by a regression model with an independent and normally distributed error term. Test 2 uses the empirical distribution of the residuals, and Test 3 assumes heteroscedasticity. The p-values of these tests appear in parentheses.
∗ Significance at the 10% level.
∗∗ Significance at the 5% level.
∗ ∗ ∗ Significance at the 1% level.

with elements equal to one divided by $z_i, i = 1, \ldots, N$ (see Eq. (5.2)), in the first column, and elements in the other k columns bootstrapped from the empirical distributions of the transformed variables $x_{ij}/z_i, i = 1, \ldots, N$, while imposing the correlation structure observed between these transformed variables. In a second step, we return to the original model $Y^* = X^*\hat{\beta} + \varepsilon^*$ by multiplying each element y_i^{TM} of Y^{TM*} as well as each element x_{ij}^{TM*} of X^{TM*} with $z_i, i = 1, \ldots, N$. By construction, the resulting model and its scenarios feature heteroscedastic errors. In a third step, we then estimate the possibilistic fuzzy regression model for each scenario consisting of Y^* and X^* and, hence, derive empirical distributions for the spread of each regression coefficient as well as for the $\bar{h}$ statistic.

7. Results

The possibilistic fuzzy regression results from Eq. (2.5) are presented in Table 5. The results from estimating the possibilistic regression model with endogenous h-certain factor given by Eq. (2.8) are presented in Table 6. While the centers of the estimated fuzzy coefficients can be interpreted similarly to standard regression coefficients, the spreads of the coefficients provide a measure for the vagueness or fuzziness of the relationship between the corresponding independent variables and the dependent variable. The difference between these two fuzzy regression models is essentially that the standard possibilistic regression minimizes the sum of the regression coefficients' spreads, whereas the revised version additionally maximizes the overall model fit measure $\bar{h}$ by readjusting the regression coefficients. Thus, a comparison between the estimates of these two fuzzy regression models reveals how robust the estimates are. Tables 5 and 6 also present the results of our three tests for the fuzziness of regression coefficients. The first test is based on the null hypothesis that the spread of the regression coefficients could have been obtained by estimating a possibilistic model with data generated by a classical regression model with an independent and normally distributed error term. The second test assumes only an independent and identically distributed error term, and the third test assumes normally distributed errors with heteroscedasticity. We run the first test for the complete sample as well as for the two sub-samples of multi-line and specialized insurance companies. The second and third tests are run depending on the outcome of the tests for normally distributed errors and homoscedasticity (see Section 5). Since we can reject the null hypothesis of normally distributed residuals as well as the null hypothesis of homoscedasticity for the OLS regression on the complete sample, we present results for all three tests for the complete sample. The same is true for the sub-sample of multi-line insurance companies. For the specialized insurance companies, we cannot reject the null hypothesis of normally distributed errors, but we can reject the null hypothesis of homoscedasticity. Thus, we present results for tests number 1 and 3 for specialized insurers.

Columns 2 and 3 of Tables 5 and 6 show the estimated fuzzy regression coefficients for the complete sample of all property–liability insurance companies. The two-year lagged *log solvency ratio*, the *capital to assets ratio*, the *return on assets*, the *line of business Herfindahl*, the (*stock + real estate*)/*assets* and the *natural log of assets* variables are positively associated with insurers' financial strength. The *net premiums/capital* and the *mutual dummy* variables are negatively associated with insurers' financial strength. These results are in line with the theoretical predictions outlined in Section 3.3. The effect of the *public dummy* variable on our financial strength measure is unclear because the estimated center based on the standard possibilistic regression model has a negative sign (Table 5), but the estimated center based on the revised possibilistic regression model has a positive sign (Table 6). Six of the ten

Table 6
Revised possibilistic regression and test for fuzziness of explanatory variables.

Variable	All insurers					Multi-line insurers					Specialized insurers			
	Center	Spread	Test 1	Test 2	Test 3	Center	Spread	Test 1	Test 2	Test 3	Center	Spread	Test 1	Test 3
Intercept	2.191	0.000				1.507	0.000				0.655	0.000		
			(0.789)	(0.710)	(0.807)			(0.750)	(0.683)	(0.788)			(0.658)	(0.783)
Log solvency ratio 2002	0.163	0.033				0.123	0.108	∗∗∗	∗∗∗	∗∗∗	0.286	0.000		
			(0.203)	(0.204)	(0.431)			(0.001)	(0.009)	(0.001)			(0.659)	(0.666)
Capital/assets	2.298	1.480	∗∗	∗	∗∗	4.247	0.000				−0.791	0.000		
			(0.020)	(0.088)	(0.031)			(0.709)	(0.701)	(0.761)			(0.692)	(0.696)
Net premiums/capital	−0.061	0.028				0.013	0.002				−0.116	0.000		
			(0.331)	(0.392)	(0.141)			(0.545)	(0.528)	(0.406)			(0.758)	(0.653)
Return on assets	0.308	0.000				1.648	0.000				−1.720	0.000		
			(0.724)	(0.716)	(0.722)			(0.729)	(0.724)	(0.725)			(0.690)	(0.686)
Line of business Herfindahl	0.076	0.000				0.508	0.000				−0.589	0.268		∗
			(0.765)	(0.746)	(0.719)			(0.769)	(0.753)	(0.736)			(0.150)	(0.088)
(Stock + real estate)/assets	0.816	0.289				1.092	0.000				0.987	0.578	∗∗	∗∗
			(0.179)	(0.268)	(0.165)			(0.751)	(0.745)	(0.754)			(0.035)	(0.035)
Natural log of assets	0.101	0.000				0.104	0.000				0.211	0.000		
			(0.749)	(0.715)	(0.705)			(0.746)	(0.702)	(0.709)			(0.668)	(0.625)
Mutual dummy variable	−0.237	0.238				−0.061	0.000				0.032	0.000		
			(0.059)	(0.212)	(0.049)			(0.855)	(0.860)	(0.857)			(0.864)	(0.856)
Public dummy variable	0.067	0.000				0.055	0.000							
			(0.922)	(0.922)	(0.922)			(0.901)	(0.900)	(0.902)				
Model fit: $\bar{h}$		0.574					0.571					0.451		
			(0.325)	(0.677)	(0.340)			(0.134)	(0.270)	(0.130)			(0.389)	(0.434)
Number of observations		114					83					31		

Note: The dependent variable is the log solvency ratio for each individual insurer, based on 2004 data. All independent variables are based on the year 2002. Test 1 for the fuzziness of regression coefficients as well as for the fuzziness of the overall sample is based on the null hypothesis that the data are generated by a regression model with an independent and normally distributed error term. Test 2 uses the empirical distribution of the residuals, and Test 3 assumes heteroscedasticity. The *p*-values of these tests appear in parentheses.
∗ Significance at the 10% level.
∗∗ Significance at the 5% level.
∗∗∗ Significance at the 1% level.

regression coefficients in the standard possibilistic regression model are actually fuzzy numbers with a positive estimated spread, and five of the regression coefficients in the revised possibilistic regression model have a positive spread.

Most important for the purpose of our paper, however, is the fact that our new test for the fuzziness of regression coefficients detects a significantly fuzzy relationship between the *capital/assets* variable and the dependent variable. All three test variants in both models reject the null hypothesis that such a large spread could have been achieved by applying the possibilistic model to a dataset generated by a classical regression relationship with a random error term. One possible reason for this finding could be that the *capital to assets* variable as well as the dependent variable are calculated with data from the insurance companies' balance sheets and, hence, reflect the book value of the insurers' assets and not their market values. The additional source of imprecision or uncertainty in the data, stemming from the discretion managers have in presenting the financial situation of their company, should result in a fuzzy relationship between the two variables. The imprecision in the data due to management discretion is not necessarily random in nature. Managers usually align their reported numbers with their corporate goals; hence, the imprecision inherent in accounting data may systematically depend on the business strategy of the company.

The possibilistic regression results for the sub-sample of multi-line insurance companies are presented in Columns 7 and 8 of Tables 5 and 6. Like the aggregate sample of all insurers, the two-year lagged *log solvency ratio*, the *capital to assets ratio*, the *net premiums to capital ratio*, the *return on assets*, the *line of business Herfindahl*, the (*stock* + *real estate*)/*assets* and the *natural log of assets* variables are positively associated with insurers' financial strength, and the *mutual dummy* variable is negatively associated with insurers' financial strength. The effect of the *public dummy* variable on our financial strength measure is unclear because the estimated center of the corresponding regression coefficient has opposing signs in the two different regression models. Two of the ten estimated regression coefficients have a positive spread in the standard as well as in the revised possibilistic regression model; and the spread of the two-year lagged *log solvency ratio* variable is significantly larger than spreads which can be derived by applying the possibilistic model to a dataset generated by a classical regression relationship with a random error term. All three test variants in both models are at least significant at the 5% significance level and indicate that the relationship between the lagged solvency ratio and the dependent variable is fuzzy in the context of the regression model analyzed.

The $\bar{h}$ is 0.530 for the standard possibilistic regression and 0.571 for the revised one indicating a good overall model fit. The *p*-values shown in Tables 5 and 6 are based on our three tests against the null hypothesis that such a high $\bar{h}$ could have been achieved by applying the possibilistic model to a dataset generated by a classical regression relationship with a random error term. For the standard possibilistic regression, all three variants of the test reject the null hypothesis at the 10% significance level, indicating that the data exhibit characteristics which fit better to a data generating process with fuzzy relationships than to one with a random error term. However, the same is not true for the revised possibilistic regression model.

Columns 12 and 13 of Tables 5 and 6 present the estimated fuzzy regression coefficients for the sub-sample of specialized insurance companies. Consistently with our expectations, the two-year lagged *log solvency ratio*, the (*stock* + *real estate*)/*assets* ratio, the *natural log of assets* and the *mutual dummy* variables are positively associated with insurers' financial strength, and the *net premiums/capital* ratio and the *line of business Herfindahl* are negatively associated with insurers' financial strength. Surprisingly, the *capital to assets ratio* and the *return on assets* are negatively associated with insurers' financial strength as well. The spread of the regression coefficient for the (*stock*+*real estate*)/*assets* ratio is positive and significant at the five per cent level in both tests and both models. Thus, we can conclude that the relationship between this variable and the log solvency ratio is fuzzy in the context of the analyzed regression model. A possible explanation for this result is that the (*stock* + *real estate*)/*assets* variable as well as the dependent variable are calculated on the basis of accounting data which is subject to some managerial discretion. This additional source of imprecision or uncertainty in the data should result in a fuzzy relationship between the two variables, even if there was a direct functional relationship between the undistorted variables. Test 3 also indicates that the relationship between the *line of business Herfindahl* index and the log solvency ratio is significantly fuzzy at the 10% level in both models. However, this result is not supported by Test 1.

Note that the regressions for the sub-sample of specialized insurance companies do not include the *public dummy* variable. Since there is only one public insurer in the sub-sample of specialized insurers, including the *public dummy* in the regression model would allow the one public insurer to have a firm specific intercept. Interestingly, both possibilistic regression models seem to have difficulties with such a specification for some data structures. To detect numerical problems in the optimization underlying the possibilistic regressions, we included constraints for the centers of the regression coefficients. These constraints are chosen to be far outside the expected range of possible estimates and only become binding if some sort of "convergence problem" occurs. For the model specification with the *public dummy*, the lower bound for the center ($lb = -100$) was binding in approximately 37% of the 100,000 scenarios used to generate the test distributions. The test results do not deviate from the results based on the model specification without the *public dummy*: The coefficient of the (*stock* + *real estate*)/*assets* variable is significantly fuzzy in both tests and both models, and the coefficient of the *line of business Herfindahl* index is significantly fuzzy in Test 3, but not in Test 1. However, we do not feel comfortable presenting these results. Therefore, we present results for the model specification without the *public dummy* variable for the sub-sample of specialized insurance companies in this paper.

8. Conclusion

This paper develops a test for the fuzziness of regression coefficients. If a regression coefficient is significantly fuzzy then the functional relationship between the corresponding independent variable and the dependent variable is fuzzy. Such a fuzzy relationship violates one of the assumptions underlying many econometric models, namely the assumption that there is a clear functional relationship which is only distorted by a random error term.

Our test procedure builds on the possibilistic fuzzy regression models of Tanaka et al. (1982) and He et al. (2007). Possibilistic regression models often use symmetric triangular fuzzy numbers as regression coefficients; such fuzzy numbers can be characterized by a center and a spread. We interpret the spread of each of the regression coefficients as a statistic measuring how fuzzy the relationship between the corresponding independent variable and the dependent variable actually is. We then derive empirical test distributions for the spreads based on the null hypothesis that such spreads could have been obtained by estimating a possibilistic regression model with a dataset generated by a classical regression model with random errors.

Like the results of a t-test for significance of regression coefficients, the results of our test for the fuzziness of regression coefficients are model specific. Changing the regression model by adding or removing variables might also change the fuzzy test results.

To illustrate our test procedure, we provide an example of a regression model in which our test detects a significantly fuzzy regression coefficient. This example is a financial strength prediction model for German property–liability insurance companies. We use variables from the literature, adapt them to the German environment if necessary, and estimate the model with accounting data. Since the German accounting principles give the management of a corporation some discretion to align their reported numbers with their corporate goals, there is an additional source of uncertainty in the data which might explain the fuzziness that our test detects.

To address the fuzziness found in a dataset, future research could focus on developing hybrid regression models incorporating both fuzzy regression coefficients and a random error term. Such models capturing fuzziness and randomness simultaneously could deepen our understanding of the various kinds of uncertainty present in our world.

Acknowledgements

The authors thank Yong Bai for his research assistance and Leon Chen, Carolyn Dehring, David Eckles, Robert E. Hoyt, Steven Pottier and Steve Miller for helpful comments on this paper.

Appendix A. Linear programming representation and a sketch of the solution for Eq. (2.5)

A matrix representation of the constraints in Eq. (2.5b) is as follows:

$$\begin{pmatrix} \text{zeros}(k+1\times k+1) & -I_{k+1\times k+1} \\ -X & -(1-h)|X| \\ X & -(1-h)|X| \end{pmatrix} \begin{pmatrix} A_C \\ A_S \end{pmatrix} \le \begin{pmatrix} \text{zeros}(k+1\times 1) \\ -Y \\ Y \end{pmatrix}$$

where $\text{zeros}(k+1\times k+1)$ is a square matrix with zeros as elements, $I_{k+1\times k+1}$ is the $k+1\times k+1$ identity matrix also denoted by $\text{eye}(k+1)$, X is the $N\times k+1$ matrix of observations from the independent variables, $|X|$ refers to the absolute value of X, Y is the $N\times 1$ vector of observations from the dependent variable, and $\{A_C = (c_0, c_1, \ldots, c_k)' A_S = (s_0, s_1, \ldots, s_k)'\}$.

The goal of the optimization is to determine the unknown $N\times k+1$ vector $A = \begin{pmatrix} A_C & A_S \end{pmatrix}'$.

Given this notation, Eq. (2.5a) can be written as

$$\begin{aligned} \text{sum}\,(|X|\times A_S) &= \text{sum}\left(\begin{pmatrix} \underset{(N\times k+1)}{\text{zeros}} & \underset{(N\times k+1)}{|X|} \end{pmatrix} \bullet \begin{pmatrix} A_C \\ A_S \end{pmatrix}\right) \\ &= \text{sum}\left(\begin{pmatrix} \underset{(N\times k+1)}{\text{zeros}} & \underset{(N\times k+1)}{|X|} \end{pmatrix} \bullet A\right). \end{aligned}$$

Note: The "sum" command in Matlab (applied to a vector) adds up all the components of the vector; the sign "•" represents the standard matrix multiplication.

Appendix B

Eq. (2.6) is developed from (2.5b) as follows:

The two conditions

Table 7
OLS regressions of squared residuals on the squared log solvency ratio.

Variable	All insurers	Multi-line insurers	Specialized insurers
Intercept	−0.413***	−0.330***	−0.078
	(−3.85)	(−3.16)	(−1.43)
Log solvency ratio 2002 squared	0.017***	0.012***	0.005**
	(5.23)	(4.18)	(2.69)
F Statistic	27.324***	17.459***	7.210**
Prob > *F*	0.000	0.000	0.012
R^2	0.196	0.177	0.199
Adj. R^2	0.189	0.167	0.171
Number of observations	114	83	31

Note: The dependent variable in these OLS regressions is the squared residual from the OLS regression of the log solvency ratio on insurer characteristics presented in Table 4. The criterion for splitting the sample insurer into multi-line and specialized insurers is having a Herfindahl index smaller than 50% vs. having a Herfindahl index greater than or equal to 50%, respectively. *T*-statistics appear in parentheses.
* Significance at the 10% level.
** Significance at the 5% level.
*** Significance at the 1% level.

$$\begin{cases} c_0 + \sum_{j=1}^{k} c_j x_{ij} + (1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right] \geq y_i \\ c_0 + \sum_{j=1}^{k} c_j x_{ij} - (1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right] \leq y_i \end{cases}$$
$$\text{for } i = 1, \ldots, N,$$

can be summarized in the following equation:

$$c_0 + \sum_{j=1}^{k} c_j x_{ij} - (1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right] \leq y_i \leq c_0 + \sum_{j=1}^{k} c_j x_{ij} + (1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right]$$

which can be rearranged as follows:

$$-(1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right] \leq y_i - \left(c_0 + \sum_{j=1}^{k} c_j x_{ij}\right) \leq (1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right]$$

$$\Rightarrow \left| y_i - \left(c_0 + \sum_{j=1}^{k} c_j x_{ij}\right)\right| \leq (1-h)\left[s_0 + \sum_{j=1}^{k} s_j |x_{ij}|\right] \quad \text{for } i = 1, \ldots, N$$

$$\Rightarrow \frac{\left| y_i - \left(c_0 + \sum_{j=1}^{k} c_j x_{ij}\right)\right|}{s_0 + \sum_{j=1}^{k} s_j |x_{ij}|} \leq (1-h) \quad \text{for } i = 1, \ldots, N.$$

Solving for h results in (2.6).

Appendix C

See Table 7.

References

BarNiv, R., McDonald, J.B., 1992. Identifying financial distress in the insurance industry: a synthesis of methodological and empirical issues. Journal of Risk and Insurance 59, 543–574.

Brockett, P.L., Cooper, W.W., Golden, L.L., Pitaktong, U., 1994. A neural network method for obtaining an early warning of insurer insolvency. Journal of Risk and Insurance 61, 402–424.

Browne, M.J., Carson, J.M., Hoyt, R.E., 1999. Economic and market predictors of insolvencies in the life–health insurance industry. Journal of Risk and Insurance 66, 643–659.

Browne, M.J., Hoyt, R.E., 1995. Economic and market predictors of insolvencies in the property–liability insurance industry. Journal of Risk and Insurance 62, 309–327.

Buckey, J., 1987. The fuzzy mathematics of finance. Fuzzy Sets and Systems 21, 257–273.

Chang, Y.-H.O., Ayyub, B.M., 2001. Fuzzy regression methods—a comparative assessment. Fuzzy Sets and Systems 22, 245–269.

Chen, R., Wong, K.A., 2004. The determinants of financial health of Asian insurance companies. Journal of Risk and Insurance 71, 469–499.

Cummins, J.D., Derrig, R.A., 1993. Fuzzy trends in property–liability insurance claim costs. Journal of Risk and Insurance 60, 429–465.

Cummins, J.D., Grace, M.F., Phillips, R.D., 1999. Regulatory solvency prediction in property–liability insurance: risk-based capital, audit ratios, and cash flow simulation. Journal of Risk and Insurance 66, 417–458.

Cummins, J.D., Harrington, S.E., Klein, R., 1995. Insolvency experience, risk-based capital, and prompt corrective action in property–liability insurance. Journal of Banking and Finance 19, 511–527.

Cummins, J.D., Nini, G.P., 2002. Optimal capital utilization by financial firms: evidence from the property–liability insurance industry. Journal of Financial Services Research 21, 15–53.

Derrig, R.A., Ostaszewski, K.M., 1995. Fuzzy techniques of pattern recognition in risk and claim classification. Journal of Risk and Insurance 62, 447–482.

Deutsch, J., 1992. Linear regression under two separate regimes: an empirical distribution for Quandt's log likelihood ratio. Applied Economics 24, 123–127.

DeWitt, G.W., 1982. Underwriting and uncertainty. Insurance: Mathematics and Economics 1, 277–285.

Diamond, P., 1988. Fuzzy least-squares. Information Sciences 46, 141–157.

Dubois, D., Prade, H., 1980. Fuzzy Sets and Systems: Theory and Applications. Academic, New York.

Ebanks, B., Kanvowski, W., Ostaszewski, K.M., 1992. Application of measures of fuzziness to risk classification in insurance, in: Computing and Information, Proceedings of the ICCI'92 Fourth International Conference. pp. 290–291.

Elango, B., Ma, Y.-L., Pope, N., 2008. An investigation into the diversification–performance relationship in the US property–liability insurance industry. Journal of Risk and Insurance 75, 567–591.

Garven, J.R., Pottier, S.W., 1995. Incentive contracting and the role of participating rights in stock insurers. Journal of Risk and Insurance 62, 253–270.

Grace, M.F., Harrington, S.E., Klein, R.W., 1998. Risk-based capital and solvency screening in property–liability insurance: hypotheses and empirical tests. Journal of Risk and Insurance 65, 213–243.

Hair, J.F., Black, W.C., Babin, B.J., Anderson, R.E., Tatham, R.L., 2006. Multivariate Data Analysis, sixth ed. Pearson Prentice Hall, Upper Saddle River, NJ.

He, Y.-Q., Chan, L.-K., Wu, M.-L., 2007. Balancing productivity and consumer satisfaction for profitability: statistical and fuzzy regression analysis. European Journal of Operational Research 176, 252–263.

Hoyt, R.E., Trieschmann, J.S., 1991. Risk/return relationships for life–health, property–liability, and diversified insurers. Journal of Risk and Insurance 58, 322–330.

Kahane, Y., Tapiero, C.S., Jacques, L., 1986. Concepts and trends in the study of insurer's solvency. In: Cummins, J.D., Derrig, R.A. (Eds.), Financial Models of Insurer Solvency. Kluwer Academic Publishers, Norwell, MA.

Kandel, A., 1986. Fuzzy Mathematical Techniques with Applications. Addison-Wesley, Boston, MA.

Kramer, B., 1996. An ordered logit model for the evaluation of Dutch non-life insurance companies. De Economist 144, 79–91.

Lamm-Tennant, J., Starks, L.T., 1993. Stock versus mutual ownership structures: the risk implications. Journal of Business 66, 29–46.

Lemaire, J., 1990. Fuzzy insurance. ASTIN Bulletin 20, 33–55.

Liebenberg, A.P., Sommer, D.W., 2008. Effects of corporate diversification: evidence from the property–liability insurance industry. Journal of Risk and Insurance 75, 893–919.

MacMinn, R.D., Witt, R.C., 1987. A financial theory of the insurance firm under uncertainty and regulatory constraints. Geneva Papers on Risk and Insurance 12, 3–20.

Pinches, G.E., Trieschmann, J.S., 1977. Discriminant analysis, classification results, and financially distressed p-l insurers. Journal of Risk and Insurance 44, 289–298.

Pottier, S.W., Sommer, D.W., 2002. The effectiveness of public and private sector summary risk measures in predicting insurer insolvencies. Journal of Financial Services Research 21, 101–116.

Pottier, S.W., Sommer, D.W., 2005. On the use of group-level financial information in insurer solvency surveillance. Working Paper Presented at the World Risk and Insurance Economics Congress.

Sánchez, J., Gómez, A., 2003. Applications of fuzzy regression in actuarial analysis. Journal of Risk and Insurance 70, 665–699.

Shapiro, A.F., 2004a. Fuzzy regression and the term structure of interest rate revisited, in: Proceedings of the 14th International AFIR Colloquium, vol. 1. pp. 29–45.

Shapiro, A.F., 2004b. Fuzzy logic in insurance. Insurance: Mathematics and Economics 35, 399–424.

Sharpe, I.G., Stadnik, A., 2007. Financial distress in Australian general insurers. Journal of Risk and Insurance 74, 377–399.

Snedecor, G.W., Cochran, W.G., 1989. Statistical Methods, 8th ed. Iowa State University Press, Ames, IA.

Sommer, D.W., 1996. The impact of firm risk on property–liability insurance prices. Journal of Risk and Insurance 63, 501–514.

Tanaka, H., Uejima, S., Asai, K., 1982. Linear regression analysis with fuzzy model. IEEE Transactions on Systems, Man and Cybernetics 12, 903–907.

Tanaka, H., Watada, J., 1988. Possibilistic linear systems and their application to the linear regression model. Fuzzy Sets and Systems 27, 275–289.

Tombs, J.W., Hoyt, R.E., 1994. The effect of product lines focus on insurer stock returns, in: Proceedings of the International Insurance Association Annual Meeting, Madrid, Spain, pp. 331–337.

Zadeh, L.A., 1965. Fuzzy sets. Information and Control 8, 338–353.

Zimmermann, A., 1996. Fuzzy Set Theory and its Applications, third ed. Kluwer Academic Publishers, Boston, MA.