

Detecting Fuzzy Predictor Variables: The Case of Insurer Solvency Surveillance in Germany

Thomas R. Berry-Stölzle¹

Marie-Claire Koissi²

Arnold F. Shapiro^{3,*}

February 2008

¹ Terry College of Business, University of Georgia, 206 Brooks Hall, Athens, GA 30602, Tel.: +1 706 542 5160, Fax: +1 706 542 4295, trbs@terry.uga.edu.

² Department of Mathematics, Western Illinois University, 1 University Circle, Macomb, IL 61455, Tel. +1 309 298 2358, mk122@wiu.edu.

³ Smeal College of Business, Penn State University, University Park, PA 16802, Tel.: +1 814 865 3961, Fax: +1 814 865 6284, afs1@psu.edu.

* Please address all correspondence to Arnold F. Shapiro.

Detecting Fuzzy Predictor Variables: The Case of Insurer Solvency Surveillance in Germany

Abstract

All econometric prediction models build on the assumption that the possible outcomes of a random event are exactly specified and observable. In reality, however, information is often subjective, incomplete, or vague. A model framework explicitly capturing such vagueness is the fuzzy set theory introduced by Zadeh (1965). This paper combines the econometric and the fuzzy world by developing a test to detect fuzzy variables in regression relationships. We illustrate this methodology by a case study on solvency prediction for German property-liability insurance companies. Our empirical analysis based on accounting data reveals that three of our predictor variables are significantly fuzzy violating the assumptions underlying standard econometric models. For these fuzzy variables, a fuzzy model provides better predictive accuracy than standard econometric models.

1. Introduction

The standard way to derive an econometric forecast for an economic phenomenon consists of three steps. First, derive a theoretical factor model explaining the phenomenon. Second, estimate the statistical version of the model using past data. Third, use the most recent data available for the independent variables to calculate a predicted value for the dependent variable. All econometric prediction models ranging from the ordinary least squares (OLS) regression to more complex panel data models build on the axioms of probability theory introduced by Kolmogorov. These axioms rest on the assumption that all possible outcomes of a random event are exactly specified and observable. In reality, however, we often face the situation that information is subjective, incomplete, or vague.

A man who is 6'1" is generally accepted as tall. Compared to the 6'7" average of basketball players in the NBA he is relatively short, but he is not "not tall" either. Thus, whether someone is described as being tall depends on the point of view of the person making this statement. The source of imprecision in this example is the absence of a sharply defined classification criteria 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. Since then, fuzzy set theory has become an active research field for applied mathematicians, and now also provides various kinds of regression models. The main difference between these models and the standard statistical ones is that a different form of uncertainty is being modeled.

The first fuzzy regression model was introduced by Tanaka et al (1982). The starting point for Tanaka's approach is that the relationship between the endogenous and the exogenous variables is vague or fuzzy in many real world applications. Therefore, Tanaka et al (1982) model the endogenous variable in a regression relationship as well as the coefficients of the

exogenous variables 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 “fuzziness” of each of the regression coefficients as a measure for how fuzzy the corresponding variable in the regression model is without having any further *a priori* assumptions about this variable.

The purpose of this paper is to combine the econometric and the fuzzy world by deriving a methodology to detect fuzzy predictor variables in regression relationships. We propose to use the spread of the regression coefficients in the possibilistic fuzzy regression model as a statistic capturing the degree of fuzziness of the corresponding explanatory variable. We then derive an empirical test distributions 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 plain vanilla regression relationship with random errors. If we can reject this null hypothesis the corresponding explanatory variable can be considered fuzzy in nature. Thus, our test allows us to determine for which variables the notion of “fuzziness” is actually meaningful. We illustrate our argument by a case study on solvency prediction for German property-liability insurance companies.

The risk of insurer insolvency and, hence, the assessment of insurers’ solvency is of utmost concern to insurance regulators. In the private sector, financial strength ratings have emerged to aid consumers in evaluating the insolvency risk of insurers. Academic research has addressed this important issue by developing and evaluating a variety of insolvency prediction models. While some of these models have examined how macro economic and market conditions influence the probability of insolvency for all insurers (see, e.g., Browne and Hoyt, 1995;

Browne, Carson, and Hoyt, 1999), the vast majority of models have focused on identifying individual insurers at high risk of experiencing financial trouble (see, e.g., Chen and Wong, 2004; Cummins, Grace, and Phillips, 1999; Sharpe and Stadnik, 2007). All these models have in common that they use firm level accounting data to predict the financial situation of an insurance company one or two years ahead. We now argue that every accounting system whether it is designed for reporting to regulators or to investors gives the management of a corporation some discretion to align their reported numbers with their corporate goals. This inevitable imprecision in accounting data is not random in nature but depends on the business strategy of a company. Such vagueness in data is best addressed with a fuzzy regression framework (Chang and Ayyub, 2001). The additional source of uncertainty in accounting data due to management discretion obviously varies across different accounting systems. To underscore our point, we chose a country for our case study whose accounting system leaves considerable leeway for management choices: Germany.

Our analysis is divided into three sections. First, we develop a solvency prediction model for German property-liability insurance companies. In the absence of a significant number of insurer insolvencies in Germany, we focus on predicting the financial strength or solvency of insurance companies with an OLS regression model. Second, we re-estimate our solvency prediction model using the possibilistic fuzzy regression methodology, and third, we explicitly test for the fuzziness of each explanatory variables in the possibilistic fuzzy regression model with our newly developed test.

Our empirical analysis provides two main results. First, three of our eight variables used to predict the solvency of German property-liability insurance companies are significantly fuzzy, and, hence, our dataset violates the assumptions underlying standard econometric mod-

els. Second, the possibilistic fuzzy regression model provides better predictive accuracy than the standard OLS regression model for our dataset. Thus, using a possibilistic fuzzy regression model in combination with our test for the fuzzyness of regression coefficients supplementary to the standard regression models provide additional insights, and, hence, can be recommended for every empirical research project.

This study should be of interest to all parties using accounting data or other vague, imprecise, subjective or judgmental data in a statistical model. Our results indicate that the fuzzy coefficients of a possibilistic regression model provide a convenient measure for the vagueness of each predictor variable. Our test for the fuzzyness of explanatory variables can be used to verify whether these variables and, hence, the dataset as a whole meet the assumptions underlying modern probability theory. If a dataset is fuzzy the use of “crisp” statistical procedures is not justified. In addition, this study should be of interest to all parties concerned with insurer solvency in Germany. At least to our knowledge, we are the first to develop an empirical solvency prediction model for German property-liability insurance companies.

This paper proceeds as follows. The next section describes the case of solvency surveillance in Germany. This section includes a brief overview of the insurer solvency surveillance literature, discusses the regulatory situation in Germany, and explains the selection of our predictor variables. Section 3 presents our dataset. In Section 4, we provide an overview of the methodology used focusing especially on the fuzzy regression framework and our test for the fuzzyness of regression coefficients. Section 5 presents our empirical results, and the final section concludes and puts forth additional research topics.

2. The Case of Solvency Surveillance in Germany

2.1 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 risk-based capital (Cummins, Harrington and Klein, 1995), NAIC FAST scores (Grace, Harrington and Klein, 1998), and financial strength ratings (Pottier and Sommer, 2002) or applied more advanced modeling techniques like neural networks (Brockett et al, 1994) or cash flow simulations (Cummins, Grace and Phillips, 1999). The vast majority of studies has focused on the US insurance industry, however, there are some exceptions: Kramer (1996) examines the financial solidity of Dutch property-liability insurance companies, Chen and Wong (2004) analyze property-liability as well as 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, the key point about all these previous studies is that they use predictor variables based on accounting data in statistical procedures without examining whether the uncertainty in their data and variables is only due to randomness or whether vagueness inherent in their measures adds an additional source of uncertainty.

Since Zadeh (1965) introduced fuzzy logic as a model to explicitly describe vagueness of set membership, his approach has gained recognition and inspired 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. The main focus of most previous fuzzy regression applications was on improving the model's fit and overall predictive power, whereas our focus is on evaluating the quality of individual predictor variables.

In its comparison of OLS and fuzzy regression results, the general approach of this study is similar to that of He, Chan and Wu (2007), although their motivation for the comparison is grounded in the inability of the OLS methodology to produce consistent estimates for their specific data set. They study how a firm's productivity as well as the satisfaction of the firm's customers affect its profitability. Since they achieve consistent estimates and a better overall model fit with a fuzzy regression model, they conclude that the vagueness inherent in consumer satisfaction data causes the OLS methodology to produce inconsistent results. We take the idea of using the fuzzy regression framework as a diagnosis tool one step further and develop a test for the fuzziness of regression coefficients in a possibilistic fuzzy regression model. Since we can interpret the "fuzziness" of each of the regression coefficients as a measure for how fuzzy the corresponding variable in the regression model is, such a test helps us to identify fuzzy variables. In an application like the prediction of insurance companies' solvency,

such a procedure provides valuable additional information as to which of the predictor variables are rather vague in nature and should be interpreted carefully.

2.2 The German Solvency Regulation

The German regulatory law dates back to 1901,¹ and establishes a federal regulatory authority with comprehensive power to control insurance companies and their business. The main duty of the German insurance authority (*Bundesanstalt für Finanzdienstleistungsaufsicht* or BaFin) is to ensure that the interests of the insured are secured and that insurers can continually fulfill the obligations they assumed.² Therefore, the BaFin is responsible for the licensing process of all insurance companies in Germany as well as for the continuous supervision of these insurers.³ In its continuous supervision, the BaFin focuses on the compliance of insurance companies with the legal framework, and on solvency surveillance. In its 2005 yearbook, the BaFin announced its intention to start developing an early warning system to systematically prioritize its on-site inspections. To date, on-site inspections are scheduled by the BaFin based on the judgment of its staff members. Thus, our research has direct practical implications.

Solvency supervision 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, thus, also monitors the sol-

¹ Reichsgesetz über die private Versicherungsunternehmung, May 12, 1901, in: Reichsgesetzblatt 139. Today's Gesetz über die Beaufsichtigung der Versicherungsunternehmen (or Versicherungsaufsichtsgesetz) is based on this law with only modest changes.

² "... die Belange der Versicherten nicht ausreichend gewahrt oder die Verpflichtungen aus den Versicherungen nicht genügend aus dauernd erfüllbar dargetan sind." (§8 Abs. 1 Ziff. 2, VAG)

³ Since the completion of the Single Market, insurers from the European Union may conduct business in Germany without setting up a German entity, they fall under the supervision of their domestic supervisory body. However, European insurers not subject to German regulation account for less than 3 percent of the market (see, e.g., the 2004 yearbook of the BaFin).

vency 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 obligated to back the distressed insurer.

The German regulatory law in conjunction with the Solvency Ordinance (*Kapitalausstattungsverordnung*) explicitly specifies a level of equity capital insurance companies are required to hold.⁴ For property-liability insurance companies, the required level of capital is a function of the insurer's underwriting risk. More precisely, the required level of capital is the maximum of the "premium index" measure of underwriting risk, the "claims index" measure of underwriting risk, and an absolute Euro value.

The *premium index* is calculated as follows:

$$\text{Premium Index} = (0.18 \cdot P^{low} + 0.16 \cdot P^{high}) \cdot \max(L^{net} / L^{total}, 0.5),$$

where P^{low} is the part of an insurer's premium volume which is smaller than or equal to 50 million Euros, P^{high} is the part of an insurer's premium volume which is greater than 50 million Euros, L^{net} denotes losses incurred net of reinsurance, and L^{total} denotes total losses incurred. Note, that the term *premium volume* in this context refers to the maximum of an insurer's gross premiums written and gross premiums earned. For all liability lines except for auto liability, the premium volume has to be multiplied by 1.5 before entering the premium index calculation reflecting an additional 50% safety loading.

⁴ The following description of the German regulatory environment 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 this capital requirements are only binding as of January 1, 2007. However, all insurers have to report to the BaFin according to this capital requirements; if an insurance company fails to meet this 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).

The *claims index* for property-liability insurance companies is given by:

$$\text{Claims Index} = (0.26 \cdot L_{av}^{low} + 0.23 \cdot L_{av}^{high}) \cdot \max(L^{net} / L^{total}, 0.5),$$

where L_{av}^{low} denotes the part of an insurer's average losses incurred which is smaller than or equal to 35 million Euros, L_{av}^{high} denotes the part of an insurer's average losses incurred which is greater than 35 million Euros, L^{net} represents the losses incurred net of reinsurance, and L^{total} denotes total losses incurred. Average losses are calculated based on the loss experience of the last three years, in some lines dealing with catastrophic natural hazards based on the last seven years. Again, for all liability lines except for auto liability, the losses incurred have to be multiplied by 1.5 before entering the claims index calculation reflecting an additional 50% safety loading.

The absolute Euro value representing the minimum level of equity capital required for all insurance companies is equal to two million Euros for insurers offering property insurance, and equal to three million Euros for insurers offering liability or credit and collateral insurance. These values are as of January 1, 2004 and have to be adjusted if the European consumer price index increases more than 5%.

The German insurance authority BaFin monitors whether the amount of equity capital actually shown on the balance sheet of an insurance company meets the required level defined as the maximum of the *premium index*, the *claims index* and the absolute minimum Euro value. Therefore, all insurance companies are required to report their equity capital to the BaFin. In Germany, there is no special accounting system like the Statutory Accounting Principles for reporting to the regulatory authority. Therefore, the reported numbers are based on the German accounting principles used in the annual reports of insurance companies. 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.⁵ If an insurance company's equity capital is below the required level the BaFin will ask the company to prepare a business plan to restore the insurer's solvency. Such a plan may either focus on increasing the company's equity capital or on reducing the capital requirement for example by ceding some of its business. However, if the actual equity capital is less than one third of the required amount, only activities to increase the insurance companies equity capital are acceptable. 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 they are required to. The average ratio of actual equity capital to required equity capital of all property-liability insurance companies supervised by the BaFin for the years 2002, 2003, 2004 and 2005 is 337%, 346%, 286% and 255% respectively. Due to this 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 specialized on the transportation insurance business line. The goal of the BaFin is to prevent insurer insolvencies through solvency surveillance, which in Germany means ranking relatively financially healthy insurance companies for further regulatory scrutiny. Since German insurers have to provide the financial

⁵ 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%.

information for the previous accounting year to the BaFin by the end of July,⁶ the BaFin would need a two-year-ahead prediction model to prioritise its on-site inspections for the following year. A solvency measure in line with the German regulatory law would be the ratio of equity capital held by an insurance company to the amount of equity capital required for this specific insurer.

Insurance companies in Germany report their financial data to the BaFin based on the German accounting principles which are designed to focus more heavily on creditor protection than on providing information to shareholders. Thus, 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 types of investments on their balance sheet. The lower of cost or market principle which applies for example to stocks⁷ requires an insurer to use the minimum of the buy 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 some other stock. Such a portfolio rebalancing transaction makes the previously hidden asset values visible on the balance sheet. This example highlights that insurance managers in Germany have a high degree of discretion in how to present the financial situation of their company. This example also raises the question whether standard statistical models using financial

⁶ According to §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.

⁷ The lower of cost or market principle as stated in §253 (3) HGB applies to stock investments which are short term in nature (Umlaufvermögen). Long term investments fulfilling certain requirements (Anlagevermögen) are subject to §253 (2) HGB instead.

data reported by insurance companies as predictor variables can deal with this additional source of uncertainty – namely management discretion – inherent in the data.

2.3 Predictors of Property-Liability Insurer Solvency

The empirical model of this article is based on predictor variables used in previous studies. Our dependent variable measuring the financial strength of an insurance company is the solvency ratio 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. The set of predictor variables used in our model is similar to the one employed by Pottier and Sommer (2005) to study insolvencies of US property-liability insurance companies. However, 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. But the goal of the BaFin is to detect financially weak insurers at a earlier stage. The BaFin can than ask distressed insurers for a business plan to restore solvency, and such a plan may indeed include the infusion of capital from their parent company. Thus, for the purpose of detecting insurers which are about to get into financial troubles, measures capturing the characteristics of their parents company or group affiliation are not helpful.

The most 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.⁸

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).

All else equal, we expect more diversified insurers to have a lower probability of getting into financial distress (Sommer, 1996). To capture the risk-reduction effects associated with business diversification, we include a line of business Herfindahl index variable in the model. This measure is calculated based on gross premiums earned and takes on values between zero and 100% with higher values representing more concentration in the insurance business.⁹

⁸ 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.

⁹ 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.

We also expect the riskiness of an insurance company's investment portfolio to affect its ruin probability. 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 and expect this measure to be negatively related to insurer solvency. It is important to note that the percentage number we calculate is based on the corresponding balance sheet items, which might substantially deviate from market values.

A variable capturing the size of the company is also included in the model. All else equal, bigger risk pools should produce less volatile claim payments. Consistent with this expectation, Cummins, Harrington and Klein (1995), Grace, Harrington and Klein (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, Harrington and Klein, 1995; Grace, Harrington and Klein, 1998). Therefore, our model includes a dummy variable equal to 1 for mutual insurance companies and 0 for all others. Since 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), we expect mutuals to have a higher solvency score on average.

In addition to mutual insurance companies and stock insurance companies, there is a third organizational form operating in the German market: public insurance companies. Public insurance companies are traditionally more focused on personal lines which tend to be less risky than commercial lines. However, since they are owned by local authorities like states

which assume unlimited liability, they might have an incentive to write more business per Euro of capital. 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.

3. Data

In our assessment of the German insurers, 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 2002-2005. There are 114 insurance companies matching our selection criteria, and these insurance companies account for 92.25% of the overall premium volume of the German property-liability insurance market. The data for the insurers in our sample is obtained from their annual reports. We construct two separate samples. The first sample includes the solvency ratios as defined by the German regulatory law for the year 2004 as well as the predictor variables described in the previous section from the year 2002. The second sample includes the solvency ratios for the year 2005 and the predictor variables from 2003. We use the first sample to estimate our models and the second sample to validate the predictive accuracy of our models.

Table 1 contains summary statistics for our two datasets, and Table 2 provides the corresponding Spearman correlations. The mean solvency ratio is 381% in 2004 and 377% in 2005 indicating that the Germans hold on average about 3.8-times the equity capital required by the regulatory law. The mean for the capital to assets ratio is 21%, which is approximately

half the size of the 49% reported by Pottier and Sommer (2005) for 1796 US property-liability insurance companies based on 1998 Statutory Accounting data. The mean premiums to capital ratio is 273% which is about 2.7-times higher than the 102% reported by Pottier and Sommer (2005). These high differences indicate that the German Accounting framework differs substantially from the Statutory Accounting Principles used by US insurers to report to their regulators.

4. Methodology

Our analysis of factors predicting insurer solvency is straight forward. We first regress 2004 solvency scores on 2002 variables, and then use the estimated regression coefficients together with 2003 variables to predict 2005 solvency scores. We perform these two steps within a standard ordinary least squares (OLS) regression framework as well as within a possibilistic fuzzy regression framework and compare the predictive accuracy of the two methodologies. In addition, we test for each predictor variable whether it is fuzzy in nature. We use this sequence of procedures to analyze the complete sample of all 114 property-liability insurance companies as well as for the two sub-samples of multi-line insurance companies ($N = 83$) and specialized insurance companies ($N = 31$).

4.1 OLS Regression Model

Assume the linear model:

$$y = X\beta + \varepsilon, \tag{4.1}$$

where y is the $N \times 1$ vector of observations on the dependent variable, X is a $N \times (k + 1)$ Matrix of observation on k explanatory variables with elements equal to one in the first column,

β is a $(k+1) \times 1$ vector of unknown coefficients and ε is a $N \times 1$ vector of independent and normally distributed error terms with zero mean and variance σ^2 . If the inverse of $X'X$ exists the OLS estimator $\hat{\beta}$ for the coefficient vector β is given by

$$\hat{\beta} = (X'X)^{-1} X'y. \quad (4.2)$$

The predictive accuracy of an OLS regression model can be evaluated with two statistics which are both based on the forecast residuals $y_i^{observed} - \hat{y}_i^{predicted}$. The root mean squared error is defined as

$$RMSE = \sqrt{\frac{1}{N} \sum_i (y_i - \hat{y}_i)^2} \quad (4.2)$$

and the Theil U statistic is given by

$$U = \sqrt{\frac{(1/N) \sum_i (y_i - \hat{y}_i)^2}{(1/N) \sum_i y_i^2}}. \quad (4.3)$$

A convenient property of the Theil U statistic is that $(1-U)$ can be interpreted as an R-squared measure which is conveniently comparable across model frameworks.

4.2 The Possibilistic Fuzzy Regression Framework

4.2.1 Definitions and notations

Definition 1: (Zadeh, 1965) A fuzzy subset \tilde{A} of a reference set X is defined by $\tilde{A} = \{a, \mu_{\tilde{A}}(a) \mid a \in X\}$, where $\mu_{\tilde{A}} : X \rightarrow [0,1]$ is called the membership function of \tilde{A} . The membership function indicates the grade at which an element a belongs to the set \tilde{A} .

For a standard (crisp) set C , the membership of an element x in C is indicated by the indicator function $\chi_C(x)$ which takes on the value 1 if x belongs to C and 0 otherwise. In contrast, fuzzy sets allow partial membership. Hence $\mu_{\tilde{A}}(a) = 0.8$, for example, means that there is a possibility (0.8) that a might be in \tilde{A} . As expected, fuzzy sets are well suited to describe groups with vague boundaries, such as the group of “relatively small” people, the group of “high risk” firms or the set of “stable interest rates”. Let us have a closer look on another example, classifying the risk capacity of insurance companies (Shapiro, 2004b: 400-401). One can use fuzzy subsets of the numerical interval $X = [0, 100\%]$ to describe the groups of companies with “high risk capacity” and “low risk capacity”. Figure 1 presents the membership function $\mu_{high}(x)$ which assigns to each insurer a grade of membership in the group of “high risk capacity” companies. Insurers with risk capacity of 50% or less are assigned a membership grade of zero, and those with risk capacity in excess of 80% are assigned a grade of one. Between those risk capacities, (50%, 80%), the grade of membership is fuzzy.

Fuzzy sets are also characterized by their α -cuts: $\tilde{A}_\alpha = \{a \mid \mu_{\tilde{A}}(a) \geq \alpha\}$, $0 \leq \alpha \leq 1$. The support of \tilde{A} is $\tilde{A}_0 = \bigcup_{\alpha > 0} \tilde{A}_\alpha$. Membership functions can be triangular, trapezoidal, Gaussian, generalized bell or a combination of these basic classes (Shapiro, 2004b).

Definition 2: (Zimmermann, 1996) When the reference set X is the set of real numbers \mathbf{R} , the convex subset \tilde{A} is called a fuzzy number.

We limit the present study to symmetric triangular fuzzy numbers.

Definition 3: (Zimmermann, 1996) A triangular fuzzy number (TFN) $\tilde{A} = (a, l_A, r_A)$ with center $a \in \mathbf{R}$ and left and right spreads (l_A, r_A) is defined by the membership function

$$\mu_{\tilde{A}}(x) = \begin{cases} 1 - \frac{|a-x|}{l_A} & \text{if } a - l_A \leq x \leq a \\ 1 - \frac{|a-x|}{r_A} & \text{if } a \leq x \leq a + r_A \\ 0 & \text{otherwise} \end{cases} \quad (4.4)$$

where “ $| \cdot |$ ” denotes the absolute value function. The associated α -cuts are: $\forall \alpha \in [0,1]$, $\tilde{A}_\alpha = [A_\alpha^1, A_\alpha^2] = [x - l_A(1-\alpha), x + r_A(1-\alpha)]$. When $l_A = r_A$, the TFN is called a symmetric triangular fuzzy number and denoted by $\tilde{A} = (a, r_A)$.

Figure 2 presents an example of a triangular fuzzy number (TFN). The notion of α -cut allows the definition of the level of inclusion for fuzzy numbers.

Definition 4: Given two fuzzy numbers \tilde{A} and \tilde{B} with α -cuts $\tilde{A}_\alpha = [A_\alpha^1, A_\alpha^2]$ and $\tilde{B}_\alpha = [B_\alpha^1, B_\alpha^2]$ respectively,

- \tilde{A} is included in \tilde{B} at level α , that is $\mu(\tilde{A} \subseteq \tilde{B}) \geq \alpha$, if $\{B_\alpha^1 \leq A_\alpha^1 \text{ and } B_\alpha^2 \geq A_\alpha^2\}$ (Sánchez and Gómez, 2003: 668).
- \tilde{A} is included in \tilde{B} , that is $\tilde{A} \subseteq \tilde{B}$, if \tilde{A} is included in \tilde{B} at level α , for all $\alpha : \forall \alpha \in [0,1], \mu(\tilde{A} \subseteq \tilde{B}) \geq \alpha$.

4.2.2 The Standard Possibilistic Regression Model

The difference between a possibilistic fuzzy regression model and an OLS regression model is that in a fuzzy model, the regression coefficients as well as the dependent variable are fuzzy numbers. Thus, the fuzzy regression equation is given by

$$\tilde{Y}_i = \tilde{A}_0 + \tilde{A}_1 x_{i,1} + \tilde{A}_2 x_{i,2} + \cdots + \tilde{A}_k x_{i,k} \quad \text{for } i = 1, \dots, N. \quad (4.5)$$

Let c_j and s_j denote the center and spread of the fuzzy coefficient $\tilde{A}_j, j = 0, \dots, k$, respectively.

Then Equation (4.5) can be rewritten as

$$\tilde{Y}_i = (c_0, s_0) + (c_1, s_1)x_{i,1} + (c_2, s_2)x_{i,2} + \cdots + (c_k, s_k)x_{i,k} \quad \text{for } i = 1, \dots, N, \quad (4.6)$$

which leads to

$$\tilde{Y}_i = (c_0 + c_1 x_{i,1} + c_2 x_{i,2} + \cdots + c_k x_{i,k}, \quad s_0 + s_1 |x_{i,1}| + s_2 |x_{i,2}| + \cdots + s_k |x_{i,k}|). \quad (4.7)$$

The main characteristic of the possibilistic fuzzy regression is that equation (4.7) 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 should be included within a specified feasible data interval. This latter requirement can be formulated as $\mu(\tilde{Y}_{x,t} \subseteq \hat{Y}_{x,t}) \geq h$, where $h \in [0,1]$ is the so called h-certain factor or “degree of belief,” and can be 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 $h = 0$ (Moskowitz and Kim, 1993; Tanaka and Watada, 1988), 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:

$$\text{Minimize } \sum_{i=1}^N \left[s_0 + \sum_{j=1}^k s_j |x_{i,j}| \right] \quad (4.8a)$$

subject to

$$\begin{cases} s_j \geq 0 & \forall j = 0, 1, \dots, k \\ c_0 + \sum_{j=1}^k c_j x_{ij} + (1-h) \left[s_0 + \sum_{j=1}^k s_j |x_{ij}| \right] \geq Y_i & , i = 1, \dots, N \\ c_0 + \sum_{j=1}^k c_j x_{ij} - (1-h) \left[s_0 + \sum_{j=1}^k s_j |x_{ij}| \right] \leq Y_i \end{cases} \quad (4.8b)$$

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

4.2.3 The Possibilistic Regression Model with Endogenous h -Certain Factor

He, Chan and Wu (2007) developed a model which determines a degree of believe h_i for every observation i endogenously within the optimization. This model is expected to provide a better fit to the data. The changes He, Chan and Wu (2007) propose to the original fuzzy regression model are based on the constraints in (4.8b), which can be rewritten as (see Appendix II).

$$h \leq 1 - \frac{|y_i - (c_0 + \sum_{j=1}^k c_j x_{ij})|}{s_0 + \sum_{j=1}^k s_j |x_{ij}|} \quad \text{for } i = 1, \dots, N \quad (4.9)$$

He, Chan and Wu (2007) denote by h_i the right hand side of Equation (4.9):

$$h_i = 1 - \frac{|y_i - (c_0 + \sum_{j=1}^k c_j x_{ij})|}{s_0 + \sum_{j=1}^k s_j |x_{ij}|} \quad \text{for all } i = 1, \dots, N. \quad (4.10)$$

For each observation i , h_i represents the membership for an observed y_i belonging to the estimated \hat{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 similar to R^2 in OLS regressions. The revised fuzzy regression model is then obtained by choosing $h = 0$ in Equation (4.9) and by adding $d_i = |y_i - (c_0 + \sum_{j=1}^k c_j x_{ij})|$

to the objective function of the minimization problem (see Equation 4.8a). This additional term forces the optimization procedure to account for a better overall model fit measured by \bar{h} .

$$\text{Minimize } \sum_{i=1}^N \left(s_0 + \sum_{j=1}^k s_j |x_{i,j}| \right) + \sum_{i=1}^N d_i \quad (4.11a)$$

subject to

$$0 \leq h_i = 1 - \frac{d_i}{s_0 + \sum_{j=1}^k s_j |x_{ij}|} \quad \forall i = 1, \dots, N \quad (4.11b)$$

$$s_j \geq 0, \forall j = 0, 1, \dots, k$$

This revised fuzzy regression model is still a linear program.

4.2.4 Evaluating the Predictive Accuracy of Fuzzy Regressions

In analogy to other regression models, predictions for the dependent variable are derived from a fuzzy regression model by multiplying the independent variables from the holdout sample with the estimated regression coefficients. If we are interested in a point prediction, we will only need to calculate the predicted values for the center of the dependent variable. The difference between the observed vales of the independent variable and the predicted values for the center is referred to as the error $\varepsilon_i = y_i^{\text{observed}} - \hat{y}_i^{\text{center}}$ of the prediction. In addition, we can also calculate predicted values for the spread of the independent variable.

To measure the quality of the predictions, we suggest, in the spirit of the R-square, the following statistic

$$\bar{r} = 1 - \frac{\left| y_i^* - (c_0 + \sum_{j=1}^k c_j x_{ij}^*) \right|}{s_0 + \sum_{j=1}^k s_j |x_{ij}^*|} \quad (4.12)$$

where $\tilde{A}_j = (c_j, s_j)$ are the fuzzy coefficients obtained with the estimation sample, and where y_i^* and x_{ij}^* can either be observations from the hold out sample to evaluate the out-of-sample predictive accuracy, or observations from the estimation sample to evaluate the in-sample predictive accuracy. The \bar{r} takes on values between 0 and 1; bigger values represent better predictive accuracy.

4.3 A Test for Fuzzyness of Explanatory Variables in Regression Relationships

A 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 corresponding variable in the regression model is. If the spread s_j of the fuzzy coefficient \tilde{A}_j in equation (4.5) is equal to zero then \tilde{A}_j is a crisp number and the corresponding explanatory variable X_j can be considered to be crisp. If the spread s_j is greater than zero then \tilde{A}_j is fuzzy, and if s_j is big enough we can conclude that X_j is fuzzy in nature. But how big has s_j to be before we can conclude that X_j is fuzzy? We address this question by deriving k empirical test distributions, one for each spread $s_j, j = 1, \dots, k$. The null hypotheses we want to reject is that the spreads $s_j, j = 1, \dots, k$ could have been obtained by estimating a possibilistic regression model with crisp data, or in other words with data generated by a standard crisp regression relationship with random errors.

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 fits better to our dataset than to a standard crisp dataset. If this is the case we can con-

clude that our original dataset is fuzzy in nature. Thus, the null hypothesis we would like to reject is that the \bar{h} obtained by the possibilistic regression model and our original dataset can as well be achieved by estimating a possibilistic regression model with crisp data, or in other words with data generated by a standard crisp regression relationship with random errors.

To obtain the empirical test distributions for $s_j, j=1, \dots, k$ and \bar{h} , we use a simulation approach similar to Deutsch (1992) and Berry-Stölzle and Born (2007). We first derive the OLS-estimator $\hat{\beta}$ of the standard regression model (see equation 4.5). We then use this estimated vector of coefficients $\hat{\beta}$ and the variance of the residuals \hat{s}_ε^2 to generate 100,000 independent scenarios. Each scenario consists of a linear model

$$Y = X\hat{\beta} + \varepsilon$$

where ε is a $N \times 1$ normally distributed random vector with zero mean and $\sigma = \hat{s}_\varepsilon$, $\hat{\beta}$ is the $(k+1) \times 1$ vector of estimated coefficients, and X is a $N \times (k+1)$ matrix with elements equal to one in the first column and all other elements generated from the following random distributions: For all continuous variables, we use the uniform distribution with values between the minimum and the maximum value observed for this specific variable in the original dataset; and for all dummy variables, we use a Bernoulli trial which generates zeros and ones. The probability of getting a one in the Bernoulli trials is calibrated to reflect the percentage of ones in the original dataset. For each scenario we estimate the possibilistic fuzzy regression model and, hence, derive empirical distributions for the spread of each regression coefficient as well as for the \bar{h} statistic. If the spread of a fuzzy coefficient is bigger than or equal to the $(1-\alpha)$ -quantile of the test distribution, then the corresponding variable can be considered fuzzy in nature with a significance level of α . In analogy, if the estimated \bar{h} statistic measuring the

overall model fit of the possibilistic regression is bigger than or equal to the $(1 - \alpha)$ -quantile of the corresponding test distribution, then the possibilistic model fits better to our dataset than to one generated by a standard regression model with a significance level of α . Such a high \bar{h} statistic indicates that the underlying dataset can be considered fuzzy in nature.

5. Results

5.1 Univariate Analysis

This section discusses some univariate results for the predictor variables across various samples. For each sample, we split the insurance companies in two groups and classify the group of insurers with a solvency ratio below the median solvency ratio as financially “weak” and all other insurers as financially “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 medians. We perform these tests separately on the estimation sample based on 2002 variables and 2004 solvency scores as well as the hold-out sample based on 2003 variables and 2005 solvency scores. In addition to analyzing the complete set of all insurance companies, we also examine the two sub-sets 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 much 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 sub-sets of multi-line insurers and specialized insurers have different characteristics. For multi-line insurers, for example, we can find a significant difference in the mean and median of the *return on assets* variable, but not for specialized insurers. These results suggest to examine multi-line insurers and specialized insurers separately.

5.2 OLS Regression Analysis

Table 4 presents the OLS regression results for the complete sample of all 114 property-liability insurance companies as well as for the two sub-samples of multi-line and specialized insurers. The independent variable in the estimations is the solvency ratio for each individual insurer based on 2004 data. All independent variables are based on the year 2002. When analyzing the complete sample, two variables are found to be significant in predicting the financial strength of property-liability insurance companies (see Column 1 of Table 4). Firms with a higher capital to assets ratio in 2002 tend to be financially stronger in 2004, and larger insurers tend to be financially stronger. Both of these results are consistent with expectations. The R-squared of 0.414 and the (1-U) statistic for in sample predictive accuracy of 0.423 indicate a good overall model fit. The out-of-sample predictive accuracy of the model is quite good as

well. The (1-U) statistic for the holdout sample consisting of 2003 insurer characteristics and 2005 solvency scores is 0.454.

Columns 2 and 3 of Table 4 present the regression results for the two sub-sets of multi-line insurance companies and specialized insurance companies. The R-squared for these two sub-samples is 0.572 and 0.667 respectively, indicating a better overall model fit than for the aggregate sample. The in- and out-of-sample predictive accuracy also increase by splitting the overall sample. The (1-U) statistic for the in sample prediction is 0.502 for multi-line insurers and 0.630 for specialized insurers, and for the out-of-sample prediction, the (1-U) statistic is 0.556 for multi-line and 0.532 for specialized insurers. Three variables are found to be significant in predicting the solvency of multi-line insurance companies. Again, insurers with a higher capital to assets ratio in 2002 tend to be financially stronger in 2004, insurers with a higher net premiums to capital ratio, and larger insurers tend to be financially stronger. The result for the net premiums to capital ratio deviates from the univariate analysis according to which financially stronger insurers have a smaller premium volume per capital. But one should keep in mind that the positive effect of a higher premium volume on the solvency score occurs after controlling for the capital to asset ratio. The capital of an insurance company can be interpreted as its capacity to assume risks. Hence, for a given level of capital insurers which use their capacity seem to perform better resulting in a higher surplus and solvency score.

For the sub-sample of specialized insurance companies, three predictor variables are significant. Consistent with our expectation, bigger firms tend to be financially stronger. Insurance companies with a higher percentage of their investment portfolio in stocks and real estate tend to be financially stronger. The higher expected return associated with these risky assets

seems to pay off. Surprisingly, we find the return on asset variable to be negatively correlated with specialized insurers' solvency scores.

5.3 Fuzzy Regression Analysis

The possibilistic fuzzy regression results from equation (4.8) are presented in Table 5. The results from estimating the possibilistic regression model with endogenous h-certain factor given by equation (4.11) are presented in Table 6. While the centers of the estimated fuzzy coefficients can be interpreted similar to standard regression coefficients, the spreads of the coefficients provide a measure for the vagueness or fuzziness of the independent variables' impact on the dependent variable. The difference between these two fuzzy regression models is essentially that the standard possibilistic regression minimizes the sum of the regression coefficient's spreads in a "one size fits all" procedure, whereas the revised version simultaneously maximizes the overall model fit 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 test for fuzziness of predictor variables. The p-values in parentheses below the spreads are based on the null hypothesis that such a spread could have been achieved by applying the possibilistic model to a dataset generated by a plain vanilla regression relationship with random errors. If we can reject this null hypothesis the corresponding explanatory variable can be considered fuzzy in nature.

Columns 1 and 2 of Table 5 and 6 show the results for the complete sample of all property-liability insurance companies. The capital to assets ratio is positively associated with our solvency measure as expected. But most strikingly, the spread of the corresponding coefficient is positive and significant at the one percent level in both models. This result provides evidence

for an additional source of uncertainty in the data for the capital to asset ratio which cannot be explained by random errors. Thus, we can conclude that this variable is fuzzy in nature. We further argue that the additional source of uncertainty inherent in accounting data stems from the discretion managers have in presenting the financial situation of their company.

Three additional explanatory variables have a straight impact on the dependent variable; their spreads are zero in both the standard and the revised possibilistic regression. Insurance companies with a higher net premiums to capital ratio, insurance companies with a smaller percentage of investments in risky assets, and bigger insurance companies tend to be financially stronger. The coefficient for the public dummy variable has a positive center in the standard but not in the revised regression. Therefore, we conclude that there is no clear evidence for an on average higher solvency score of public insurance companies. In both regressions, the coefficient of the mutual dummy variable has a center of zero and the biggest spread of all estimated coefficients indicating that the relationship between the mutual organizational form and the solvency of an insurance company is fuzzy. However, this spread is not significant.

The \bar{h} is 0.609 for the standard possibilistic regression and 0.610 for the revised one indicating a good overall model fit. The p-value below the \bar{h} in Tables 5 and 6 are based on our test 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 plain vanilla regression relationship. For the standard possibilistic regression, we can reject this null hypothesis at the 10 percent significance level. Therefore, our dataset exhibits fuzzy characteristics and can be considered fuzzy in nature. For the revised possibilistic regression, however, we cannot reject this null hypothesis.

The results for the sub-sample of multi-line insurance companies (see columns 3 and 4 of Tables 5 and 6) are similar to the results for the aggregate sample of all insurers. The capital

to assets ratio is positively associated with insurer solvency in both possibilistic regression models. The spread of the corresponding coefficients are positive and significant at the one percent level indicating that the capital to assets ratio is fuzzy. Two additional variables have a straight impact on the dependent variable. Insurance companies with a higher net premiums to capital ratio and bigger insurance companies tend to be financially stronger. The coefficient for the *(stock + real estate)/assets* variable has a negative center in the standard but not in the revised regression. Therefore, we conclude that there is no clear evidence for a negative correlation between the percentage of risky assets in an insurance companies investment portfolio and its future solvency. In both regressions, the coefficient of the mutual dummy variable has a center of zero and the biggest spread of all estimated coefficients, and this spread is also significant for the standard possibilistic regression model. This result provides some evidence that that the relationship between the mutual organizational form and the solvency of a multi-line insurance company is unclear, vague or in other words: fuzzy. The \bar{h} of the standard possibilistic regression is 0.566, the \bar{h} of the revised possibilistic regression is 0.609 indicating a good model fit, and in both cases this model fit is significantly better for our dataset than for a dataset generated by a plain vanilla regression model with random errors.

Columns 5 and 6 of Tables 5 and 6 present the estimation results for the sub-sample of specialized insurance companies. There are four variables which influence the dependent variable in both regression models. Firms with a higher capital to asset ratio, firms with a lower return on assets, more diversified firms and bigger firms tend to be financially stronger. The coefficient for the *natural log of assets* variable also has a positive spread in both models providing some evidence for the fuzzyness of this variable. However, this spread is not significant. The coefficients for the *line of business Herfindahl* variable as well as for the *public dummy*

variable have non-zero centers in the standard possibilistic regression, but not in the revised possibilistic regression. Therefore, we conclude that there is no clear relationship between these two variables and our solvency measure. The coefficient for the *(stocks + real estate)/assets* variable, however, has its center at zero but a positive and significant spread. This result provides evidence for an additional source of uncertainty in the data for the *(stocks + real estate)/assets* ratio which cannot be explained by random errors. Hence, we can conclude that this variable is fuzzy. Since we use accounting data to calculate the variable, we argue that the additional imprecision in our variable stems from the discretion managers have in presenting the financial situation of their company. The \bar{h} measures of the standard as well as the revised possibilistic regression on the sub-sample of specialized insurance companies are 0.413 and 0.511, respectively, indicating a good overall model fit.

Figure 3 shows the out of sample predictions for the standard possibilistic regression model. The predictive accuracy of the fuzzy model measured with \bar{r} is 56% for the complete sample of all property-liability insurance companies. This compares favorably with a (1-U) statistic of 45% for the OLS regression.

6. Conclusion

All econometric prediction models ranging from the ordinary least squares (OLS) regression to more complex panel data models build on the assumption that all possible outcomes of a random event are exactly specified and observable. In reality, however, we often face the situation that information is subjective, incomplete, or vague. A model framework explicitly capturing such vagueness is the fuzzy set theory introduced by Zadeh (1965). Since then, fuzzy set theory has become an active research field for applied mathematicians, and now also pro-

vides regression models. The purpose of this paper is to combine the econometric and the fuzzy world by deriving a methodology to detect fuzzy predictor variables in regression relationships. We propose to use the spread of the regression coefficients in the possibilistic fuzzy regression model as a statistic capturing the degree of fuzziness of the corresponding explanatory variable. We then derive an empirical test distributions 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 plain vanilla regression relationship with random errors. If we can reject this null hypothesis the corresponding explanatory variable can be considered fuzzy in nature. Thus, our test allows us to determine for which variables the notion of “fuzziness” is actually meaningful. We illustrate our argument by a case study on solvency prediction for German property-liability insurance companies.

Standard solvency prediction models use firm level accounting data to predict the financial situation of a company one or two years ahead. Since every accounting system whether it is designed for reporting to regulators or to investors gives the management of a corporation some discretion to align their reported numbers with their corporate goals, predictor variables based on accounting data exhibit some imprecision or fuzziness. We develop a solvency prediction model for German property-liability insurance companies and explicitly test for the fuzziness of each explanatory variable. Our empirical analysis provides two main results. First, three of our eight variables used to predict the solvency of German property-liability insurance companies are significantly fuzzy, and, hence, our dataset violates the assumptions underlying standard econometric models. Second, the possibilistic fuzzy regression model provides better predictive accuracy than standard econometric models for our dataset.

In summary, the possibilistic fuzzy regression model in combination with our test for the fuzzyness of regression coefficients can be used as a tool to detect fuzzy variables. This diagnostic ability helps us to understand which predictor variables are reliable and which might result in rather vague predictions.

Appendix I

Linear Programming Representation and Sketch of Solution for Equation (4.6)

A matrix representation of the conditions (5.b) is as follows:

$$\begin{pmatrix} \mathit{zeros}(9 \times 9) & -I_{9 \times 9} \\ -X & -(1-h)|X| \\ X & -(1-h)|X| \end{pmatrix} \begin{pmatrix} A_C \\ A_S \end{pmatrix} \leq \begin{pmatrix} \mathit{zeros}(9 \times 1) \\ -Y \\ Y \end{pmatrix}$$

where $\mathit{zeros}(9 \times 9)$ is a square matrix with zeros as elements, $I_{9 \times 9}$ is the (9x9) identity matrix also denoted by $\mathit{eye}(9)$, X is the (114x9) matrix defined in Model 1, $|X|$ refers to the absolute value of X , Y is the (114x1) output vector, and $\{A_C = (c_0, c_1, \dots, c_8)'$ $A_S = (s_0, s_1, \dots, s_8)'\}$.

The unknown vector (18x1) to determine is $A = (A_C \ A_S)'$.

Then, the sum to minimize in Eq. (5.a) can be represented by

$$\mathit{sum}(|X| \times A_S) = \mathit{sum} \left(\begin{pmatrix} \mathit{zeros} & |X| \\ (114 \times 9) & (114 \times 9) \end{pmatrix} \times \begin{pmatrix} A_C \\ A_S \end{pmatrix} \right) = \mathit{sum} \left(\begin{pmatrix} \mathit{zeros} & |X| \\ (114 \times 9) & (114 \times 9) \end{pmatrix} \times A \right)$$

Note: the “sum” command in Matlab (applied to a vector) adds-up all the components of the vector; the sign “ \times ” represents the standard matrix multiplication.

Appendix II

Equation (4.9) is developed from (4.8b) as follows:

The two conditions

$$\begin{cases} c_0 + \sum_{j=1}^8 c_j x_{ij} + (1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right] \geq y_i \\ c_0 + \sum_{j=1}^8 c_j x_{ij} - (1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right] \leq y_i \end{cases} \text{ for } i=1, \dots, 114,$$

can be summarized in the following equation

$$c_0 + \sum_{j=1}^8 c_j x_{ij} - (1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right] \leq y_i \leq c_0 + \sum_{j=1}^8 c_j x_{ij} + (1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right]$$

which can be rearranged as follows:

$$\begin{aligned} & -(1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right] \leq y_i - (c_0 + \sum_{j=1}^8 c_j x_{ij}) \leq (1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right] \\ \Rightarrow & \left| y_i - (c_0 + \sum_{j=1}^8 c_j x_{ij}) \right| \leq (1-h) \left[s_0 + \sum_{j=1}^8 s_j |x_{ij}| \right] \quad \text{for } i=1, \dots, 114 \\ \Rightarrow & \frac{\left| y_i - (c_0 + \sum_{j=1}^8 c_j x_{ij}) \right|}{s_0 + \sum_{j=1}^8 s_j |x_{ij}|} \leq (1-h) \quad \text{for } i=1, \dots, 114. \end{aligned}$$

Solving for h results in (4.9).

References

- BarNiv, R., and J. B. McDonald, 1992, Identifying Financial Distress in the Insurance Industry: A Synthesis of Methodological and Empirical Issues, *Journal of Risk and Insurance*, 59: 543-574.
- Berry-Stölzle, T. R., and P. Born, 2007, The Effect of Regulation on Underwriting Cycles: the Case of Germany, Working Paper, University of Georgia.
- Brockett, P. L., W. W. Cooper, L. L. Golden, and U. Pitaktong, 1994, A Neural Network Method for Obtaining an Early Warning of Insurer Insolvency, *Journal of Risk and Insurance*, 61 (3): 402-424.
- Browne, M. J., J. M. Carson, and R. E. Hoyt, 1999, Economic and Market Predictors of Insolvencies in the Life-Health Insurance Industry, *Journal of Risk and Insurance*, 66 (4): 643-659.
- Browne, M. J., and R. E. Hoyt, 1995, Economic and Market Predictors of Insolvencies in the Property-Liability Insurance Industry, *Journal of Risk and Insurance*, 62 (2): 309-327.
- Buckey, J., 1987, The fuzzy mathematics of finance, *Fuzzy Sets and Systems*, 21: 257-273.
- Chang, Y.-H. O., and B. M. Ayyub, 2001, Fuzzy regression methods – A comparative assessment, *Fuzzy Sets and Systems*, 22 (3): 245-269.
- Chen, R., and K. A. Wong, 2004, The Determinants of Financial Health of Asian Insurance Companies, *Journal of Risk and Insurance*, 71: 469-499.
- Cummins, J. D., and R. A. Derrig, 1993, Fuzzy Trends in Property-Liability Insurance Claim Costs, *Journal of Risk and Insurance*, 60 (3): 429-465.
- Cummins, J. D., M. F. Grace, and R. D. Phillips, 1999, Regulatory Solvency Prediction in Property-Liability Insurance: Risk-Based Capital, Audit Ratios, and Cash Flow Simulation, *Journal of Risk and Insurance*, 66 (3): 417-458.
- Cummins, J. D., S. E. Harrington, and R. Klein, 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., and G. P. Nini, 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., and K. M. Ostaszewski, 1995, Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification, *Journal of Risk and Insurance*, 62 (3): 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 (3): 141-157.
- Dubois, D., and H. Prade, 1980, *Fuzzy Sets and Systems: Theory and Applications* (New York: Academic).
- Ebanks, B., W. Kanvowski, and K. M. Ostaszewski, 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.
- Garven, J. R., and S. W. Pottier, 1995, Incentive Contracting and the Role of Participating Rights in Stock Insurers, *Journal of Risk and Insurance*, 62: 253-270.

- Grace, M. F., S. E. Harrington, and R. W. Klein, 1998, Risk-Based Capital and Solvency Screening in Property-Liability Insurance: Hypotheses and Empirical Tests, *Journal of Risk and Insurance*, 65: 213-243.
- He, Y.-Q., L.-K. Chan, and M.-L. Wu, 2007, Balancing productivity and consumer satisfaction for profitability: Statistical and fuzzy regression analysis, *European Journal of Operational Research*, 176: 252-263.
- Kahane, Y., C. S. Tapiero, and L. Jacques, 1986, Concepts and Trends in the Study of Insurer's Solvency, in: Cummins, J. D., and R.A. Derrig, eds., *Financial Models of Insurer Solvency* (Norwell, MA: Kluwer Academic Publishers).
- Kandel, A., 1986, *Fuzzy Mathematical Techniques with Applications* (Boston, MA: Addison-Wesley).
- Kramer, B., 1996, An Ordered Logit Model for the Evaluation of Dutch Non-Life Insurance Companies, *De Economist*, 144: 79-91.
- Lamm-Tennant, J., and L. T. Starks, 1993, Stock versus Mutual Ownership Structures: The Risk Implications, *Journal of Business*, 66: 29-46.
- Lemaire, J., 1990, Fuzzy insurance, *ASTIN Bulletin*, 20 (1): 33-55.
- MacMinn, R. D., and R. C. Witt, 1987, A Financial Theory of the Insurance Firm under Uncertainty and Regulatory Constraints, *Geneva Papers on Risk and Insurance*, 12: 3-20.
- Moskowitz, H., and K. Kim, 1993, On assessing the H value in fuzzy linear regression, *Fuzzy Sets and Systems*, 58 (3): 303-327.
- Pinches, G. E., and J. S. Trieschmann, 1977, Discriminant Analysis, Classification Results, and Financially Distressed P-L Insurers, *Journal of Risk and Insurance*, 44: 289-298.
- Pottier, S. W., and D. W. Sommer, 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., and D. W. Sommer, 2005, On the Use of Group-Level Financial Information in Insurer Solvency Surveillance, Working Paper, University of Georgia.
- Sánchez, J., and A. Gómez, 2003, Applications of Fuzzy Regression in Actuarial Analysis, *Journal of Risk and Insurance*, 70 (4): 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., and A. Stadnik, 2007, Financial Distress in Australian General Insurers, *Journal of Risk and Insurance*, 74 (2): 377-399.
- Sommer, D. W., 1996, The Impact of Firm Risk on Property-Liability Insurance Prices, *Journal of Risk and Insurance*, 63: 501-514.
- Tanaka, H., S. Uejima, and K. Asai, 1982, Linear regression analysis with fuzzy model. *IEEE Transactions on Systems, Man and Cybernetics*, 12 (6): 903-907.
- Tanaka, H., and J. Watada, 1988, Possibilistic linear systems and their application to the linear regression model, *Fuzzy Sets and Systems*, 27 (3): 275-289.
- Zadeh, L. A., 1965, Fuzzy sets, *Information Control*, 8: 338-353.
- Zimmermann, A., 1996, *Fuzzy Set Theory and its Applications*, third ed. (Boston, MA: Kluwer Academic Publishers).

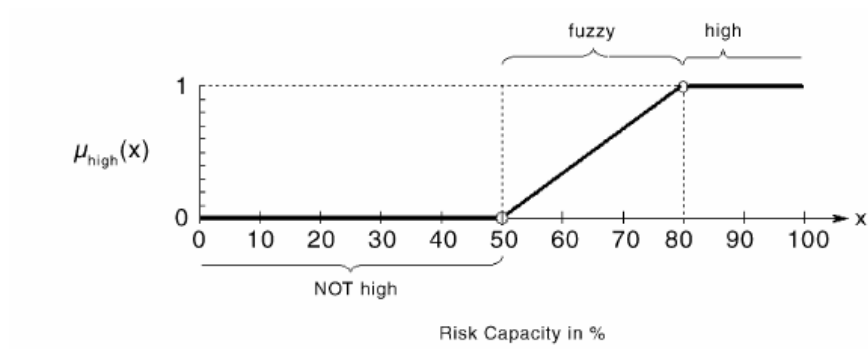


Figure 1: Membership Function for “High Risk Capacity” Companies (Shapiro, 2004: 401)

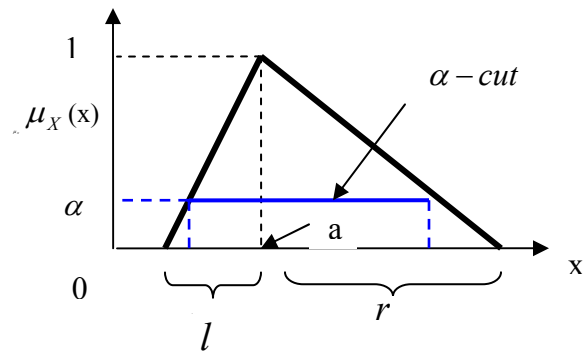


Figure 2: Triangular Fuzzy Number (TFN)

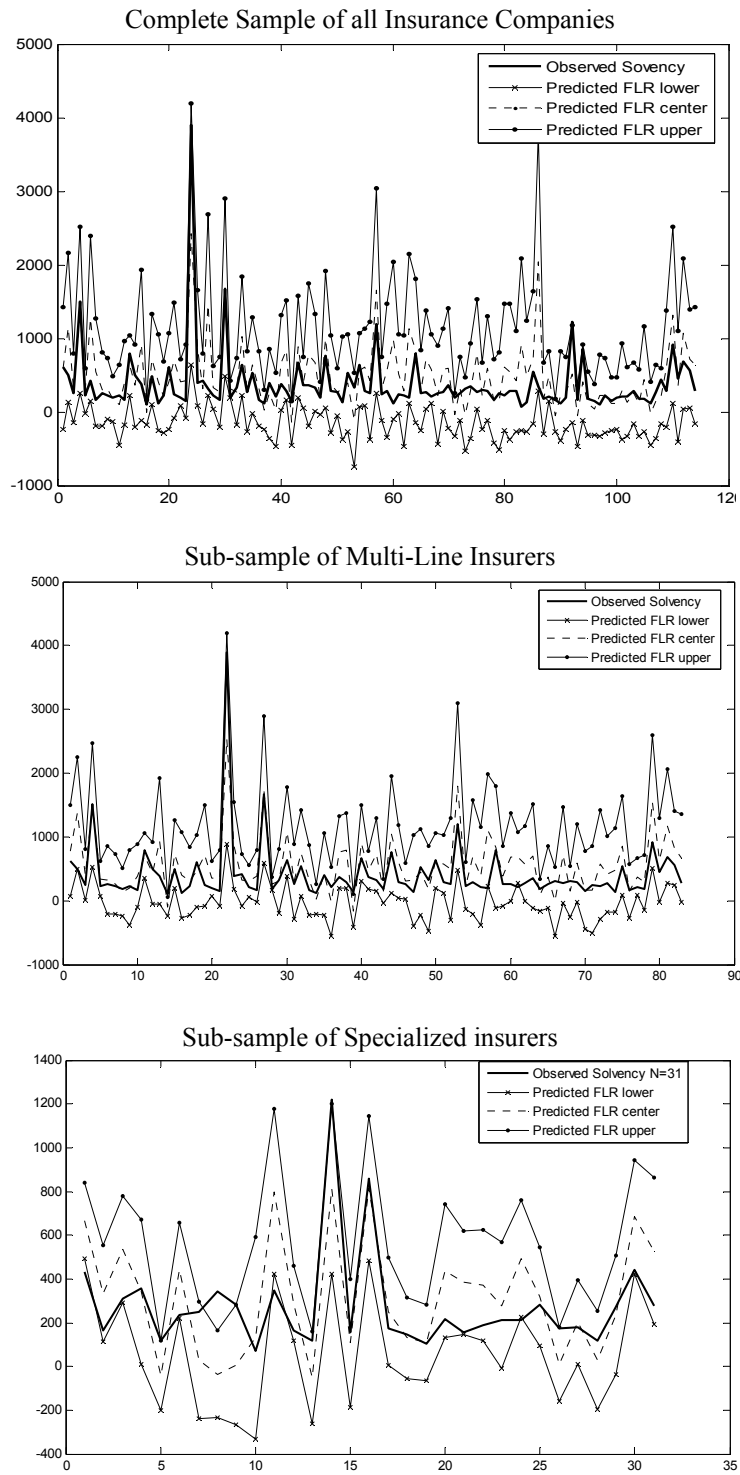


Figure 3: Observed and out-of-sample prediction for Tanaka model of 2005 solvency scores

Table 1. Summary Statistics

| Variable | 2002 | | | | | |
|------------------------------|------|--------|-----------|--------|--------|---------|
| | N | Mean | Std. Dev. | Min | Median | Max |
| Solvency ratio (t + 2) | 114 | 360.04 | 415.23 | 69.93 | 246.65 | 3880.49 |
| 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 | | | | |

| Variable | 2003 | | | | | |
|------------------------------|------|--------|-----------|--------|--------|---------|
| | N | Mean | Std. Dev. | Min | Median | Max |
| Solvency ratio (t + 2) | 114 | 377.16 | 427.39 | 69.49 | 262.88 | 3892.92 |
| Capital/assets | 114 | 20.95 | 9.85 | 4.01 | 19.29 | 60.05 |
| Net premiums/capital | 114 | 273.66 | 169.40 | 33.15 | 253.34 | 1470.85 |
| Return on assets | 114 | 3.22 | 5.15 | -20.62 | 2.87 | 15.88 |
| Line of business Herfindahl | 114 | 41.07 | 29.59 | 12.03 | 31.58 | 100.00 |
| (Stock + real estate)/assets | 114 | 27.24 | 17.23 | 0.00 | 25.19 | 85.33 |
| Natural log of assets | 114 | 19.66 | 1.35 | 17.12 | 19.60 | 23.63 |
| Mutual (%) | 114 | 38.60 | | | | |
| Public (%) | 114 | 14.04 | | | | |

Note: Mutual is a dummy variable equal to 1 if the insurance company is a mutual insurer. Public is a dummy variable 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 percent. The solvency ratios are based on the years 2004 and 2005, respectively. All euro values are inflation adjusted with 2004 as the basis year.

Table 2. Spearman Correlations

| Variable | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|------------------------------|-----|-----------|-----------|--------|-----------|-----------|----------|-----------|
| | | | | | | | | 2003 |
| Capital/assets | (1) | -0.722*** | 0.290*** | 0.123 | 0.088 | -0.225** | 0.075 | 0.302*** |
| Net premiums/capital | (2) | -0.767*** | -0.208** | 0.078 | -0.258*** | -0.196** | 0.107 | -0.352*** |
| Return on assets | (3) | 0.130 | -0.135 | 0.152 | 0.038 | -0.008 | 0.127 | 0.001 |
| Line of business Herfindahl | (4) | 0.141 | 0.014 | 0.167* | -0.153 | -0.512*** | 0.125 | -0.218** |
| (Stock + real estate)/assets | (5) | 0.027 | -0.199** | -0.109 | -0.166* | 0.353*** | -0.133 | 0.256*** |
| Natural log of assets | (6) | -0.170* | -0.187** | 0.030 | -0.508*** | 0.412*** | -0.185** | 0.121 |
| Mutual dummy variable | (7) | 0.117 | 0.008 | 0.103 | 0.128 | -0.202** | -0.179* | -0.320*** |
| Public dummy variable | (8) | 0.304*** | -0.365*** | -0.121 | -0.195** | 0.257*** | 0.123 | -0.320*** |
| | | | | | | | | 2002 |

Note: The table presents Spearman correlations for the variables based on 2002 and 2003 data, separately. ***, **, and * denotes statistical significance at the 1, 5, and 10 percent levels respectively.

Table 3. Univariate Differences

| 2002 | All Insurers | | Multi-line Insurers | | Specialized Insurers | |
|------------------------------|--------------|-----------------------|---------------------|-----------------------|----------------------|-----------------------|
| | Strong | Weak | Strong | Weak | Strong | Weak |
| Capital/assets | | | | | | |
| Mean | 24.67 | 17.64 ^{***} | 24.73 | 17.00 ^{***} | 25.04 | 18.34 |
| Median | 22.85 | 16.9 ^{***} | 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 |
| 2003 | | | | | | |
| Capital/assets | | | | | | |
| Mean | 25.51 | 16.39 ^{***} | 25.29 | 15.86 ^{***} | 25.13 | 18.26 [*] |
| Median | 22.88 | 15.95 ^{***} | 23.39 | 15.95 ^{***} | 19.58 | 15.5 |
| Net premiums/capital | | | | | | |
| Mean | 196.30 | 351.02 ^{***} | 190.01 | 345.72 ^{***} | 225.43 | 362.38 ^{***} |
| Median | 189.30 | 324.32 ^{***} | 176.79 | 308.45 ^{***} | 220.79 | 360.33 ^{**} |
| Return on assets | | | | | | |
| Mean | 4.24 | 2.21 ^{**} | 4.53 | 1.80 ^{***} | 3.57 | 3.07 |
| Median | 3.88 | 1.88 ^{***} | 3.60 | 1.33 ^{***} | 4.28 | 3.73 |
| Line of business Herfindahl | | | | | | |
| Mean | 36.32 | 45.82 [*] | 25.68 | 24.21 | 79.11 | 89.68 |
| Median | 29.36 | 36.05 | 22.69 | 20.80 | 81.50 | 100 ^{**} |
| (Stock + real estate)/assets | | | | | | |
| Mean | 31.84 | 22.64 ^{***} | 32.69 | 21.53 ^{***} | 32.32 | 22.16 |
| Median | 31.87 | 20.93 ^{***} | 31.89 | 20.76 ^{***} | 32.79 | 17.79 ^{**} |
| Natural log of assets | | | | | | |
| Mean | 19.95 | 19.38 ^{**} | 20.11 | 19.74 | 19.32 | 18.56 ^{**} |
| Median | 19.93 | 19.20 [*] | 20.04 | 19.76 | 19.29 | 18.61 |
| Mutual (%) | 35.09 | 42.11 | 35.71 | 46.34 | 25.00 | 40.00 |
| Public (%) | 24.56 | 3.51 ^{***} | 33.33 | 2.44 ^{***} | 0.00 | 6.67 |
| N | 57 | 57 | 42 | 41 | 16 | 15 |

Note: The criteria for splitting the sample insurer in multi-line and specialized insurers is having a Herfindahl index smaller than 50% vs. having a Herfindahl index greater or equal to 50%, respectively. Insurers with a solvency ratio below the median solvency ratio are classified as “weak”, 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. ***, **, and * denotes statistical significance at the 1, 5, and 10 percent levels respectively.

Table 4. OLS Regression of Solvency Ratios on Insurer Characteristics

| Variable | All Insurers | Multi-line Insurers | Specialized Insurers |
|------------------------------|-------------------------|-------------------------|-----------------------|
| Intercept | -2741.890*** (-4.41) | -3016.191*** (-4.41) | -2005.272* (-1.91) |
| Capital/assets | 25.158*** (5.96) | 45.58*** (7.99) | -0.094 (-0.02) |
| Net premiums/capital | 0.075 (0.33) | 0.572** (2.21) | -0.554 (-1.67) |
| Return on assets | -1.286 (-0.26) | -1.627 (-0.22) | -5.901* (-1.73) |
| Line of business Herfindahl | 0.257 (0.22) | -0.959 (-0.23) | -0.768 (-0.44) |
| (Stock + real estate)/assets | -1.154 (-0.54) | -0.604 (-0.21) | 3.804* (1.93) |
| Natural log of assets | 129.713*** (4.57) | 117.791*** (3.84) | 127.759** (2.67) |
| Mutual dummy variable | 66.732 (0.95) | 48.242 (0.58) | 15.987 (0.22) |
| Public dummy variable | -3.408 (-0.03) | -121.870 (-1.04) | -194.295 (-1.17) |
| R squared | 0.414 | 0.572 | 0.667 |
| Number of observations | 114 | 83 | 31 |
| In sample prediction | | | |
| Root mean squared error | 316.39 | 300.64 | 132.95 |
| (1 – Theil’s U) | 0.423 | 0.502 | 0.630 |
| Out of sample prediction | | | |
| Root mean squared error | 310.58 | 279.49 | 166.76 |
| (1 – Theil’s U) | 0.454 | 0.556 | 0.532 |

Note: The dependent variable in the OLS regression is the solvency ratio for each individual insurer based on 2004 data. All independent variables are based on the year 2002. The criteria for splitting the sample insurer in multi-line and specialized insurers is having a Herfindahl index smaller than 50% vs. having a Herfindahl index greater or equal to 50%, respectively. T-statistics appear in parentheses. ***, **, and * denotes statistical significance at the 1, 5, and 10 percent levels respectively. The estimated regression coefficients are then used to predict insurer solvency scores in sample, as well as for our holdout sample consisting of 2003 insurer characteristics and 2005 solvency ratios.

Table 5. Standard Possibilistic Regression and Test for Fuzzyness of Explanatory Variables

| Variable | All Insurers | | Multi-line Insurers | | Specialized Insurers | |
|------------------------------|--------------|----------------------|---------------------|----------------------|----------------------|-------------------|
| | Center | Spread | Center | Spread | Center | Spread |
| Intercept | -3589.998 | 0.000 (0.988) | -3414.955 | 0.000 (1.000) | -4915.272 | 0.000 (0.851) |
| Capital/assets | 49.329 | 29.230*** (0.000) | 51.789 | 25.241*** (0.000) | 9.644 | 0.823 (0.353) |
| Net premiums/capital | 0.484 | 0.000 (1.000) | 0.448 | 0.000 (1.000) | 0.000 | 0.000 (1.000) |
| Return on assets | 0.000 | 0.000 (0.915) | 0.000 | 0.000 (1.000) | -9.166 | 0.000 (0.849) |
| Line of business Herfindahl | 0.000 | 0.000 (1.000) | 0.000 | 0.000 (1.000) | -0.798 | 0.000 (0.852) |
| (Stock + real estate)/assets | -5.412 | 0.000 (0.898) | -0.518 | 0.000 (1.000) | 0.000 | 3.472* (0.051) |
| Natural log of assets | 156.075 | 0.000 (0.966) | 140.987 | 0.000 (1.000) | 269.732 | 7.574 (0.165) |
| Mutual dummy variable | 0.000 | 83.555 (0.227) | 0.000 | 182.682* (0.065) | 4.222 | 0.000 (0.892) |
| Public dummy variable | 15.539 | 0.000 (0.925) | 0.000 | 0.000 (1.000) | 2.882 | 0.000 (0.550) |
| Model fit: \bar{h} | | 0.609* (0.054) | | 0.566* (0.095) | | 0.413 (0.510) |
| Number of observations | | 114 | | 83 | | 31 |

Note: The dependent variable is the solvency ratio for each individual insurer based on 2004 data. All independent variables are based on the year 2002. p-values based on the test for the fuzzyness of regression coefficients as well as for the fuzzyness of the overall sample appear in parentheses. ***, **, and * denotes statistical significance at the 1, 5, and 10 percent levels respectively.

Table 6. Revised Possibilistic Regression and Test for Fuzzyness of Explanatory Variables

| Variable | All Insurers | | Multi-line Insurers | | Specialized Insurers | |
|------------------------------|--------------|-----------|---------------------|-----------|----------------------|----------|
| | Center | Spread | Center | Spread | Center | Spread |
| Intercept | -3589.998 | 0.000 | -3439.767 | 0.000 | -3782.666 | 0.000 |
| | | (0.879) | | (0.835) | | (0.725) |
| Capital/assets | 49.329 | 29.230*** | 51.910 | 28.493*** | 9.456 | 0.000 |
| | | (0.000) | | (0.000) | | (0.776) |
| Net premiums/capital | 0.484 | 0.000 | 0.483 | 0.000 | 0.000 | 0.000 |
| | | (0.767) | | (0.765) | | (0.763) |
| Return on assets | 0.000 | 0.000 | 0.000 | 0.000 | -6.016 | 0.000 |
| | | (0.759) | | (0.763) | | (0.768) |
| Line of business Herfindahl | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.773 |
| | | (0.771) | | (0.769) | | (0.333) |
| (Stock + real estate)/assets | -5.412 | 0.000 | 0.000 | 0.000 | 0.000 | 7.138*** |
| | | (0.773) | | (0.771) | | (0.001) |
| Natural log of assets | 156.075 | 0.000 | 139.279 | 0.000 | 205.587 | 0.031 |
| | | (0.780) | | (0.775) | | (0.362) |
| Mutual dummy variable | 0.000 | 83.555 | 0.000 | 46.665 | 0.453 | 0.000 |
| | | (0.226) | | (0.325) | | (0.883) |
| Public dummy variable | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 |
| | | (0.920) | | (0.910) | | (0.984) |
| Model fit: \bar{h} | | 0.610 | | 0.609* | | 0.511 |
| | | (0.169) | | (0.059) | | (0.250) |
| Number of observations | | 114 | | 83 | | 31 |

Note: The dependent variable is the solvency ratio for each individual insurer based on 2004 data. All independent variables are based on the year 2002. p-values based on the test for the fuzzyness of regression coefficients as well as for the fuzzyness of the overall sample appear in parentheses. ***, **, and * denotes statistical significance at the 1, 5, and 10 percent levels respectively.