



MAX-PLANCK-INSTITUT FÜR SOZIALRECHT UND SOZIALPOLITIK
MAX PLANCK INSTITUTE FOR SOCIAL LAW AND SOCIAL POLICY

mea *Munich Center for the Economics of Aging*

Asymmetric Information in the Market for Automobile Insurance: Evidence from Germany

Martin Spindler, Joachim Winter, Steffen Hagmayer

08-2012

MEA DISCUSSION PAPERS



Alte Nummerierung: 259-12

Asymmetric Information in the Market for Automobile Insurance: Evidence from Germany*

Martin Spindler

Max Planck Institute for Social Law and Social Policy and University of Munich, Germany

Joachim Winter

Department of Economics, University of Munich, Germany

Steffen Hagmayer

HUK-COBURG, Coburg, Germany

June 12, 2012

Abstract: Asymmetric information is an important phenomenon in insurance markets, but the empirical evidence on the extent of adverse selection and moral hazard is mixed. Because of its implications for pricing, contract design, and regulation, it is crucial to test for asymmetric information in specific insurance markets. In this paper, we analyze a recent data set on automobile insurance in Germany, the largest such market in Europe. We present and compare a variety of statistical testing procedures. We find that the extent of asymmetric information depends on coverage levels and on the specific risks covered which enhances the previous literature. Within the framework of Chiappori et al. (2006), we also test whether drivers have realistic expectations concerning their loss distribution, and we analyze the market structure.

Keywords: Asymmetric information; automobile insurance; parametric tests; imperfect competition.

JEL classification codes: D82, C12, G22.

*Address correspondence to: Martin Spindler, Max Planck Institute for Social Law and Social Policy, Amalienstrasse 33, 80799 Munich, Germany; E-mail: Spindler@mea.mpsoc.mpg.de, Phone: +49 89 38602 343. Joachim Winter, University of Munich, Economics Department, 80539 Munich, Germany; E-mail: Joachim.Winter@lrz.uni-muenchen.de. Steffen Hagmayer, HUK Coburg Insurance Group, 96448 Coburg, Germany; Email: Steffen.Hagmayer@huk-coburg.de. Spindler gratefully acknowledges financial support from the Deutsche Forschungsgemeinschaft through GRK 801 and in particular from The Geneva Association through a research grant. We thank Bernard Salanié for valuable comments and Thomas Yee for assistance in applying the R package VGAM. We are grateful to participants of the EEA/ESEM annual meeting 2011 in Oslo, of the annual conference of the German Economic Association (VfS) 2011 in Frankfurt am Main and of the annual seminar of the EGRIE 2011 in Vienna. Parts of this paper were written while Martin Spindler was visiting Columbia University, which he thanks for its hospitality.

1 Introduction

Since Akerlof (1970), the consequences of asymmetric information, in particular adverse selection and moral hazard, have been explored in a vast body of research. The initial gap between the theoretical developments and empirical studies of asymmetric information has recently become narrower. In particular, insurance markets have proved a fruitful and productive field for empirical studies, for two reasons. First, the data are well-structured: insurance contracts are usually highly standardized; they can be described completely by a relatively small set of variables; and data on the insured person's claim history, i.e., the occurrence of claims and the associated costs, is stored in the database of an insurance company. Second, insurance companies have hundreds of thousands or even millions of clients and therefore the samples are sufficiently large to conduct powerful statistical tests. The markets for automobile insurance, annuities and life insurance, crops insurance, as well as long-term care and health insurance provide large samples of standardized contracts for which performances are recorded and are well suited for testing the theoretical predictions of insurance theory. Chiappori and Salanié (1997) provide a detailed justification for using insurance data to test contract theory. Cohen and Siegelman (2010) present a comprehensive overview of approaches for testing for adverse selection in insurance markets, covering a large number of empirical studies in different insurance branches.

In statistical terms, the notion of asymmetric information implies a positive (conditional) correlation between coverage and risk. Several different methods how to test for asymmetric information have been proposed in the literature. In this paper, we apply an array of such tests to detailed contract-level data from the German car insurance market.

Our study contributes to the existing literature in several respects. First, we present the first study analyzing the German car insurance market. The German car insurance market is the largest in Europe and therefore for many insurance companies the most important sales market for their insurance policies. We had unique access to the data set of one of the largest insurance companies in the field of automobile insurance in Germany.

Second, the literature has reached an almost complete consensus that asymmetric information is not prevalent in automobile insurance. Our analysis shows that this finding does not hold in general; in particular, we show that the institutional arrangements of a market and the structure of the contracts have a great influence on whether the insureds have an informational advantage that they can possibly exploit. Because of a special arrangement that holds in the car insurance in Germany, we can show that the extent of asymmetric information depends on the specific kind of risks which is covered.

Third, we compare several tests for asymmetric information that have been proposed in the literature. Chiappori and Salanié (2000) propose tests for the positive correlation property, Dionne et al. (2001) use a two stage approach and Kim et al. (2009) modify a multinomial approach. Most studies only apply a selection of these tests. We apply all tests on the same data set. We find that they give consistent results, not in a statistical but in a qualitative sense: They all give the same answer to the underlying question of whether there is asymmetric information in a particular market. The tests we use are based on different statistical strategies, i.e., they translate the formal definition of asymmetric information differently into a statistical framework. Our finding that these tests deliver robust results suggests that the choice of a particular parametric method is only of secondary importance; the more important aspect is whether the institutional arrangements are taken into account when the tests are interpreted (as argued above).

Finally, by applying the framework of Chiappori et al. (2006) we can also test (i) whether consumers know their loss distribution; (ii) whether in this market the non-increasing profit assumption holds, i.e., whether contracts with higher coverage earn not higher profits; and, most importantly, we can test (iii) whether some form of generalized positive correlation property holds (which is also valid when there are differences in risk preferences). The first two statements are interesting in their own right (although they serve as assumptions for the last one). Specifically, test (i) addresses an important practical question in insurance economics: Can the insureds correctly estimate their loss distribution or do they overestimate

or underestimate their risk? Test (ii) provides empirical evidence on the structure of the German car insurance market: Our results show that the non-increasing profit condition holds, which indicates that there was strong competition among insurance firms in the period for which we observe our data. This finding is in line with informal descriptions of this particular market that we discuss in section 3.

The rest of the paper is structured as follows. Section 2 outlines the theory of asymmetric information and summarizes the empirical literature relevant for our paper. Section 3 describes the institutional arrangements that govern car insurance in Germany and the structure of this specific insurance market. In section 4, we describe the data set. In section 5, we review briefly the parametric tests used in this field and present the results for our data set. In section 6, we introduce the generalized positive correlation property and some related tests and present the results. Final remarks and conclusions are contained in Section 7.

2 Asymmetric Information in Insurance Markets: Theory and Evidence

In their seminal paper Rothschild and Stiglitz (1976) introduce the notion of adverse selection in insurance markets, which has since then been extended in many directions.¹ In the basic model, the insureds have private information about the expected claim, exactly speaking about the probability that a claim with fixed level occurs, while the insurers do not have this information. Thus there are two groups with different claim probabilities, the “bad” and “good” risks. The agents have identical preferences which are moreover perfectly known to the insurer. Additionally, perfect competition and exclusive contracts are assumed. Exclusive contracts mean that an insured can buy coverage only from one insurance company. This allows firms to implement nonlinear (especially convex) pricing schemes which are typical under asymmetric information. In this setting, insurance companies offer a menu of contracts

¹For a detailed survey on adverse selection and the related moral hazard problem, see Dionne, Doherty and Fombaron (2000) and Winter (2000), respectively.

in equilibrium: a full insurance which is chosen by the “bad” risks and a partial coverage which is bought by the “good” risks. In general, contracts with more comprehensive coverage are sold at a higher (unitary) premium.

Clearly, one expects a positive correlation between “risk” and “coverage” (conditional on observables). Since the assumptions in the Rothschild and Stiglitz model are very simplistic and normally not fulfilled in real applications, an important question to address is how robust this coverage-risk correlation is. Chiappori et al. (2006) show that the positive correlation property extends to much more general models under a suitably defined notion. In particular, the notion of positive correlation is generalized in this context. On competitive markets, this property is also valid in a very general framework entailing heterogeneous preferences, multiple level of losses, multidimensional adverse selection plus possible moral hazard and even non-expected utility theory.² In the case of imperfect competition, some form of positive correlation must hold if the agent’s risk aversion becomes public information. In the case of private information the property does not necessarily hold (see Jullien et al. (2007)).

While adverse selection concerns “hidden information”, moral hazard deals with “hidden action”. Moral hazard occurs when the expected loss (accident probability or level of damage) is not exogenous, as assumed in the adverse selection case, but depends on some decision or action made by the insured (e.g., effort spent to prevent a damage) which is neither observable nor contractible. Higher coverage leads to decreased effort and therefore to higher expected loss. Thus, moral hazard also predicts a positive correlation between coverage and risk.

To summarize, the theory of asymmetric information on insurance markets predicts a positive correlation between (appropriately defined) “risk” and “coverage”, a prediction which is quite robust across different theoretical models. Nevertheless, there is one important difference: Under adverse selection, the risk of the potential insured affects the choice of the

²The generalized correlation approach, which we employ in our empirical analysis, is also consistent with recent extensions of the asymmetric information framework for insurance markets, such as differences in risk preferences (see, for example, Finkelstein and McGarry, 2006, or Fang et al. 2008). This approach allows for differences in risk aversion and holds if the market is competitive, which as we argue in the next section, was the case for the German car insurance market in the period we study.

contract, whereas under moral hazard the chosen contract influences behavior and therefore the expected loss. However, there exists reversed causality in both cases.³ It seems that the empirical insurance literature has concentrated on testing for adverse selection whereas the moral hazard aspect has received only minor attention; see, inter alia, Cohen and Siegelman (2010).

In order to test for asymmetric information, the researcher needs to have access to the same information which is also available to the insurer (and used for pricing). The theory of adverse selection predicts that the insurance company offers a menu of contracts to indistinguishable individuals. Individuals are (*ex ante*) indistinguishable for the insurer if they share the same characteristics. Therefore the positive risk-coverage correlation is valid only conditional on the observed characteristics. In general, different classes of observationally equivalent individuals will be offered different menus of contracts with different prices according to their risk exposure.⁴ The mechanisms described above are valid only within each class.

Despite the scarcity of data sets in empirical contract theory, the automobile insurance market has been analyzed extensively. Amongst others, automobile insurance markets in France (Salanié and Chiappori (2000, 2006) and Richaudeau (1999)), Israel (Cohen (2005)), Canada (Dionne et al. (2001)), Korea (Kim et al. (2009)), Japan (Saito (2009)) and the Netherlands (Zavadil (2011)) have been analyzed. In one of the first studies, Puelz and Snow (1994) confirm the existence of asymmetric information, but Dionne et al. (2001) show that this might be due to misspecification of their model. In general, there is a tendency to confirm absence of asymmetric information, e.g., Salanie and Chiappori (2000), Kim et al. (2009), Dionne et al. (2001) and Zavadil (2011). Evidence for asymmetric information has been found only in data from Israel, and only for experienced drivers, but this result can be contributed to a special institutional feature of this market: insurance companies in Israel cannot gather information about the driving history of their new customers. This gives

³To disentangle moral hazard from adverse selection is a difficult but important problem in the empirical literature. An early attempt is Dionne et al. (2004); for an overview of different strategies see Cohen and Siegelman (2010).

⁴For the theory of risk classification under asymmetric information see Crocker and Snow (2000).

consumers who change their insurer some advantage. The empirical analysis of the German car insurance market also suggests that institutional arrangements matter when it comes to asymmetric information on insurance markets.

3 Automobile Insurance in Germany

Like in many other countries, a third-party vehicle insurance is mandatory for all vehicles in Germany. This is the so called *KFZ-Haftpflicht*, a compulsory liability insurance that covers damage inflicted to other drivers and their cars. In addition, insurance companies offer two types of non-compulsory coverage for damages to the own car. The first one is called *Teilkasko* (TK), the second *Vollkasko* (VK).⁵ The first (partial) type of non-compulsory insurance, TK, covers own damages and losses caused by theft, natural disasters (storm, hail, lightning strike, flood), collusion with furred game and so on. The second (comprehensive) type of non-compulsory insurance, VK, covers accidental damage on the own car, even if caused by oneself, and damages caused by vandalism of strangers. For both types, the insured can choose a deductible.

In the German car insurance, there is also a uniform experience rating system (*Schadenfreiheitsrabatt*) which, however, applies only to the compulsory liability insurance and to VK but not to TK. The number of years without accident is counted separately for the two types (compulsory and VK) and according to these numbers every insured is divided into a class (*Schadensfreiheitsklasse*). which is associated with a bonus coefficient, b_t , which serves as a proxy for past experience. For any year t , the premium is defined as the product of a base amount and this coefficient. The base amount can be set freely by the insurance companies according to their risk classification conditional on characteristics of the insured (such as age, sex, profession, location) but it cannot itself be related to past experience. Suppose the bonus

⁵The term *Kasko* is derived from the Spanish word *casco*, which denotes the hull of a ship, *inter alia*. The German compounds nouns terms *Teilkasko* and *Vollkasko* thus refer to partial and comprehensive coverage of damages to the own car, respectively. We prefer not to translate these terms but use the short-hands TK and VK.

Table 1: An overview of the German car insurance market in 2008

	“Haftpflicht”	“Teilkasko”	“Vollkasko”
number of insured cars in million	39.69	12.6	20.76
number of claims in million	2.57	1.30	3.43
claims expenditure in billion Euro	9.22	0.95	5.00
average claim in Euro	3,600	730	1,460

source: GDV (2010).

coefficient is b_t at the beginning of the t th period. Then the occurrence of an accident during the period leads to a categorization into another class and, for example, an increase of 25 percent at the end of the period (i.e., $b_{t+1} = 1.25b_t$), whereas an accident-free year leads to a reduction of the coefficient according to the new class. Additionally, the coefficient is to be restricted to lie between 230 % and 30 % in the compulsory insurance and between 125 % and 30 % in the VK non-compulsory insurance.

In 2009, the size of German motor insurance markets was about 20 billion Euro. 39.7 million cars were covered by the compulsory liability insurance; this is thus the total number of registered cars. A detailed overview of key figures for the year 2008 is given in Table 1.

In 2009 the premium income was 20,057 million Euro and expenditures for claims were 19,420 million Euro. There are 104 companies in the market competing for contracts, thus the market is broader market than in many other countries in western Europe. A very important statistic for insurance companies is the so called “combined ratio”, cost and claims divided by the premium income. This figure was in the last years about 100 %, in 2008 and 2009 the industry average was slightly above 100 %. A combined ratio over 100 indicates that the insurer is making an underwriting loss. The reason is that in the last five years there was a price battle in this market with cutting rates in each year leading to insurance rates similar to the level of those of the early 1980s (see Bloomberg (09/06/2010)). Between 2004 and 2009, the average premium decreased by 15.9 % (see GenRe (2010)). A detailed analysis of the German car insurance and the prevailing price war in the last years is given in GenRe

(2010). While most insurance markets show a tendency towards oligopolies, in the last years the German car insurance market was very competitive, close to perfect competition as the figures above and recent market surveys indicate.⁶ We study this structure of this market in section 6 by testing the non-increasing profit property, described above (this property holds in most settings of competitive markets but is more general than a simple null-profit condition).

4 The Data Set

For our analysis we had access to the database of the insurance contracts of a major company in Germany. We analyzed the data for the TK and VK (partial and comprehensive) non-compulsory insurance separately, due to the different scopes of indemnity and liability rules described above. In fact, these institutional differences are important for the interpretation of the results. We analyzed both the whole portfolio of contracts and a subsample restricted to “young drivers”⁷. The concentration on beginners enables us to rule out learning effects with respect to individual risk, which might arise over time on both sides of the market – the insured and the insurance company.

As the database is quite large, we restrict our analysis in both cases to random samples from the universe of all contracts.⁸ We use data for the year 2009. The data set contains information about each contract for a full contract year. The sample size in the TK is $n = 5,321$ for the whole portfolio and $n = 5,647$ for the beginners, in the VK $n = 7,200$ resp. $n = 6,466$ for the beginners. The level of deductibles in the TK are 0, 150, 300 and 500 Euro and 0, 300, 1,000 and 2,500 Euro in the VK. As 2,500 Euro is very seldom chosen we omit this level of deductible. In the TK case we use the number of claims exceeding 500

⁶GenRe (2010), Bloomberg (2009), Reuters (2009), Reuters (2010), and Handelsblatt (2010) provide information on the German car insurance market, especially on the existence of a price war and on the toughness of competition in this market over the past few years.

⁷In actuarial usage, the expression “young driver” refers not to the age of the driver but to the driving experience.

⁸For the analysis of the whole portfolio, the random sample was drawn from the set of all contracts which were signed after January 1, 2007. As there are changes in the product menu and contract conditions from time to time, this ensures that the contracts are comparable and differ only in the chosen deductible and not with regard to other contract conditions.

Euro and in the VK the number of claims exceeding twice the highest level of deductible, i.e., $2 \cdot 1,000 = 2,000$ Euro as a measure for the ex post risk. In the TK there is no incentive not to claim accidents as there is no bonus / malus coefficient. In the VK case there is an incentive not to claim all accidents as accidents lead to a worsening of the the risk classification and therefore to a higher premium. But we argue that claims higher than 2,000 Euro are filed in any case.⁹ In general, as accidents below the deductible are not reported to the insurance company and cannot be observed we have to restrict to claims being higher than certain thresholds as mentioned above. Otherwise also in the absence of asymmetric information a positive correlation would be possibly detected.

A detailed statistical analysis of the chosen deductibles and the number of accidents according to the chosen deductibles for both the whole portfolio and the novice drivers broken down into the TK and VK types of non-compulsory insurance is provided in Tables 2 to 5. These tables show the typical pattern for the distribution of claims in the car insurance: Most insurees do not file any claim; more claims than two are very seldom.

Our data set stems from the company with one of the highest market shares, which is also diversified across all regions in Germany, across all occupations and ages. The data should therefore be fairly representative of whole market for automobile insurance in Germany.

5 Testing for Asymmetric Information

5.1 Statistical Procedures

In this section, we present several different parametric methods that have been proposed in the literature to test for asymmetric information. In an econometric sense, we want to test whether there is a positive correlation between risk and coverage. We concentrate on parametric tests which are well established in this field and can be implemented by most

⁹At this level all accidents are filed, as the reimbursement from the insurance company for the damage is in any case higher than the discounted increase in premiums according to internal calculations of the insurance company (personal communication with actuaries). Another possibility would be to use the approach of Dionne and Gagné (2001) or a Heckman selection bias model, but both require additional assumptions and / or variables.

Table 2: Number of accidents according to the choice of deductible in the TK for the whole portfolio

level of deductible / number of accidents	0	1	2	sum
0	552 (0.914)	50 (0.083)	4 (0.003)	604
150	4349 (0.953)	208 (0.046)	7 (0.001)	4564
300	107 (0.964)	4 (0.036)	0 (0.000)	111
500	40 (1.00)	0 (0.000)	0 (0.000)	40

Notes: Figures in brackets denote the relative frequency of the number of accidents for the corresponding level of deductible.

Table 3: Number of accidents according to the choice of deductible in the VK for the whole portfolio

level of deductible / number of accidents resp. share	0	1	2	sum
0	54 (0.982)	1 (0.018)	0 (0.000)	55
150	486 (0.972)	14 (0.028)	0 (0.000)	500
300	5182 (0.977)	119 (0.022)	2 (0.001)	5303
500	962 (0.969)	31 (0.031)	0 (0.000)	993
1000	129 (0.985)	2 (0.015)	0 (0.000)	131

Notes: Figures in brackets denote the relative frequency of the number of accidents for the corresponding level of deductible.

Table 4: Number of accidents according to the choice of deductible in the TK for the novice drivers

level of deductible / number of accidents	0	1	2	sum
0	377 (0.887)	44 (0.103)	4 (0.010)	425
150	4723 (0.940)	288 (0.057)	15 (0.003)	5026
300	148 (0.961)	6 (0.039)	0 (0.000)	154
500	40 (1.00)	2 (0.000)	0 (0.000)	40

Notes: Figures in brackets denote the relative frequency of the number of accidents for the corresponding level of deductible.

Table 5: Number of accidents according to the choice of deductible in the VK for the novice drivers

level of deductible / number of accidents resp. share	0	1	2	3	sum
0	3 (1.000)	0 (0.000)	0 (0.000)	0 (0.000)	13
150	205 (0.995)	1 (0.005)	0 (0.000)	0 (0.000)	206
300	4619 (0.964)	167 (0.035)	6 (0.001)	1 (0.000)	4793
500	1194 (0.954)	56 (0.045)	2 (0.001)	0 (0.000)	1252
1000	129 (0.960)	8 (0.040)	0 (0.000)	0 (0.000)	202

Notes: Figures in brackets denote the relative frequency of the number of accidents for the corresponding level of deductible.

statistical software packages.¹⁰ In the following, the (row) vector X denotes the exogenous variables which are used for risk classification by the insurance company, Y denotes the type of the chosen contract, e.g., the chosen deductible, and Z measures the damage risk. Risk is measured as *ex post* risk, e.g., by the number of accidents or the damage payments by caused the insured. All variables are defined at the level of individual contracts which are indexed by i (the index is omitted if there is no confusion).

5.1.1 Unrelated Probit regressions

The first approach is to define two Probit models, one for the choice of the coverage Y_i (either basic coverage or comprehensive coverage – that is, TK or VK) and the other for the occurrence of an accident Z_i (either no accident or at least one accident with fault):

$$\begin{cases} Y_i = \mathbf{1}(X_i\beta + \varepsilon_i > 0) \\ Z_i = \mathbf{1}(X_i\gamma + \eta_i > 0) \end{cases} \quad (5.1)$$

where ε_i and η_i are independent standard normal errors, and β and γ are coefficient vectors (as columns). First, these two Probit models are estimated independently and then the generalized residuals $\hat{\varepsilon}_i$ and $\hat{\eta}_i$ ¹¹ are calculated. These are required for the following test

¹⁰Spindler and Su (2011) present a new nonparametric test for asymmetric information which is, however, not straightforward to implement.

¹¹For example, the generalized residual $\hat{\varepsilon}_i$ estimates $E(\varepsilon_i|Y_i)$. See Gourieroux et al. (1987) for the definition of generalized residuals in limited dependent models and applications.

Table 6: Description of the variables

name	description	number of categories	
		“Teilkasko”	“Vollkasko”
commitment to work-shop	yes/no	2	2
profession	categorical	9	9
region	categorical	6	4
type of vehicle	categorical	6	6
no-claims bonus	bonus / malus coefficient; categorical	–	8
kilometers per year	categorical	9	9
age of car when being bought	categorical	8	7
lodging of the car over night combined with housing	categorical	12	12
driver	age of the driver and potential drivers – categorical	18	25
keeper of the car	categorical	5	5
payment method	categorical	6	6
bonus	yes/no	2	2
protection against upgrading	yes /no, in the case of an accident the no-claims bonus is not raised	–	2
deductible	different possible values	4	5
number of accidents	discrete		
payment for damages	continuous		

Note: For the young drivers the number of categories is slightly different, i.e., smaller, as they have a little bit lower variation in their characteristics.

statistic

$$W_n = \frac{(\sum_{i=1}^n \hat{\varepsilon}_i \hat{\eta}_i)^2}{\sum_{i=1}^n \hat{\varepsilon}_i^2 \hat{\eta}_i^2}. \quad (5.2)$$

Under the null of conditional independence, $\text{cov}(\varepsilon_i, \eta_i) = 0$ and W_n is distributed asymptotically as $\chi^2(1)$, as shown by Gouriéroux et al. (1987).

Chiappori and Salanié (1997, 2000) introduced this approach. One drawback is that information is lost as Y and Z have to be defined as binary variables. This loss of information might disguise asymmetric information as indicated in Spindler and Su (2011).

5.1.2 Bivariate Probit regression

A related approach is to estimate a bivariate Probit model in which ε_i and η_i are distributed as bivariate normal with correlation coefficient ρ which has to be estimated, and then to test whether $\rho = 0$ or not. In order to test this hypothesis the Wald-, Score- or LR-test can be used.

5.1.3 Two-stage regressions

Multinomial approach (Kim et al. (2009))

Depending on the number of categories of the choice variable Y , this procedure varies a little bit. In the case of a dichotomous Y in the first stage, a bivariate Probit regression of the choice of contract on the exogenous variables, i.e., the variables used for risk classification, is run. The Probit equation is of the form

$$Y_i = \mathbf{1}(X_i\beta + \varepsilon_i > 0) \quad (5.3)$$

with ε_i iid and standard normal. Then the generalized residuals $\hat{\varepsilon}_i$ are obtained:

$$\hat{\varepsilon}_i = \frac{\phi(X_i\hat{\beta})}{\Phi(X_i\hat{\beta})(1 - \Phi(X_i\hat{\beta}))} \left[Y_i - \Phi(X_i\hat{\beta}) \right], \quad (5.4)$$

where ϕ and Φ are the density and cumulative distribution functions of the standard normal distribution, respectively, and $\hat{\beta}$ is the estimated coefficient vector. As an unexplained proba-

bility of making a corresponding coverage choice, $\hat{\varepsilon}_i$ captures the extent of private information in the binary choice of Y_i , conditional on the observables.

In the case of more than two categories of Y_i , an ordered multinomial choice model is applied. For example Y_i is equal to 0 if policyholder i chooses no optional coverage (i.e., liability only), 1 for some optional coverage and 2 for all optional coverage / full coverage. For the kind of insurance we will analyze (separately for TK and VK) Y_i is equal to 0 if someone chooses the highest possible deductible, equal to 1 if someone chooses some medium deductible and 2 for no deductible.

A multinomial choice model is given by

$$Y_i = \begin{cases} 0 & \text{if } Y_i^* \leq \mu_1 \\ 1 & \text{if } \mu_1 < Y_i^* \leq \mu_2 \\ 2 & \text{if } Y_i^* > \mu_2 \end{cases} \quad (5.5)$$

where Y_i^* denotes a latent variable representing the policyholder's utility associated with insurance coverage, and μ_1 and μ_2 are unknown thresholds for observed categories. The ordered multinomial choice model can be estimated using an ordered Probit regression.

With an ordered multinomial variable $Y_i \in \{0, 1, 2\}$, the above procedure is not directly applicable. In order to obtain the unexplained probabilities equivalent to the one in the binary choice model, the choice of three contracts is split up in two binary choices.¹² Therefore they define two auxiliary variables Y_i^1 and Y_i^2 . $Y_i^1 = 0$ if $Y_i = 0$ and $Y_i^1 = 1$ if $Y_i \in \{1, 2\}$. Accordingly, $Y_i^2 = 0$ if $Y_i \in \{0, 1\}$ and $Y_i^2 = 1$ if $Y_i = 2$. Then the generalized residuals are calculated by

$$\hat{\varepsilon}_i^1 = \frac{\phi(X_i\hat{\beta} - \hat{\mu}_1)}{\Phi(X_i\hat{\beta} - \hat{\mu}_1)(1 - \Phi(X_i\hat{\beta} - \hat{\mu}_1))} \left[Y_i^1 - \Phi(X_i\hat{\beta} - \hat{\mu}_1) \right] \quad (5.6)$$

$$\hat{\varepsilon}_i^2 = \frac{\phi(X_i\hat{\beta} - \hat{\mu}_2)}{\Phi(X_i\hat{\beta} - \hat{\mu}_2)(1 - \Phi(X_i\hat{\beta} - \hat{\mu}_2))} \left[Y_i^2 - \Phi(X_i\hat{\beta} - \hat{\mu}_2) \right] \quad (5.7)$$

¹²The exposition follows Kim et al. (2009), also the interpretation of the residuals below follows their presentation.

with $\hat{\beta}$ the estimated coefficient vector and $\hat{\mu}_1, \hat{\mu}_2$ the estimated thresholds for observed categories.

$\hat{\varepsilon}_i^1$ and $\hat{\varepsilon}_i^2$ can be interpreted as the unexplained probabilities associated with Y_i^1 and Y_i^2 . For example, $\hat{\varepsilon}_i^1$ estimates the unexplained probability of choosing any optional coverage ($Y_i \in 1, 2$) over no optional coverage ($Y_i = 0$). $\hat{\varepsilon}_i^2$ can be interpreted in an analogous way. If we include these two residuals in the "accident equation" in the 2nd step as regressors, the regression coefficient of $\hat{\varepsilon}_i^1$ captures the effect of information asymmetry in the choice between no optional coverage ($Y_i = 0$) and some optional coverage ($Y_i = 1$) and the coefficient of $\hat{\varepsilon}_i^2$ captures the effect of information asymmetry in the choice between some optional coverage ($Y_i = 1$) and all optional coverage ($Y_i = 2$).

Therefore in a second step we run a negative binomial regression or alternatively Poisson regression of the number of accidents Z_i in the contract year on the exogenous regressors including the generalized regressors calculated according to 5.6 and 5.7. The distribution of the number of accidents Z_i in the case of the negative binomial regression is given by

$$\mathbf{P}(Z_i) = \frac{\Gamma(Z_i + \frac{1}{\sigma^2}) \left[\sigma^2 \exp(X_i \hat{\beta}_0 + \hat{\varepsilon}_i \hat{\beta}_\varepsilon) \right]^{Z_i}}{\Gamma(\frac{1}{\sigma^2}) \Gamma(Z_i + 1) \left[1 + \sigma^2 \exp(X_i \hat{\beta}_0 + \hat{\varepsilon}_i \hat{\beta}_\varepsilon) \right]^{Z_i + \frac{1}{\sigma^2}}} \quad (5.8)$$

with Γ the Gamma function and $\hat{\beta}_0, \hat{\beta}_\varepsilon$ the estimated coefficient vectors. $\hat{\varepsilon}_i$ is defined as $(\hat{\varepsilon}_i^1, \hat{\varepsilon}_i^2)$.

Statistically significant and positive components of $\hat{\beta}_\varepsilon$ indicate the existence of asymmetric information between the parties.

As a two stage nonlinear estimation procedure is used, one has to apply the Murphy-Topel standard error estimates in the second stage (Murphy and Topel (1985)). As in the second regression regressors are included which are estimated themselves in the first step, one has to account for this additional source of uncertainty and correct the induced bias in the variance. An adapted version which is tailored to the situation above can be found in the appendix of Kim et al. (2009).

Allowing for nonlinearities (Dionne et al. (2001))

Dionne et al. (2001) choose the following procedure. In the first step, $\hat{E}(Z|X)$ is computed by the estimation of a negative binomial regression of the distribution of accidents by using basic rating variables of the insurer as regressors. In the second step, a Probit model with the chosen deductible as independent variable is run. The exogenous variables are the same as in the first step, plus the expected number of claims estimated from the first step and the actual number of accidents. Thus, the regressor $\hat{E}(Z|X)$ in the second equation controls for possible nonlinearities.

In one of the first empirical studies, Puelz and Snow (1994) consider an ordered logit formulation for the deductible choice variable and find strong evidence for the presence of asymmetric information in the market for automobile collision insurance in Georgia. But Dionne et al. (2001) show that this correlation might be spurious because of the highly constrained form of the exogenous effects or the misspecification of the functional form used in the regression. They propose to add the estimate $\hat{E}(Z_i|X_i)$ of the conditional expected value of Z_i given X_i as a regressor into the ordered logit model to take into account the nonlinear effect of the risk classification variables, and by accounting for this, they find no residual asymmetric information in the market for Canadian automobile insurance.¹³

5.2 Results

In this section, we present our results. In order to characterize the robustness of the inferences on asymmetric information that a researcher might draw for a given market, we apply all the parametric tests presented in the previous section. We apply them separately for the TK and VK types of insurance, and for both samples (whole portfolio and young drivers). The results for the whole portfolio are summarized in Table 7, the results for the novice drivers in Table 8. As the results for both groups are, perhaps surprisingly, relatively similar, we first discuss the results for the whole portfolio in some detail before comparing the results for both samples.

¹³Chiappori and Salanié (2000) also estimated a nonparametric model which confirmed the results from their parametric models.

In the TK, the picture seems to be clear-cut. Both tests building up on the two Probits and the bivariate Probit reject the null hypothesis of conditional independence at a significance level of $\alpha = 0.01$. When using a two-step estimation procedure, the number of accidents has a significant influence in predicting the choice of deductible ($\hat{\beta}_{accidents} = 0.517$). In order to take into account nonlinearities, we also include the expected number of accidents, additional to the number of accidents. The expected number of accidents is estimated according to a Poisson regression in which all variables used for risk classification by the insurance company are included. We also applied a Negative Binomial regression, but as the results are similar to the Poisson regression and as we find no indication of overdispersion, we omit them here. While in Dionne et al. (2001) the addition of the expected number of accidents made the coefficient of the number of accidents insignificant, our results are not changed so that the actual number of accidents remains a significant predictor of the choice of the deductible. Therefore, the two step estimation confirms the existence of asymmetric information in the TK as well. The multinomial approach shows that there is no asymmetric information in the choice between the highest possible deductible and some lower deductible. This information is contained in the coefficient of ε_1 . In the choice between some deductible and full insurance, presumably the most important decision, the generalized residual ε_2 has a significant positive influence which again indicates asymmetric information.

The interpretation of the results for comprehensive coverage (VK) is similar but the estimates are quite different, as can be seen in Table 7. The estimated coefficient of ρ is clearly not significantly different from zero and the test statistic W using the generalized residuals does not reject conditional independence on a significance level $\alpha = 0.01$. In the two step estimation, the number of accidents is not a significant variable, regardless of taking into account nonlinearities. Also the multinomial approach shows no indication of asymmetric information if one examines the detailed choice between certain levels of deductible. The results for the young drivers are interpreted in an analogous way. A comparison of Table 7 and 8 shows that the pattern for the whole portfolio and the novice drivers is surprisingly similar and robust.

To sum up, we detect the existence of asymmetric information in the TK, but not in the VK, for both experienced and young drivers.

Table 7: Results for the whole portfolio

Test Procedure	TK	VK
Two Probits	$W = 21.9$	$W = 0.01$
	reject conditional independence ($\alpha = 0.01, 0.05$)	do not reject conditional independence ($\alpha = 0.01, 0.05$)
Bivariate Probit	$\rho = 0.286$ (0.049)	$\rho = -0.1403$ (0.126)
	reject conditional independence ($\alpha = 0.01, 0.05$)	do not reject conditional independence ($\alpha = 0.01, 0.05$)
2 step estimation (Dionne et al. (2009))	$\hat{\beta}_{accidents} = 0.517^{*,**}$ (0.083)	$\hat{\beta}_{accidents} = 0.036$ (0.094)
	$\begin{cases} \hat{\beta}_{accidents} = 0.509^{*,**}(0.083) \\ \hat{\beta}_{exp_acc} = 1.4(1.098) \end{cases}$	$\begin{cases} \hat{\beta}_{accidents} = 0.019(0.095) \\ \hat{\beta}_{exp_acc} = 1.725(1.090) \end{cases}$
Multinomial approach	$\beta_{\hat{\varepsilon}_1} = 1.721$ (2.038/2.056)	$\beta_{\hat{\varepsilon}_1} = 0.338$ (0.332/0.332)
	$\beta_{\hat{\varepsilon}_2} = 0.352^{*,**}$ (0.079/0.149)	$\beta_{\hat{\varepsilon}_2} = 0.053$ (0.386/0.386)

Notes: *, ** indicate significance at the 1% and 5%, respectively.

Figures in brackets indicate the standard errors, and for the multinomial approach additionally the Murphy-Topel standard errors.

Table 8: Results for novice drivers

Test Procedure	TK	VK
Two Probits	$W = 15.4$ reject conditional independence ($\alpha = 0.01, 0.05$)	$W = 3.86$ do not reject conditional independence for $\alpha = 0.01$ and reject for $\alpha = 0.05$
Bivariate Probit	$\rho = 0.227$ (0.066) reject conditional independence ($\alpha = 0.01, 0.05$)	$\rho = -0.0101$ (0.108) do not reject conditional independence ($\alpha = 0.01, 0.05$)
2 step estimation (Dionne et al. (2009))	$\hat{\beta}_{accidents} = 0.388^{*,**}$ (0.075) $\begin{cases} \hat{\beta}_{accidents} = 0.386^{*,**}(0.076) \\ \hat{\beta}_{exp_acc} = 0.640(1.375) \end{cases}$	$\hat{\beta}_{accidents} = -0.123$ (0.076) $\begin{cases} \hat{\beta}_{accidents} = -0.125(0.076) \\ \hat{\beta}_{exp_acc} = 1.306(1.871) \end{cases}$
Multinomial approach	$\beta_{\hat{\varepsilon}_1} = 0.1241$ (0.2676) $\beta_{\hat{\varepsilon}_2} = 0.4136^{*,**}$ (0.074)	$\beta_{\hat{\varepsilon}_1} = 0.038$ (0.167) $\beta_{\hat{\varepsilon}_2} = -1.785$ (4.086)

Notes: *, ** indicate significance at the 1 % and 5 %, respectively.

Figures in brackets indicate the standard errors, and for the multinomial approach additionally the Murphy-Topel standard errors.

6 The Generalized Positive Correlation Property

6.1 Statistical Procedures

As mentioned earlier, Chiappori et al. (2006) extend the positive correlation property (for suitably defined notions) to much more general models, entailing heterogeneous preferences, multiple level of losses, multidimensional adverse selection, plus possibly moral hazard and even nonexpected utility. They generalize the notion of positive correlation between risk and coverage. The most important assumptions they make are the “realistic expectations” assumption and the “non-increasing profit”(NIP) assumption. Realistic expectations means that when agents choose a contract, they correctly assess their accident probability and loss distribution. With other words, they know their loss distribution. The NIP assumption states that if a

contract C_2 covers more than a contract C_1 , then the expected profits generated by C_2 are not higher than profits under C_1 : $\pi(C_2) \leq \pi(C_1)$.

These assumptions (or rather specific predictions that follow from them, as explained below) can be tested with our contract data. The results are of interest in their own right, as they answer questions related to the market structure and to the self-assessment of drivers. We present only the special case for contracts with deductibles. In the following, C_1 and C_2 are contracts with straight deductibles, where C_2 covers more than C_1 , i.e., the deductible for contract 1 is higher, $d_1 \geq d_2$. P_1 , P_2 denote the corresponding premia, R_1 , R_2 denote the indemnities for every possible claim under each contract and can be approximated by loss minus the deductible in our case.

Under the null that the insureds know their loss distribution, i.e., have realistic expectations, the following property should hold in the case of straight deductibles

$$P_2 - P_1 \geq q_1(d_1 - d_2)$$

q_1 denotes the probability that the loss L is above the deductible d_1 under C_1 : $q_1 = \text{Prob}(L > d_1)$.

The condition of NIP can be formulated in this special case as

$$P_1/(1+t) - P_2/(1+t) \geq (1+\lambda)(R_1 - R_2)$$

with t the tax rate and λ the load factor.

Finally, under the two assumptions mentioned above (and some other weak conditions) a generalized positive correlation property holds which in our case can be written as¹⁴

$$(1+K)(E_2 - E_1) \geq d_2(q_2 - q_1).$$

K is defined as $(1+t)(1+\lambda) - 1$, E_i denotes the expected loss under contract i . This can be approximated by the payment of compensation in the case of an accident plus the

¹⁴Chiappori et al. (2006) give further explanations and an intuition for the result.

corresponding deductible. q_i is the probability of a claim under contract C_i .

It is important to mention that the inequalities above are only valid for individuals with the same observables X and therefore have to be checked for all cells separately.

In order to test if these inequalities hold, one forms a test statistic T for each cell. In the case of realistic expectations T is given by

$$T = P_2 - P_1 - q_1(d_1 - d_2)$$

for the NIP

$$T = P_1/(1+t) - P_2/(1+t) - (1+\lambda)(R_1 - R_2)$$

and finally for the positive correlation property

$$T = (1+K)(E_2 - E_1) - d_2(q_2 - q_1).$$

As these test statistics are calculated for every cell, we get an empirical distribution for each test statistic. Under the null hypothesis that $T(X) = 0$ for all X , these numbers should be distributed as a standard normal distribution, $N(0, 1)$. There are several ways to come to a test decision. One way is to calculate the standardized mean and to conduct a t -test of the hypothesis that the mean is different from zero. Additionally, we report the share of the number of positive signs of our test statistic as this provides additional information on the dispersion of the distribution. As our data cover several levels of deductibles in the TK and VK types of insurance, we compare them pairwise.

6.2 Results

In this section, we present the results of the above tests for TK and VK. We applied the tests to the whole portfolio of contracts in the year 2009. In order to save space, we only present the results if one of the deductibles involved in the comparisons contains at least more than 5% of the contracts of the total portfolio.

The first question to address is whether the customers have realistic expectations. Our

results show that for both the TK and VK and for the different choices of deductibles, the null can be rejected at any reasonable significance level and therefore the customers have realistic expectations about their true loss distribution.

Table 9: Realistic expectations - VK 2009

coverage 1	coverage 2	number of cells	std. mean	t-statistic	share of pos. signs
300	150	6,843	1.90	157.55	0.998
500	150	6,595	2.19	178.23	1.000
500	300	8,730	1.89	176.43	0.999

coverage in Euro. sample size appr. 2.3 mio.

Table 10: Realistic expectations - TK 2009

coverage 1	coverage 2	number of cells	std. mean	t-statistic	share of pos. signs
150	0	1,880	1.72	74.68	0.999

coverage in Euro. sample size appr. 1.2 mio.

The second question to be addressed is whether the non-increasing profit assumption holds. This is also interesting with regard to market structure and strength competition on this market. Tables 11 and 12 show that the mean is significantly different from zero and therefore the null can be rejected, so the non-increasing profit property holds. However, the share of observations with positive sign falls below 50 %. Therefore, the situation is not completely clear as the value of the test statistic alone suggests, but nevertheless we see a tendency that the NIP assumption is fulfilled.

Finally, we present the results of the general positive correlation tests, summarized in Tables 13 and 14. For the TK, we can reject the null hypothesis (at a significance level of 1 %) and therefore confirm the result of asymmetric information in this part of the automobile insurance. The share of cells with positive sign is approximately 75.6 %. This result is in line with the results we obtained by applying the parametric procedures in the previous section.

Table 11: Non increasing profit assumption (NIP) - VK 2009

coverage 1	coverage 2	number of cells	std. mean	t-statistic	share of pos. signs
300	150	6,843	0.04	3.70	0.38
500	150	6,595	0.07	5.56	0.44
500	300	8,730	0.02	2.09	0.62

coverage in Euro. sample size appr. 2.3 mio.

Table 12: Non increasing profit assumption (NIP)- TK 2009

coverage 1	coverage 2	number of cells	std. mean	t-statistic	share of pos. signs
150	0	1,880	0.06	2.57	0.450

coverage in Euro. sample size appr. 1.2 mio.

For the VK, the picture is not so clear-cut. The tests confirm the existence of a positive correlation property. This contradicts the previous section at a first glance. But the number of cells with positive sign is quite low. For the comparisons of deductibles, the share goes down to about one fourth – the overwhelming number of cells has a negative sign. This makes it plausible that despite the value of the test statistic, the positive correlation property does not hold for most cells. Therefore, the existence of asymmetric information seems to be doubtful or at least the effect is not strong. Moreover, a detailed look at the cells reveals that the high value of the test statistics is driven by some cells with a sparse number of observations which could be classified as outliers.

Summing up, we can confirm that drivers have realistic expectations with regard to their loss distribution and that in this market (and this period) the NIP condition holds. Moreover, we show that the generalized positive correlation property holds for TK. For VK the picture is not so clear-cut. Depending on the chosen test, evidence for VK that this property also holds is quite weak.

Table 13: Generalized positive correlation - VK 2009

coverage 1	coverage 2	number of cells	std. mean	t-statistic	share of pos. signs
300	150	6,843	0.09	7.56	0.410
500	150	6,595	0.15	12.26	0.495
500	300	8,730	0.08	7.45	0.659

coverage in Euro. sample size appr. 2.3 mio.

Table 14: Generalized positive correlation - TK 2009

coverage 1	coverage 2	number of cells	std. mean	t-statistic	share of pos. signs
150	0	1,880	0.40	17.42	0.756

coverage in Euro. sample size appr. 1.2 mio.

7 Discussion and Concluding Remarks

In this paper, we analyzed a contract-level data set on automobile insurance in Germany. This market is of interest not only because it is the largest such market in Europe, but also because particular contractual arrangements allow us to analyze asymmetric information with respect to the different kinds of risk which are usually covered by car insurance.

Our first conclusion is that in TK (partial insurance), asymmetric information exists. This is confirmed by the tests in section 5 and 6. By weighing all test results, we would however deny the existence of asymmetric information in the VK (comprehensive insurance), or perhaps the effect is very weak compared to the TK. These findings hold for both young drivers and for the whole portfolio of contracts.

To understand the differences in our findings for these two types of coverage, we must consider the types of risk that are covered and how filed claims affect the future premium. The VK covers damages to the own car or own body which are not covered in the TK, in particular damages in the case of an accident for which the driver is at his own fault. (Damages to the other party are covered by the *Haftpflichtversicherung* which we did not analyze in this

paper.) In contrast to the TK, there is a bonus/malus coefficient in the VK, i.e., an accident caused by the insured leads to a worsening of this coefficient and thus to a higher premium in the following years. Thus, there is an incentive in the VK not to file all accidents. We therefore limited our analysis to claims which are twice as high as the highest deductible, i.e., we considered only damages above 2,000 Euro. Damages of this magnitude are in nearly all cases filed as a claim, since the reimbursement will be higher than the discounted sum of the increased future premiums. In the VK, we find no convincing evidence for the existence of asymmetric information.

When we restrict our analysis to young drivers, i.e., drivers with no driving experience, it is reasonable to assume that they do not have an informational advantage concerning their accident probability. This means that in the VK, it is reasonable to rule out adverse selection and therefore the analysis can be interpreted as a test of moral hazard. But our results show that there is no positive correlation as predicted by the theory of moral hazard. This finding might be due to the existence of an experience rating, a deterrent that does not exist in the TK. If one is not willing to maintain the assumption that young drivers do not have an informational advantage, then different effects must cancel to produce an insignificant effect, which is possible but not very likely in our view.

When we consider the whole portfolio of insurance contracts, the pattern is the same. Drivers have an informational advantage in the TK. The absence of evidence for asymmetric information in the VK might be explained by the experience rating system: The bonus/malus coefficient is both a good proxy for the ability and accident history of a driver, and an effective scheme to induce careful driving.

Going one step further, one could presume that the risks covered by the TK cannot be influenced by individual behavior to the same extent as the risks covered by the VK. Therefore one might argue that moral hazard is of minor importance in the TK, in which case the positive correlation would be driven by adverse selection. This would imply that the insureds are aware of specific individual risk factors, such as local variation in weather patterns, and choose the

level of deductible (of full insurance) according to their risk exposure. More research along these lines would require data on, say, regional variation in risk exposure that is not reflected in regional premium variation. We leave this issue to future work.

Finally, an important finding we presented in this paper is that the insureds know their loss distribution, i.e., they have realistic expectations. This rules out the possible explanation that insureds have different abilities in assessing the risks related to the TK or VK. We can also conclude that drivers do not over- or underestimate their risk as it is often conjectured. We also checked the non-increasing profit assumption and found that it is confirmed. Usually insurance markets are characterized by an oligopolistic market structure, but this result confirms the informal descriptions of the German car insurance market in the 1990s: In this specific period, this market seems to have been characterized by almost perfect competition.

References

- Akerlof, G. 1970. The market for lemons: quality uncertainty and the market mechanism. *Quarterly Journal of Economics* 84, 488–500.
- Bloomberg, 2009. Allianz wont cut auto premiums, braced to lose market share. Published on www.bloomberg.com, 11/13/2009.
- Chiappori, P.-A., Heckman, J. J., 1999. Testing for moral hazard using dynamic data. Working paper, University of Chicago.
- Chiappori, P.-A., Salanié, B. 1997. Empirical contract theory: the case of insurance data. *European Economic Review* 41, 943–950.
- Chiappori, P.-A., Salanié, B., 2000. Testing for asymmetric information in insurance markets. *Journal of Political Economy* 108, 56–78.
- Chiappori, P.-A., Jullien, B., Salanié, B., Salanié, F., 2006. Asymmetric information in insurance: general testable implications. *Rand Journal of Economics* 37, 783–798.
- Cohen, A., 2005. Asymmetric information and learning: evidence from the automobile insurance market. *Review of Economics and Statistics* 87, 197–207.
- Cohen, A., Siegelman, P., 2010. Testing for adverse selection in insurance markets. *Journal of Risk and Insurance* 77, 39–84.
- Crocker, K.-J., Snow, A., 2000. The theory of risk classification. In: G. Dionne, editor, *Handbook of Insurance*, Chapter 8, 245–276. Kluwer Academic Publishers.
- Dionne, G., Doherty, N., Fombaron, N., 2000. Adverse selection in insurance markets. In: G. Dionne, editor, *Handbook of Insurance*, Chapter 7, 185–244. Kluwer Academic Publishers.

- Dionne, G., Gagné, R., 2001. Deductible contracts against fraudulent claims: evidence from automobile insurance. *Review of Economics and Statistics* 83, 290–301.
- Dionne, G., Gauthier C., Vanasse, C. 2001. Testing for evidence of adverse selection in the automobile insurance market: A comment. *Journal of Political Economy* 109, 444–453.
- Dionne, G., Michaud, P.-C., Dahchour, M., 2004. Separating moral hazard from adverse selection in automobile insurance: Longitudinal evidence from France. *Cahiers de recherche* 0420, CIRPEE.
- Fang, H., Keane, M. P., Silverman, D., 2008. Sources of advantageous selection: Evidence from the Medigap insurance market. *Journal of Political Economy* 116, 303–350.
- Finkelstein, A., McGarry, K., 2006. Multiple dimensions of private information: Evidence from the long-term care insurance market. *American Economic Review* 96, 938–958.
- GenRe 2010. Die Kfz-Versicherung im Jahr 2010: Ende des Preiskampfs oder Preiskampf ohne Ende?
- Gesamtverband der Deutschen Versicherungswirtschaft (GDV), 2010. Statistisches Taschenbuch der Versicherungswirtschaft 2010.
- Gourieroux, C., Monfort, A., Renault, E., Trognon, A., 1987. Generalised residuals. *Journal of Econometrics* 34, 5–32.
- Handelsblatt, 2010. Allianz opfert die Marktführerschaft in der Auto-Versicherung. Published on www.handelsblatt.com, 09/01/2010.
- Jullien, B., Salanié, B., Salanié, F., 2007. Screening risk averse agents under moral hazard. *Economic Theory* 30, 151–169.
- Kim, H., Kim, D., Im, S., Hardin, J. W., 2009. Evidence of asymmetric information in the automobile insurance market: dichotomous versus multinomial measurement of insurance coverage. *Journal of Risk and Insurance* 76, 343–366.

- Muermann, A., Straka, D., 2011. Asymmetric information in automobile insurance: new evidence from telematics data. Working paper, Vienna University of Economics and Business.
- Murphy, K. M., Topel, R. H., 1985. Estimation and Inference in Two-Step Econometric Models. *Journal of Business and Economic Statistics* 3, 370–379.
- Puelz, R., Snow, A., 1994. Evidence on adverse selection: Equilibrium signaling and cross-subsidization in the insurance market. *Journal of Political Economy* 102, 236–257.
- Reuters, 2009. German motor insurance competition still hard – HUK. Published on www.reuters.com, 01/28/2009.
- Reuters, 2010. German car insurance prices up, market murky – HUK. Published on www.reuters.com, 05/17/2010.
- Richaudeau, D., 1999. Automobile insurance contracts and risk of accident: an empirical test using French individual data. *Geneva Papers on Risk and Insurance Theory* 24, 97–114.
- Rothschild, M., Stiglitz, J. E., 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. *Quarterly Journal of Economics* 90, 630–649.
- Saito, K., 2009. Testing for asymmetric information in the automobile insurance market under rate regulation. *Journal of Risk and Insurance* 73, 335–356.
- Spindler, M., Su, L., 2011. Nonparametric testing for asymmetric information. Working Paper, University of Munich and Singapore Management University.
- Winter, R., 2000. Optimal insurance under moral hazard. In: G. Dionne, editor, *Handbook of Insurance*, Chapter 6, 155–184. Kluwer Academic Publishers.

Yee, T. W., 2009. VGAM: Vector Generalized Linear and Additive Models. R package version 0.7-9, <http://www.stat.auckland.ac.nz/~yee/VGAM>.

Zavadil, T., 2011. Do the better insured cause more damage? Testing for asymmetric information in car insurance. Working Paper, National Bank of Slovakia, Bratislava.