HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY of York

WP 15/15

Risk selection under public health insurance with opt-out

Sebastian Panthöfer

September 2015

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Risk Selection under Public Health Insurance with Opt-out

Sebastian Panthöfer*
Universidad Carlos III de Madrid
spanthoefer@gmail.com

June 21, 2015

Abstract

This paper examines risk selection in a parallel public and private health insurance system in which some, but not all, individuals can purchase substitutive private insurance by opting out of otherwise mandatory public insurance. Using a theoretical model, I show that public insurance is adversely selected when insurers and insureds are symmetrically informed about health-related risks, and that there can be any type of selection (advantageous or adverse) when insureds are privately informed. Using the German Socio-Economic Panel, I present evidence on the selection between public and private health insurance in Germany, which is one of the countries with such a health insurance system. I find that: (1) public insurance is adversely selected, (2) individuals adversely select public insurance based on self-assessed health and advantageously select public insurance based on risk aversion, and (3) there is evidence suggesting the presence of asymmetric information between private insurers and their clients.

Keywords: Public and private health insurance, risk selection, asymmetric information
JEL classification: D82, H51, I13

*I am grateful to Matilde Machado for continued advice and support, and to Sebastian Bauhoff, María Inés Berniell, Jesús Carro, Luis Corchón, Anett Erdmann, Natalia Fabra, Andrés García-Suaza, Ángel Hernando-Veciana, Federico Masera, Jan Nimczik, Kathleen Nosal, Achim Wambach, three anonymous referees, and participants of the 5th ASHEcon Conference, CDSE Seminar Mannheim, Madrid Microeconomics Graduate Workshop, and the UC3M 3rd-year Workshop for helpful comments. Financial support from the Spanish Plan Nacional de I+D+I (ECO2012-31748) is gratefully acknowledged.

1 Introduction

Both mandatory public health insurance and freely competitive markets for private health insurance have their disadvantages. Public health insurance often does not leave much room for consumer choice. It frequently involves a one-size-fits-all policy, such as in the case of the UK National Health Service. In response to inadequate public health insurance coverage, individuals turn to private insurers and end up holding two health insurance policies.[1] On the other hand, in a freely competitive market for private health insurance there may be too little pooling of risks and insufficient access to health insurance from a social point of view. Access to health insurance can become severely limited in the presence of asymmetric information between insurers and insureds, which may manifest itself in contracts with too little coverage (Rothschild and Stiglitz 1976) or in a complete break down of some segments of the insurance market (Akerlof 1970).

Parallel public and private health insurance systems can overcome these drawbacks to a certain extent. However, whenever private insurers compete with a public option, there is the concern that the former are able to cream skim the best risks, leaving the public option adversely selected; a prominently discussed example in this context is the competition between traditional Medicare and private Medicare Advantage plans.[2] In this paper, I contribute towards the understanding of this issue by examining the risk selection between public and private health insurance under the German health insurance system, which is peculiar in that it allows only a subset of the population to opt out of public insurance, where those who opt out do not contribute towards the financing of the public plan.[3] I will refer to this system as public health insurance with opt-out.

I begin by constructing a theoretical model of public health insurance with opt-out in which a budget-balanced public plan that is financed by risk-unrelated contributions coexists along with a market for private health insurance, in which private insurers compete à la Rothschild and Stiglitz (1976). The purpose of the model is twofold. First, it shows that, contrary to the conventional wisdom, private insurers are not always able to cream skim the best risks, even though they can adjust coverage levels whereas the public plan is a single policy with a fixed amount of coverage. Instead, the model predicts that the type of selection depends on the informational scenario. Under symmetric information, selection against the public plan is adverse: those who are eligible to opt out and high risk stay in public insurance, because they profit from the implicit subsidies they receive from pooling with those who cannot opt out, while those who are eligible to opt out and low risk buy private insurance.

[1] For example, 44.9% of Australians, 15.6% of Italians, and 10% of the British hold private health insurance, although all three countries provide universal health care (Colombo and Tapay 2004). Many of these private policies are duplicative, i.e., they cover services which are already covered by universal insurance.

[2] See, for example, Cao and McGuire (2003), Brown *et al.* (2014), and Newhouse *et al.* (2012).

[3] While peculiar, Germany is not the only country with such a health insurance system: Argentina, Chile, Colombia, and Peru all have similar arrangements.

Under asymmetric information, public insurance can be adversely or advantageously selected. There can be advantageous selection into public insurance because private insurers offer less insurance coverage to healthier individuals in order to screen different risks. This can lead to a situation in which some good risks, who are offered relatively little coverage in private insurance, stay in public insurance, whereas some bad risks, who are offered relatively generous coverage in private insurance, opt out. The second purpose of the model is that it suggests a test of asymmetric information, which I will conduct in the empirical part.

Drawing on data from the German Socio-Economic Panel (SOEP), I provide empirical evidence on the type and sources of selection under the German public health insurance with opt-out scheme. Since private insurers are allowed to discriminate prices based on some observables, and since the price of public insurance does not depend on these observables, it becomes necessary to distinguish between price discrimination based on observables and selection. I separate the two by estimating a bivariate probit model of health insurance choice and health care use in the spirit of the *positive correlation test* (Chiappori and Salanié 2000). To overcome the well-known issue that the *positive correlation test* does not distinguish between moral hazard and selection, I choose hospitalisations to measure health care use. I show that hospitalisations are not affected by health insurance status (public or private) in the German health insurance system, which allows me to attribute the test results to selection. I find that public insurance is heavily adversely selected in Germany. Subsequently, I turn to identifying sources of selection between public and private health insurance. I follow Finkelstein and Poterba (2014) and search for *unused observables*, variables that are observed by the econometrician but which the insurers do not observe or do not use in the premium calculation, that are correlated with both health insurance choice and health care use after controlling for observables. I find two *unused observables* that matter: self-assessed health gives rise to adverse selection against the public plan, and risk aversion leads to advantageous selection in favour of the public plan. The former finding is consistent with the idea that private health insurers are able to cream skim the best risks by offering contracts with less coverage than the public plan, which should appeal only to those with high self-assessed health. The latter finding confirms earlier studies which have found that individuals self-select based on risk aversion (Finkelstein and McGarry 2006, Buchmueller *et al.* 2013).

Finally, I show that while private insurers are able to attract individuals who have a high self-assessed health on average, they in fact attract two kinds of customers: individuals with high and very high self-assessed health and individuals with very low self-assessed health, where the former are more numerous. Public insurance, on the other hand, is more attractive to those who have self-assessed health which is neither high nor low. Such a nonmonotone pattern between health insurance choice and health status can arise in the theoretical model only under asymmetric information, which leads me to conclude that there is some form of information asymmetry between private insurers and their clients. However, the evidence in favour of this result is merely suggestive, given that the coefficients that reveal this pattern are imprecisely estimated.

This paper is related to several distinct literatures. The theoretical model relates closely to Olivella and Vera-Hernández (2013), who incorporate a public insurance plan into the canonical model of a competitive health insurance market due to Rothschild and Stiglitz (1976). The main difference between their model and the model presented in this paper lies in the financing of the public plan: Olivella and Vera-Hernández (2013) consider a tax-financed public plan, whereas I study a budget-balanced and contribution-financed public plan. The latter differs from the former in two important aspects. First, public insurance has a price attached to it. Second, this price is determined by the characteristics of those who choose public insurance.

This paper moreover contributes to the literature on the selection between a public option and competing private health insurance plans. One widely studied example in this literature is the choice of elderly citizens in the U.S. between traditional fee-for-service Medicare and private Medicare Advantage (MA) plans. Brown *et al.* (2014) document that individuals who switch into MA plans turn out to be significantly less costly than observationally equivalent individuals who stay in traditional Medicare. Newhouse *et al.* (2012) show that changes to the risk adjustment formula and restrictions on switches have reduced, but not eliminated, selection in the Medicare program. Beyond the U.S. context and in a health insurance system that is very similar to the one studied in this paper, Sapelli and Vial (2003) find that public health insurance in Chile is adversely selected. Two recent empirical studies analyse the German health insurance system. Grunow and Nuscheler (2014) study switches between public and private health insurance and find that individuals who have experienced a negative health shock show an increased propensity to switch from private to public insurance. Bünnings and Tauchmann (2014) investigate the decision to opt out of public insurance and find that young and healthy individuals are disproportionately more likely to opt out. Both studies point towards adverse selection against public insurance, but they do not relate health insurance choice with realised risk events and do not distinguish between variables that are observed by private insurers (and therefore affect premiums) and variables that are not observed by private insurers.

Finally, this paper is also related to the literature on testing for asymmetric information in insurance markets. The starting point for this literature has been the insight that, after controlling for observables, standard models of asymmetric information predict a positive correlation between insurance coverage and realised risk ex post (Chiappori and Salanié 2000). However, initial studies were unable to document such positive correlation and could therefore not reject the absence of symmetric information. Motivated by these findings, follow-up studies have considered multidimensional private information and selection on multiple contract dimensions. Cohen and Siegelman (2010) and Einav *et al.* (2010) survey the literature on testing for asymmetric information in insurance markets.

The rest of the paper is structured as follows. Section 2 lays out the theoretical model. Section 3 summarises relevant institutional details of the German health insurance system. Section 4 describes the data. Section 5 explains the empirical approach and reports the evidence on selection and asymmetric information. Section 6 concludes. All proofs are delegated to an Online Appendix.

2 A Model of Public Health Insurance with Opt-Out

This section introduces a model of public health insurance with opt-out. The model is geared towards the German health insurance system, which features income-dependent contributions to public insurance, a budget-balanced public plan, an opt-out policy that is based on income, and risk rating in private insurance. However, the model's result that private insurers are not always able to cream skim the best risks applies much more generally, to all parallel public and private health insurance systems with risk-unrelated contributions in public insurance and a market for private health insurance in which price discrimination is possible.

As in Olivella and Vera-Hernández (2013), a public insurance plan coexists along with a market for private insurance which is modelled as in Rothschild and Stiglitz (1976). Individuals must choose exactly one insurance plan, be it public or private. Private insurers observe some categorical information about the individual applicants and are allowed to discriminate prices based on these observables. The model describes the market for health insurance within a risk class, in which individuals are observationally equivalent except for the two features described below.

Within a risk class, there is a measure one of individuals who differ along two dimensions: their probability of becoming sick and their income. There are $n \geq 2$ different risks, $0 < p_1 < \ldots < p_n < 1$, where $p_i \in \{p_1, \ldots, p_n\}$ is the probability with which an illness occurs. Under symmetric information, a consumer's probability of falling sick is publicly observable, whereas it is private to the consumer under asymmetric information. In the case of falling sick, individuals suffer a monetary loss d, against which they can insure themselves. There are two levels of income, $y_H > y_L$, which are publicly observable. The right to opt out of public insurance is granted based on income. Individuals who earn the low income, y_L, are mandatorily insured in public insurance, while individuals who earn the high income, y_H, can stay in public insurance or opt out and buy private insurance instead.

Low-income earners play only a subordinated role, given that they do not choose their insurance contract. They are completely characterized by three parameters: their fraction of the population λ_L, $0 < \lambda_L < 1$, their income, y_L, and their average risk p_L, $p_1 \leq p_L \leq p_n$. The fraction of individuals who earn y_H and are of risk p_i is denoted by λ_i, where $\lambda_i > 0$ for all $i \in \{1, \ldots, n\}$. The fraction of high-income earners in the population is $\sum_{i=1}^{n} \lambda_i = 1 - \lambda_L$.

Henceforth, I refer to an individual with risk p_i and income y_H as a type p_i.

An insurance contract is a vector $I = (\alpha, \beta)$, where α is the insurance premium and β, $1 \geq \beta \geq 0$, is the co-insurance rate, i.e., the fraction of the damage which the insurer does not cover.[4] Expected utility of type p_i holding insurance policy $I = (\alpha, \beta)$ is given by

$$U(I, p_i) = p_i u(y_H - \beta d - \alpha) + (1 - p_i) u(y_H - \alpha),$$

where u is strictly increasing, strictly concave, and twice continuously differentiable.

The public insurance plan $I^{pub} = (\tau y, \eta)$ is announced at the outset and constitutes a committed offer. It consists of a contribution rate τ, which multiplied by the income yields the insurance premium, and a co-insurance rate η, which is the fraction of the damage that will be not be covered by the public insurer. Public health insurance entails two forms of redistribution: from the rich to the poor, as premiums increase with income, and from the healthy to the sick, as premiums do not depend on risk. Public insurance is financed through the contributions of its members. To maintain a balanced budget, the government must set the contribution and co-insurance rates such that revenues equal expected cost:

$$\tau\left(\lambda_L y_L + \sum_{i=1}^{n} \lambda_i s(p_i) y_H\right) = (1 - \eta) d \left(\lambda_L p_L + \sum_{i=1}^{n} \lambda_i s(p_i) p_i\right), \tag{1}$$

where $s(p_i)$ equals one (zero) if type p_i joins public insurance (private insurance). Note that all individuals who can choose between public and private health insurance (the high-income earners) pay the same premium for public insurance: τy_H.

After the public plan is announced, $m \geq 2$ private insurers simultaneously offer contracts. Private insurers are risk neutral, incur no administrative cost, and expect the following profit from selling the contract $I = (\alpha, \beta)$ to an individual with risk p_i:

$$\pi(I, p_i) = \alpha - p_i (1 - \beta) d.$$

Observing the menu of insurance plans available to them, high-income earners maximise their expected utility by choosing between public insurance and the best private contract which is available to them. The following tie-breaking assumption is made to avoid the indeterminacy of the equilibrium strategy profile that arises in the knife-edge case in which one type is indifferent between public and private insurance.

Assumption 1. *A high-income earner who is indifferent between public insurance and the best available private insurance contract chooses public insurance.*

Equilibrium of the health insurance market is defined as follows.

[4] The model could equivalently be formulated in terms of copayments.

Definition 1. An equilibrium is a strategy profile $s^* = [s^*(p_i)]_{i \in \{1,\ldots,n\}}$ and a set of contracts C^*, which includes the public plan, such that:

1. Every contract in C^* is selected by some consumer.
2. No contract in C^* yields negative profits.
3. There is no single contract outside of C^* that, if offered, will be selected by consumers and will generate nonnegative profits for the insurer.
4. For $i \in \{1,\ldots,n\}$: $s^*(p_i) = 1$ if $I^{pub} \in \arg\max_{I \in C^*} U(I, p_i)$, and $s^*(p_i) = 0$ otherwise.
5. The government budget is balanced: equation (1) holds for $s = s^*$.

The first three conditions are analogous to the equilibrium conditions in Rothschild and Stiglitz (1976). The fourth condition requires high-income earners to decide optimally between public and private insurance, incorporating the tie-breaking rule posited in Assumption 1. The fifth condition guarantees that the public budget is balanced in equilibrium. In light of the fact that the majority of high-income earners in Germany stay in public insurance, I consider only equilibria in which at least some high-income earners choose public insurance, i.e., $s^*(p_i) = 1$ for some $p_i \in \{p_1, \ldots, p_n\}$.[5]

Following Einav *et al.* (2010), I say that the public plan is adversely selected if the expected cost of insuring the high-income earners who choose public insurance in equilibrium is higher than the expected cost of insuring the population of high-income earners.

Definition 2. I^{pub} is adversely selected if $\mathbb{E}_p[pd \mid s^*(p) = 1] > \mathbb{E}_p[pd]$.

Conversely, I say that there is advantageous selection into public insurance if the reverse inequality holds.

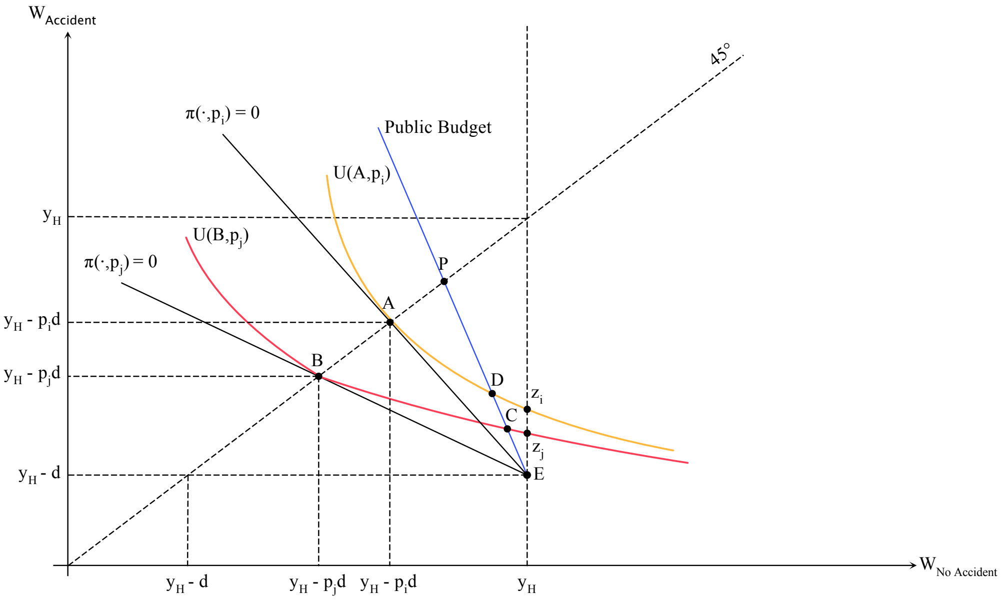
2.1 Equilibrium under Symmetric Information

Consider first the market for private health insurance. Under symmetric information, private insurers know the risk of each applicant and can offer a corresponding contract. Following Rothschild and Stiglitz (1976), competition drives insurance companies to offer n actuarially fair contracts, one for each type, each offering full insurance.

When deciding between public and private health insurance, the high-income earners take into account two factors: the co-insurance rate in public insurance and the public-private premium differential. The co-insurance rate matters because individuals are risk averse. The

[5]Under both symmetric and asymmetric information, equilibria exist in which all high-income earners purchase private insurance. For example, if the public plan offers no coverage at all ($\eta = 1$), then all high-income earners will choose private insurance under both informational scenarios.

Figure 1. *Insurance Market under Symmetric Information*



Notes: $A = (p_i d, 0)$ and $B = (p_j d, 0)$, respectively, are the private insurance contracts offered to two risks p_i and p_j, where $p_i < p_j$. For a given contribution rate τ, the public insurance contract lies on the line EP. The exact position on EP depends on the co-insurance rate of the public plan. Full insurance ($\eta = 0$) corresponds to P and no insurance ($\eta = 1$) to E. z_i and z_j are used in the proof of Lemma 1.

public-private premium differential is the difference in the price of one unit of coverage under public and private insurance. For a given type p_i, the premium differential is determined by the difference between incomes, which determines the amount of income redistribution, and the difference between p_i and the average risk in public insurance, which determines the gains from risk pooling.

The following Lemma shows that the incentives to join public insurance are stronger for individuals with bad health, the intuition being that these individuals experience higher gains from risk pooling in public insurance.

Lemma 1. *For any given public plan under symmetric information, if p_i joins public insurance, then all types $p_j > p_i$ join public insurance.*

After performing the standard change of variables (see the Online Appendix) in Figure 1, the insurance contracts of two types $p_i < p_j$ are depicted in the space of final wealths. In the proof of Lemma 1, I show that the indifference curve of p_i through his designated private insurance contract, A, lies strictly above the one of type p_j through his designated private contract, B. Assuming that the public insurance budget is balanced on the line running through the points E and P, we can see that both types join public insurance if the public plan is located on or above D, only p_j joins public insurance if the public plan is located on or above C and below D, and none of the two types joins public insurance if the public plan is located below C. This leads us to the health insurance market equilibrium.

Proposition 1. *Fix a co-insurance rate η for the public plan and suppose that some high-income earners choose to remain in public insurance. The health insurance market equilibrium under symmetric information, if it exists, is unique and characterised by a threshold type p_j, $p_1 < p_j \leq p_n$. The equilibrium strategy profile is such that $s^*(p_i) = 0$ for all $p_i < p_j$ and $s^*(p_k) = 1$ for all $p_k \geq p_j$, and the equilibrium set of contracts is the following:*

$$C^* = \left\{ I^{pub} = \left((1-\eta)\, d \frac{\lambda_L p_L + \sum_{k=j}^{n} \lambda_k p_k}{\lambda_L y_L + \sum_{k=j}^{n} \lambda_k y_H} y, \eta \right), I^{priv} = (p_i d, 0)_{i \in \{i : s^*(p_i) = 0\}} \right\}.$$

The public plan is adversely selected in this equilibrium, given that only the highest risks remain in public insurance. Formally, $\mathbb{E}_p[pd \mid s^*(p) = 1] = \mathbb{E}_p[pd \mid p \geq p_j] > \mathbb{E}_p[pd]$, where the inequality follows from the fact that $p_j > p_1$. The healthiest high-income earners (type p_1) do not join public insurance in equilibrium because they cannot gain from risk pooling and lose from income redistribution under the public plan.

There are parameter constellations in which the government cannot set a budget-balancing contribution rate for a given amount of coverage in public insurance, meaning that no equilibrium according to Definition 1 exists. This scenario arises when a certain type p_i wants to join public insurance at some given contribution and co-insurance rates, but if the government adjusts the contribution rate to balance the budget for when p_i joins public insurance, p_i no longer wants to join public insurance at the new contribution rate. The non-existence of equilibrium is unrelated to the tie-breaking assumption made above. It is due to the discreteness of the type distribution, which implies that the entry of any type into public insurance has a discrete effect on the public budget. A straightforward solution to the non-existence problem is to adopt a continuous distribution of types. However, the model becomes intractable with a continuum of types in the asymmetric information case, and the conclusions (at least in the symmetric information case) are not affected. My preferred solution is to assume that if the parameter constellation is such that the scenario described above arises, the government does no longer maintain a perfectly balanced budget, but rather finances the budget gap through a lump-sum payment (a tax).[6] I refrain from modelling this scenario explicitly.

2.2 Equilibrium under Asymmetric Information

Now consider the case in which the insureds are privately informed about their personal risk p_i, while the insurers only know the distribution of risks in the population, $(\lambda_i)_{i \in \{1,\ldots,n\}}$. Let us focus first on the private sector and abstract away from the public plan for a moment. From Rothschild and Stiglitz (1976), we know that, in equilibrium, less risky types will be

[6] In fact, a small part (<5%) of public health insurance in Germany is financed through general taxation.

offered contracts with higher co-insurance rates, competition drives profits on each contract down to zero, and contracts have to be separating (otherwise, there would exist a profitable deviation that lies in cream-skinning the good risks out of a pooling contract). We also know that local (downward) incentive-compatibility constraints are sufficient for separation if expected utility satisfies the Spence-Mirrlees strict single crossing property.[7] Finally, we know that an equilibrium in pure strategies may fail to exist.[8]

The equilibrium properties of the private contract schedule remain basically unchanged in the presence of a public insurance plan. The only difference is that, with a coexisting public plan, some incentive-compatibility constraints are not determined by making a type indifferent between two private contracts, but between a private contract and the public plan.

When deciding between public and private insurance, high-income earners face the same trade-off as under symmetric information. They weigh the difference in coverage between public and private insurance against the public-private premium differential. Only that now, under asymmetric information, coverage in private insurance is no longer necessarily higher than coverage in public insurance. This feature gives rise to the possibility of advantageous selection into public insurance.

The set of empirically relevant equilibria (those in which some high-income earners remain in public insurance) is characterised below.

Proposition 2. *Fix a co-insurance rate η for the public plan and suppose that some high-income earners choose to remain in public insurance. The health insurance market equilibrium under asymmetric information, if it exists, can feature two scenarios:*

1. *Adverse selection: there exists a threshold type p_j, $p_1 < p_j \leq p_n$,* such that $s^*(p_i) = 0$ *for all $p_i < p_j$* and $s^*(p_k) = 1$ *for all $p_k \geq p_j$, and the equilibrium set of contracts is*

$$C^* = \left\{ I^{pub} = \left((1-\eta)\, d \frac{\lambda_L p_L + \sum_{k=j}^{n} \lambda_k p_k}{\lambda_L y_L + \sum_{k=j}^{n} \lambda_k y_H} y, \eta \right), I^{priv} = \left(p_i \left(1-\beta_i\right) d, \beta_i\right)_{i \in \{i: s^*(p_i)=0\}} \right\},$$

where $I_i^{priv} = (p_i(1-\beta_i)d, \beta_i)$ is the private contract which will be chosen by type p_i.

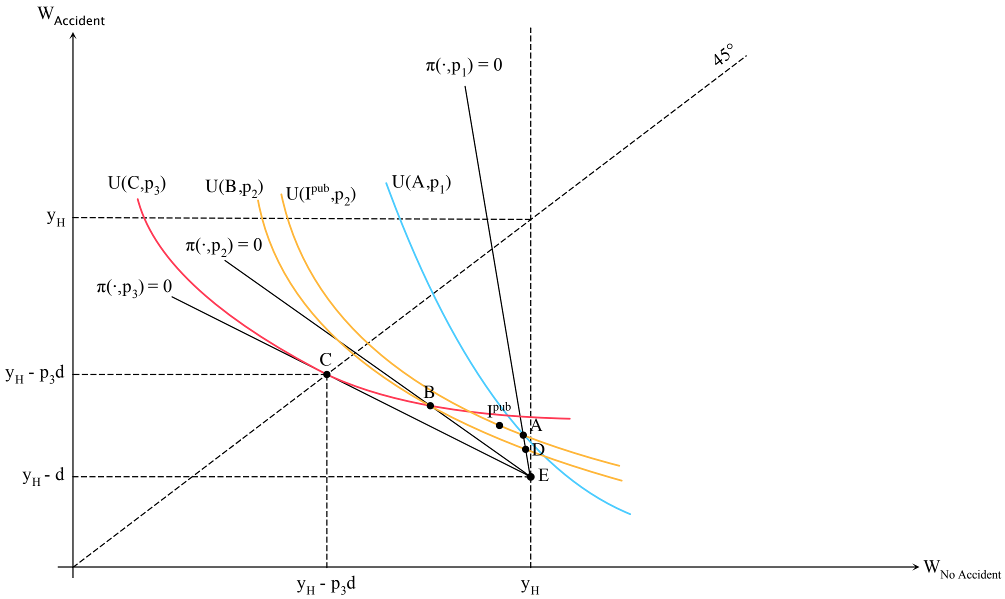
2. *Nonmonotone selection: there exist two threshold types, p_j and p_k, $p_1 < p_j \leq p_k < p_n$, such that $s^*(p_i) = 0$ for all $p_i < p_j$, $s^*(p_r) = 1$ for all $p_j \leq p_r \leq p_k$, and $s^*(p_l) = 0$ for all $p_l > p_k$, and the equilibrium set of contracts is*

$$C^* = \left\{ I^{pub} = \left((1-\eta)\, d \frac{\lambda_L p_L + \sum_{r=j}^{k} \lambda_r p_r}{\lambda_L y_L + \sum_{r=j}^{k} \lambda_r y_H} y, \eta \right), I^{priv} = \left(p_i \left(1-\beta_i\right) d, \beta_i\right)_{i \in \{i: s^*(p_i)=0\}} \right\},$$

[7] Which it does: $U_\alpha(I,\cdot)/\left|U_\beta(I,\cdot)\right|$ is strictly increasing for all I and has the same sign for all I and p.

[8] With 2 types, the unique equilibrium ceases to exist when there are too many low risks in the market. The conditions for equilibrium existence in the case of $n \geq 2$ types are more demanding. Riley (1985) identifies a set of sufficient conditions for this case.

Figure 2. *Equilibrium under Asymmetric Information with Fragmentation of Risks*



Notes: A, B, and C are the incentive-compatible, actuarially fair private insurance contracts offered to types p_1, p_2, and p_3, respectively, where $p_1 < p_2 < p_3$. I^{pub} is the public plan. Types p_1 and p_3 prefer private over public insurance. Type p_2 prefers public over private insurance. The incentive-compatible contract for type p_1, A, is designed such that type p_2 is indifferent between this contract and public insurance. In the absence of a public plan, private insurers would offer the contract D to type p_1.

where $I_i^{priv} = (p_i(1-\beta_i)d, \beta_i)$ *is the private contract which will be chosen by type* p_i.

The equilibrium schedule of private contracts satisfies the following: (i) uniqueness: I^{priv} *is unique, (ii) no distortion at the top: if* $s^*(p_n) = 0$*, then* $\beta_n = 0$, (iii) *distortion at the bottom: if* $s^*(p_i) = 0$ *and* $p_i < p_n$*, then* $\beta_i > 0$*, (iv) incentive compatibility: if* $s^*(p_i) = s^*(p_{i+1}) = 0$*, then* $U[(p_{i+1}(1-\beta_{i+1})d, \beta_{i+1}), p_{i+1}] = U[(p_i(1-\beta_i)d, \beta_i), p_{i+1}]$, and *if* $s^*(p_i) = 0$ *and* $s^*(p_{i+1}) = 1$*, then* $U(I^{pub}, p_{i+1}) = U[(p_i(1-\beta_i)d, \beta_i), p_{i+1}]$*, (v) monotonicity: if* $s^*(p_i) = s^*(p_j) = 0$ *and* $p_j > p_i$*, then* $\beta_j < \beta_i$*, and (vi) positive insurance: if* $s^*(p_i) = 0$*, then* $\beta_i < 1$.

The first scenario is the only outcome which is possible when all market participants are symmetrically informed. The second scenario, which I call nonmonotone selection, occurs exclusively under asymmetric information. It corresponds to a situation where some bad risks, who are offered relatively generous coverage in private insurance, take up private insurance, while some less riskier types join public insurance, because they are offered only relatively little coverage by private insurers. The reason for why this scenario can occur in the presence of asymmetric information is simple. In response to having inferior information than clients, private insurers offer screening contracts, which include less coverage for the good risks. These contracts introduce a new trade-off for the high-income earners, who are now no longer guaranteed full insurance in the private sector.

Figure 2 illustrates the third scenario for the case of three types, $p_1 < p_2 < p_3$. In equilibrium, type p_2 chooses public insurance, and types p_1 and p_3 purchase private insurance. The equilibrium set of contracts is given by $C^* = \{I^{pub}, A, C\}$. Selection into public insurance can be adverse or advantageous, depending on the fractions of the three types. Formally, $\mathbb{E}_p[pd \mid s^*(p) = 1] = p_2 d$ and $\mathbb{E}_p[pd] = (\lambda_1 p_1 + \lambda_2 p_2 + \lambda_3 p_3)d/(\lambda_1 + \lambda_2 + \lambda_3)$, so that the public plan is adversely selected if $\lambda_1(p_2 - p_1) > \lambda_3(p_3 - p_2)$ and advantageously selected if the reverse inequality holds. This example generalises to any number of types.

One special feature of the model with a parallel public and private health insurance market is that some types can obtain more comprehensive coverage than in a purely private health insurance market. In the example of Figure 2, the incentive-compatible contract that can be offered to type p_1 is A. In the absence of a public plan, private insurers must reduce the coverage for type p_1 to the level of contract D, in order to guarantee incentive compatibility. This interaction between public and private health insurance does not arise under symmetric information, where the public plan solely determines which of the private contracts are chosen in equilibrium.

Against the common expectation that private health insurers will cream skim the best risks and leave the bad risks in public insurance, the model shows that the selection between public and private insurance depends on the information scenario. Corollary 1 summarises the model's predictions, which provide the basis for the one-sided test of asymmetric information performed below.

Corollary 1. *Under symmetric information, health status and health insurance choice are monotonously related and lead to adverse selection against the public sector. Under asymmetric information, the relationship between health status and health insurance choice can be monotone or U-shaped, and the public sector can be adversely or advantageously selected.*

3 Public Health Insurance with Opt-out in Germany

This section gives an overview over the German health insurance system and describes the institutional features that are relevant for the analysis to follow.

About 90% of the German population holds public health insurance, which is largely financed through the contribution of its members. A small part of the cost of public insurance (less than 5%) is financed through general taxation. Public health insurance is provided by sickness funds, which operate on a non-profit basis. For roughly 75% of the population, public insurance is mandatory. The remaining 25% can decide between opting out of public insurance and remaining *voluntarily insured* in the public system. About 10% of the German population has decided to opt out of public insurance and holds substitutive private health insurance, which is provided by for-profit insurance companies. Opting out of public insurance becomes possible when gross labor income is above the so-called *compulsory insurance threshold* (53,550 Euros in 2014), or when an individual is exempted based on his occupation. The two most important groups that are exempted from the public insurance mandate are civil servants and individuals in self-employment. Once an individual has decided to opt out of public insurance, reentry into the public system is restricted and becomes possible only when both criteria that determine the eligibility to opt out are no longer satisfied.[9] There are virtually no uninsured individuals. This is in part due to the fact that Germany imposes an individual health insurance mandate since 2009. However, the number of uninsured has historically always been low, irrespective of the mandate.

The most important difference between public and private insurance lies in the premium calculation. Premiums in public insurance depend only on labor income: the publicly insured pay a fraction of their gross wage (8.2% in 2014) up to the *contribution ceiling*, above which contributions are zero.[10] Premiums in private insurance are risk rated and not tied to income. They are fixed at the initial enrollment and cannot be adjusted in response to health events after the signing of the contract, implying that private insurance clients are protected against reclassification risk. Private contracts are lifetime contracts which cannot be cancelled by the insurer, unless premiums are not paid.

Risk rating in private insurance is conducted on the basis of mandatory health questionnaires.[11] Most insurers elicit the same information from potential consumers. Applicants have to report height and weight, disability status, chronic diseases, pregnancy status, a potential HIV infection, ambulatory treatments within the last three years, stationary treatments within the last five years, prescribed pharmaceuticals within the last 3 years, absences from work which have lasted longer than 14 days during the last three years, and psychological therapies within the last 10 years. Based on this information, insurers can apply risk surcharges or deny coverage for certain diseases and chronic conditions, such as diabetes, cancer, or hypertension.

Contract customisation is another important aspect in which public and private insurance differ. Private insurance customers can select a contract that is individually optimal. Private insurers typically offer several benefits packages and let customers choose the level of their annual deductible; some private contracts also include co-insurance rates. Public insurance, on the other hand, offers little room for consumer choice. Almost all of the benefits that are covered by public health insurance are dictated by the regulator, and consumers do not have much choice in changing the cost-sharing rules of their contract. The standard contract in

[9] Individuals aged 55 years or older can never switch from private to public health insurance.

[10] The employee contribution is matched with a contribution of similar size from the employer (7.3% in 2014). The employer pays this contribution also if the employee is privately insured, which means that the employer contribution does not distort the choice between public and private health insurance.

[11] As false reporting in the questionnaires can lead to withholding of benefits or termination of the contract, private insurance applicants have strong incentives to answer truthfully.

public health insurance imposes little out-of-pocket expenses on its customers, mostly in the form of moderate co-payments for pharmaceuticals and hospital stays.

Besides health status, a few other personal characteristics determine the relative price between public and private insurance. The relative price of public insurance decreases with the age of the insured for two reasons. First, health deteriorates so that risk surcharges may apply. Second, private insurers are legally mandated to build up old-age provisions in order to keep premiums constant over the life-cycle. As the time period over which these old-age provisions can be built up becomes shorter the older the applicant, premiums have to increase. Public insurance is relatively cheap for families because non-working spouses and dependent children below 26 years of age are insured free of charge in public insurance. Women face a lower relative price of public insurance than men because private insurers charge women higher premiums throughout the sample period.[12] Civil servants pay a high relative price for public insurance because they and their dependent family members receive partial reimbursements of medical expenses through the so-called *Beihilfe*, which reduces the cost of private insurance by up to 80% but does not affect the price of public insurance. The relative price of public insurance increases with the income of self-employed individuals and civil servants until the *contribution ceiling* is reached. Individuals who qualify to opt out of public insurance based on their income pay a fixed premium for public insurance, which does not depend on income.[13]

4 Data and Descriptive Statistics

The empirical analysis is based on data from the German Socio-Economic Panel, a long-running, annual panel which elicits information from a representative sample of the population living in Germany. At the individual level, the SOEP contains subjective and objective health measures, health insurance details, as well as a wide array of socio-economic variables.

The sample period consists of the years from 1999 to 2011. Observations from prior to 1999 had to be excluded because important covariates are missing. The sample consists of individuals aged 20 to 60 who are eligible to purchase private insurance, not insured through one of their family members, and not receiving subsidies towards private insurance through being a civil servant or a dependent relative of a civil servant. Individuals under 20 years of age are excluded to avoid distortions which may arise due to family insurance. Individuals above the age of 60 are excluded to avoid distortions arising from various retirement schemes. Individuals who are insured through one of their family members are not considered because they do not make an active decision. Civil servants and their relatives are excluded because

[12]However, as of December 2012, gender-based discrimination of insurance premiums is prohibited in the European Union by a ruling of the European Court of Justice.

[13]This is because the *compulsory insurance threshold* lies above the *contribution ceiling*.

Table 1. *Descriptive Statistics of Selected Variables*

Health measures	Public	Private	Socio-economic factors	Public	Private
Self-assessed health	3.594	3.717	Number of children	0.835	0.707
	(0.831)	(0.793)		(1.049)	(0.961)
Disability	0.054	0.021	Married	0.836	0.776
	(0.225)	(0.143)		(0.371)	(0.417)
Chronic illness	0.315	0.214	Female	0.254	0.220
	(0.465)	(0.410)		(0.436)	(0.414)
Hospitalised within a year	0.076	0.063	Age	44.737	44.141
	(0.265)	(0.243)		(8.535)	(8.358)
Doctor visits per quarter	1.785	1.477	Income	54,307	65,566
	(3.078)	(2.948)		(36,732)	(66,023)
Sick leave >6 weeks	0.033	0.022	Self-employed	0.378	0.611
within a year	(0.178)	(0.147)		(0.485)	(0.488)
Observations	16,881	9,813	Observations	16,881	9,813

Notes: Means and standard deviations (in parentheses) are calculated based on all observations from 1999 to 2011 of individuals aged 20 to 60 who are eligible to purchase private insurance, not insured through one of their family members, and not receiving subsidies towards private insurance as a civil servant. Two-sided t-tests reject the null hypothesis of no difference in means between public and private insurance at the 1% level of significance for all variables. Number of observations varies by variable and sample.

their choice of health insurance is fundamentally different from the rest of the population due to the subsidies they receive. The sample consists of 26,694 person-year observations from 6,701 individuals. Individuals in the sample are either self-employed or dependent workers with incomes above the *compulsory insurance threshold*.

Table 1 contains means and standard deviations (in parentheses) of selected variables. From these statistics, we can infer that private and public insurance clients are remarkably different in terms of health status. The privately insured report to be of better health, are less likely to be disabled, have a lower prevalence of chronic illnesses, and are less likely to take a sick leave of more than six weeks. Moreover, the privately insured use less medical services: they go less often to the doctor and are less likely to be hospitalised.

Private and public insurance clients differ also with respect to their socio-economic characteristics. Private insurance clients are less likely to be married or female, and, on average, they are younger and have less children. These factors are hardly a surprise given the institutional incentives. The privately insured also have substantially higher incomes, which may be due to the fact that the income-rich can more easily afford to buy coverage for their dependent family members. Lastly, there is a higher share of self-employed individuals among the privately insured than among the publicly insured, indicating that private insurance is more attractive on average to the self-employed than public insurance.

5 Evidence on Selection and Asymmetric Information

In this section, I present evidence on the nature and magnitude of selection between public and private health insurance in Germany. I show that the public sector is heavily adversely selected, and I identify two sources of selection. Last but not least, I present evidence which suggests that private insurers and their clients are asymmetrically informed.

5.1 Empirical Approach

The first goal of the analysis to follow is to distinguish between price discrimination based on observables and selection, and to quantify their relative importance. To this end, I include observables that determine the relative price between public and private health insurance as independent variables in a bivariate probit model of health care utilisation and health insurance choice, similar in spirit to the standard *positive correlation test* for asymmetric information (Chiappori and Salanié 2000). Price discrimination based on observables, also referred to as selection on observables, is then captured by the coefficient vector associated to the observables, and selection is captured by the correlation of the error terms.

However, as is well known, the *positive correlation test* does not distinguish between moral hazard and adverse selection. To address this issue, I choose a measure of health care use that is arguably less susceptible to moral hazard: hospitalisations.[14] The validity of my empirical approach hinges on the identifying assumption that, after controlling for observables, hospitalisations do not depend on health insurance status (public or private) other than through selection, which I assume to be the case in the following.

There are a couple of reasons why the identifying assumption is reasonable in the context of the German health insurance system. First, the cost of a hospitalisation is roughly the same for a privately and a publicly insured patient. The publicly insured pay a moderate copayment of 10 Euros per day of hospital stay, up to a maximum of 280 Euros a year. In private insurance there are typically no copayments for hospital stays. However, many private contracts include a deductible, implying that the cost of a hospitalisation in private insurance is generally different from zero. The most popular deductibles are, in that order, 0 Euros, 300

[14]Early empirical support for this hypothesis stems from the RAND Health Insurance Experiment, which does not find evidence of moral hazard in inpatient spending (Manning *et al.* 1987). A few studies have capitalised on this result in order to test for adverse selection in health insurance (Olivella and Vera-Hernández 2013, de la Mata *et al.* 2014). On the other hand, a couple of recent studies point towards some degree of moral hazard in hospitalisations (Finkelstein *et al.* 2012, Dardanoni and Li Donni 2012). However, the treatment and control group differ substantially in the contexts of these studies. As I will argue below, the difference between public and private health insurance in the case of Germany is much less pronounced.

Euros, and 600 Euros. On average, it likely that the difference in the cost of a hospitalisation between public and private insurance is, if anything, relatively small. The small difference in price implies that, even if people have a price elasticity of demand for hospitalisations which is different from zero, the moral hazard effect of health insurance should be modest. Second, several studies have examined the incentive effect of private health insurance coverage on hospitalisations in Germany, all of which conclude that there is no statistically significant effect. The most convincing study is the one of Hullegie and Klein (2010), who exploit the *compulsory insurance threshold* in a regression discontinuity design, with which the authors can control for selection into public and private health insurance. Using SOEP data, Hullegie and Klein (2010) find no incentive effect of health insurance status on hospital stays. Two earlier studies use count data methods and also do not find an effect of the type of health insurance on hospitalisations (Geil *et al.* 1997, Riphahn *et al.* 2003).

Following Bünnings and Tauchmann (2014), I model health insurance choice as a hazard model in discrete time with private insurance as the absorbing state: in each period, an individual either *survives* (stays in public insurance), *fails* (switches to private insurance), or drops out of the sample (no longer qualifies to purchase private insurance). German legislation dictates the choice of such a hazard model as it generally restricts private insurance clients to stay in private insurance. During the sample period, 741 individuals opt out of public insurance and become privately insured. It should be noted that the hazard model implies a considerable drop in the effective number of observations, given that individuals who have private insurance from the first period when they enter the sample are not included in the estimations, and given that observations of individuals who opt out during the sample period are exploited only up to the point at which they opt out.

After presenting the results from the basic bivariate probit model of health insurance choice and health care use, I turn to identifying sources of selection following the approach of Finkelstein and Poterba (2014), and I test for asymmetric information. To conclude the empirical part, I discuss potential confounding factors and offer a sensitivity analysis.

5.2 Magnitude and Sources of Selection

I estimate variants of the following bivariate probit model:

$$\text{OPTOUT}_{it} = \mathbf{1}\left(a_1 + b_1 X_{it} + c_1 Z_{it} + e_{1it} > 0\right). \tag{2}$$

$$\text{HOSPITAL}_{it+1} = \mathbf{1}\left(a_2 + b_2 X_{it} + c_2 Z_{it} + e_{2it} > 0\right). \tag{3}$$

OPTOUT_{it} is a binary variable that captures the choice of individual i to opt out of public insurance in period t conditional on being eligible to opt out. HOSPITAL_{it+1} is a binary variable that equals one if individual i is hospitalised in period $t+1$. X_{it} is a vector of observable attributes of individual i that determine the relative price between private and public insurance at time t.[15] Z_{it} is a vector of *unused observables*, variables which the econometrician observes but which the insurers do not observe or do not use in the premium calculation. I consider the following five *unused observables*: self-assessed health, risk aversion, smoking status, weekly hours of work, and income. $e_{it} = (e_{1it}, e_{2it})$ is a vector of error terms.

The error terms $e_{it} = (e_{1it}, e_{2it})$ are assumed to be independently, identically, and normally distributed across individuals. They may, however, be correlated over time for a given individual. In the main analysis, I will estimate a pooled bivariate probit model and compute the standard errors using a block bootstrap procedure with clustering at the individual level, allowing for valid inference in the presence of serially correlated errors. Because some individuals, when they become eligible to opt out of public insurance, may decide to stay in public insurance once and for all, without reevaluating the relative price between public and private insurance in future periods, I will, as a robustness check, reestimate the model on the subsample of individuals who are in their first or second year of being eligible to opt out. Notwithstanding this, all individuals in the sample can opt out of public insurance any time, where the price of opting out is determined by the observables.

In a first step, I will estimate the model with $b = (b_1, b_2) = 0$ and $c = (c_1, c_2) = 0$ imposed, which yields a measure of the sum of the effects of price discrimination based on observables and selection. Subsequently, I will estimate the model with $c = 0$ imposed. With this specification, price discrimination based on observables is captured by the coefficient vector b, and selection is captured by the correlation of the error terms. Finally, I will estimate the full model, without any restrictions imposed, where the coefficient vector c identifies potential sources of selection on unobservables.

Table 2 reports the correlation of the residuals for the three variants of the bivariate probit model. From column 1 of Table 2, we can infer that the combination of price discrimination based on observables and selection implies that the public sector is adversely affected: individuals who opt out of public insurance in period t are significantly less likely to be hospitalised in period $t + 1$. The estimated correlation coefficient of column 2 shows that, when we control for observable differences across individuals that affect the relative

[15] X contains disability status, an indicator of chronic illness, the number of hospitalisations within the last five years, doctor visits per quarter within the last three years, an indicator of body mass index (BMI) above 30, absences from work for more than 6 weeks within the last three years, number of children, marital status, gender, an indicator of being a women of child-bearing age, four indicators capturing age, an indicator of self-employment, an interaction term between self-employment and the natural logarithm of income, and a set of year dummies. The year dummies are included to capture changes of the public contribution rate, which occur almost every year. All variables, apart from the chronic illness indicator and the obesity indicator, are predetermined. The chronic illness indicator equals one if an individual reports to suffer from a chronic disease in at least one of the three years from 2009 to 2011, which are the only years for which this information is available. The obesity indicator equals one if the average BMI over the years 2002, 2004, 2006, 2008, and 2010 of a person is above 30. To the extent that chronic conditions and BMI do not fluctuate much over time, these two indicators represent a reasonable approximation of the missing values.

Table 2. *Magnitude of Selection*

	No controls (1)	Control for observables (2)	Control for observables and *unused observables* (3)
$\hat{\rho}$	-0.149***	-0.113**	-0.111**
LR test of $\rho = 0$ (p-value)	0.002	0.033	0.041
N	8,301	8,301	8,301

Notes: Table reports the correlation of the residuals from bivariate probit estimations of equations (2) and (3). Column 1 reports the correlation for the model with $b = c = 0$. Column 2 reports the correlation for the model with $c = 0$. Column 3 reports the correlation for the model with five *unused observables*. All p-values are bootstrapped (1000 replications). * p<0.1, ** p<0.05, *** p<0.01.

price between private and public insurance, there remains less unexplained correlation between the choice to opt out and future hospitalisations, as one would expect. The difference between the estimates of the correlation coefficients in columns 1 and 2 measures the effect of the price discrimination based on observables. The selection effect is roughly three times as large as the effect of the price discrimination based on observables (-0.113 vs. -0.036). The five *unused observables* do not seem to explain much of the observed adverse selection, considering how similar are the estimated correlation coefficients of columns 2 and 3. However, as we will see below, some *unused observables* are a source of adverse selection against the public sector (implying a less negative $\hat{\rho}$), while other *unused observables* are a source of advantageous selection in favour of the public sector (implying a more negative $\hat{\rho}$).

I do not discuss in detail the effect of price discrimination based on observables. Most of the observables have a foreseeable impact insofar that they lead to adverse selection against the public sector. This is in particular true for the health-related observables that private insurers elicit, such as the existence of chronic conditions and the number of past doctor visits. The interested reader is referred to Table 5 in the Appendix, which contains the estimates for the coefficient vector b.

Table 3 reports the results on the sources of selection. *Unused observables* whose estimated coefficients are of the opposite (same) sign across the two outcomes and statistically significantly affect both outcomes constitute a source of adverse (advantageous) selection against (in favour of) the public sector. The preferred specification is shown in column 6. From this specification, we can infer that two *unused observables* matter for the selection between public and private health insurance: self-assessed health and risk aversion.

Self-assessed health is reported on a five-point scale from one (=bad) to five (=excellent)

Table 3. *Sources of Selection*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: OPTOUT						
Self-assessed health	0.008**					0.010***
	(0.003)					(0.003)
Risk aversion		-0.005***				-0.004***
		(0.001)				(0.001)
Smoking			0.015***			0.015***
			(0.005)			(0.005)
Weekly hours of work				-0.002**		0.000
				(0.001)		(0.001)
Log income					-0.019***	-0.022***
					(0.004)	(0.004)
Dependent variable: HOSPITAL						
Self-assessed health	-0.018***					-0.019***
	(0.004)					(0.005)
Risk aversion		-0.005**				-0.006***
		(0.002)				(0.002)
Smoking			0.004			-0.000
			(0.007)			(0.007)
Weekly hours of work				0.003**		0.003
				(0.001)		(0.002)
Log income					0.002	-0.002
					(0.006)	(0.007)
N	9,014	9,017	8,401	8,918	9,017	8,301

Notes: Table reports average marginal effects of the *unused observables* from bivariate probit estimations of equations (2) and (3). Means of OPTOUT and HOSPITAL vary across columns. Bootstrapped standard errors (1000 replications) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

and should correlate with both the health insurance choice of forward-looking consumers and future hospitalisations. In fact, we see that self-assessed health explains a part of the adverse selection against the public sector: after controlling for observables, individuals with higher self-assessed health have a lower risk of being hospitalised and are more likely to choose private insurance. This self-selection based on self-assessed health is consistent with the idea that private insurers screen their customers. Healthy individuals choose private contracts with high deductibles, in return for low insurance premiums, whereas sick individuals prefer public insurance, which involves moderate cost sharing. One caveat applies to this result, however. The observables which are included in X can potentially not control for all the observable differences across individuals which affect the relative price between public and private health insurance. As such, it could be that self-assessed health picks up some of the effects of the observables, meaning that the coefficients above are an upper bound on the impact of the unobservable (for insurers) part of health status on health insurance choice and hospitalisations. I will return to this issue in Section 5.4.

Risk aversion is reported on a scale from zero to ten, where individuals are asked to assess their aversion towards risk in general.[16] I follow Bünnings and Tauchmann (2014) and others and consider risk preferences as fixed over time, using the average value of an individual's responses. Risk aversion has been shown to correlate with both health insurance choice and health care utilisation after controlling for observables (Finkelstein and McGarry 2006, Buchmueller *et al.* 2013). It is therefore not surprising to see that risk aversion also is a source of selection in the present context. Table 3 shows that risk aversion gives rise to advantageous selection in favour of the public sector: after controlling for observables, risk-averse individuals are less willing to opt out of public insurance and less likely to be hospitalised. Regarding the choice of health insurance in Germany, there are two reasons to believe that risk-averse individuals prefer public over private insurance. First, public insurance exposes insureds to less financial risk on average, since out-of-pocket costs are on average lower in public insurance. Second, switching to private insurance implies uncertainty about future premiums, as changes in family status translate into premium changes in private insurance. For example, a privately insured couple who become parents has to pay for their child in private health insurance, whereas the child is insured free of charge in public insurance. The estimates confirm the findings from other contexts that risk-averse individuals tend to be less risky, in the sense that they are less likely to be hospitalised. Possible explanations for this result are that the risk-averse use more preventive care, or that they abstain from undertaking risky activities. The finding of advantageous selection based on risk aversion is also mirrored in the findings of Schmitz (2011), who shows that risk-adverse males advantageously select supplementary insurance for hospital stays in Germany.

Smoking, hours of work, and income do not represent sources of selection because they

[16] Dohmen *et al.* (2011) confirm that this question is a good measure of risk aversion in several domains.

are not statistically significantly correlated with both health insurance choice and hospitalisations after controlling for observables. While hours of work do significantly affect health insurance choice when they enter as the only *unused observable*, the effect is likely due to a spurious correlation with income. Once income is controlled for, the effect of hours of work on health insurance choice vanishes. For the same reason, we can expect the estimated positive coefficient of income on hospitalisations in column 4 to be biased, because hours of work are omitted in this specification. The preferred estimates of column 6 suggest that earning more and working fewer hours are inversely related to the probability of being hospitalised, as one would commonly expect. There is no obvious explanation for the strong negative (positive) relationship between income (smoking) on the one hand and the choice of private insurance on the other hand. Finally, we can conclude that hours of work have no impact on health insurance choice, and that smoking is unrelated to hospitalisations.

5.3 Testing for Asymmetric Information

The standard test of asymmetric information analyses the consumer's decision between two different insurance contracts, where one contract offers more generous coverage than the other, and the occurrence of subsequent risk events (Chiappori and Salanié 2000). The null hypothesis of symmetric information is rejected if the correlation between the two error terms that are associated to insurance choice and risk occurrence is statistically different from zero. In the present setting, however, individuals choose between public and private health insurance, which are not ordered by coverage, albeit public insurance offers more coverage on average. We can therefore not proceed as usual.

Nevertheless, the theoretical model of public health insurance with opt-out offers a one-sided test of asymmetric information. Recalling Corollary 1, we see that both advantageous selection into public insurance and a U-shaped relationship between the choice of private insurance and unobservable (to the insurers) health status occur only under asymmetric information. Hence, either of the two indicates the presence of asymmetric information.

A positive correlation of the residuals after controlling for observables signals the presence of advantageous selection into public insurance. But, the residuals are not positively correlated, as we have seen in Table 2. Nevertheless, we may still conclude that there is asymmetric information, in case we find a U-shaped relationship between health insurance choice and unobservable health status. Self-assessed health is a natural candidate to test for such a relationship, given that it inherently is private information of the assessors. In order to pick up a potential nonmonotone effect of self-assessed health, I replace the previously considered linear term with indicators for the five categories of self-assessed health.

Table 4 shows the estimated average marginal effects of the five indicators of self-assessed health after controlling for observables. The probability that a hospitalisation occurs strictly

Table 4. *Testing for Asymmetric Information*

Dependent variable:	OPTOUT		HOSPITAL	
SAH=1	0.005	(0.123)	0.023	(0.027)
SAH=2	-0.006	(0.011)	0.018	(0.011)
SAH=3 (omitted)	—	—	—	—
SAH=4	0.006	(0.005)	-0.026***	(0.007)
SAH=5	0.019***	(0.007)	-0.028**	(0.013)
N	9,014		9,014	

Notes: Table reports average marginal effects of the five categories of self-assessed health from a bivariate probit estimation of equations (2) and (3). 1=bad, 2=poor, 3=satisfactory, 4=good, 5=excellent. Means of *OPTOUT* and *HOSPITAL* are 0.041 and 0.085, respectively. Bootstrapped standard errors (1000 replications) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

decreases in self-assessed health, meaning that the residual information contained in self-assessed health indeed is predictive of health status. On the other hand, the probability of opting out of public insurance is nonmonotonously related to self-assessed health: the relationship is U-shaped, where the probability of opting out of public insurance is lowest for individuals with a self-assessed health status of two (=poor). According to the theoretical model, this U-shaped relationship indicates the presence of asymmetric information between private insurers and their clients. However, because there are few observations with self-assessed health equal to one, the estimates for this category are very imprecise. Therefore, the evidence in favour of information asymmetry should be considered as merely suggestive. More encouragingly, the U-shaped pattern between the choice to opt out and self-assessed health is not affected by the potential issue that self-assessed health reflects some information that is observable to private insurers but not observed by the econometrician. If anything, the U-shaped pattern is reinforced, because the estimates would be biased downwards for the categories one and two and biased upwards for the categories four and five.

5.4 Discussion

This section discusses potential threats to the validity of the empirical approach and examines the robustness of the conclusions from the preceding two sections to these threats. The detailed results of this sensitivity analysis are collected in Tables 6-8 in the Appendix. The qualitative conclusions do not change in any of the alternative specifications. Due to large variations in the number of observations under different specifications and the resulting loss in statistical power, the estimates' statistical significance departs occasionally from those of the baseline estimates.

The most pressing issue is the potential mismatch between the actual observables that determine the relative price between public and private and the observables which are taken into account in the empirical analysis. Some variables in the SOEP are less informative than the corresponding questions from the insurer questionnaires. For example, private insurers know not only whether a hospitalisation occurred in the past, but also the reason for the hospitalisation. Moreover, some items from the insurer questionnaires are not included in the SOEP questionnaire, or they are included only in some years. For this reason, private insurers can be expected to have more information about insurance applicants than what is controlled for with the vector of observables in the preceding two sections. A clear-cut distinction between the price discrimination based on observables and the selection on unobservables is therefore not possible. Instead, the estimates of the selection on observables and unobservables should be viewed as a lower and an upper bound, respectively, of the actual effects. What can be concluded irrespective of the potential mismatch is that the public sector is adversely affected, given that individuals adversely select public insurance on both observables and unobservables. To assess how sensitive the results of the preceding sections are to the specification of the observables, I add two additional insuree characteristics to the vector of observables: a physical component scale score and a mental component scale score, which are both described in detail in Andersen *et al.* (2007). The two scales are available only for the even-numbered years from 2002 onwards, and I linearly impute them for the odd-numbered years in between. Almost certainly, private insurers will not use these scales to calculate premiums. But, they may pick up some of the effects of variables that are missing in the SOEP data but observed by insurers.

Another issue, which has also been mentioned before, is that some individuals may only actively consider opting out of public insurance in the first few periods when they become eligible to opt out. In order to show that the pooled estimates are not driven by repeated observations on individuals who do not make an active decision between public and private insurance, I reestimate the models on the subsample of individuals who are in the first or second year of being eligible to choose private insurance.

To assess the measure of health care use chosen for the baseline model, which is an indicator that equals one if a person is hospitalised in the year after the making choice between public and private insurance, I consider two alternatives dependent variables. The first alternative is an indicator which equals one if a person is hospitalised within two years after health insurance choice. The second alternative is an indicator which equals one if a person is hospitalised from period $t + 1$ onwards until the end of the sample period.

Finally, there appears to be some measurement error in income in the SOEP data, a point which has been raised by Hullegie and Klein (2010). To address this issue, I reestimate the models on the subsamples of individuals with incomes of 5,000 and 10,000 Euro, respectively, above the *compulsory insurance threshold*, in order to guarantee that all individuals in the sample can actually choose to opt out of public insurance.

6 Conclusion

There are a number of benefits associated to the German public health insurance with opt-out scheme. These include the socialisation of risks in public insurance vis-à-vis a system relying only on private health insurance, public insurance being an insurance of last resort to the seriously sick and the poor, a reduced number of insureds having double coverage compared to a system with a National Health Service such as the UK, and an increased competition for clients, which creates strong incentives for public health insurance providers to increase efficiency and quality. On the other hand, there are two major drawbacks. First, the difference in pricing between public and private insurance allows private insurers to attract individuals which are observably better risks. Second, private insurers may also attract individuals with better unobservable health, who look for contracts with higher degrees of cost sharing than the public plan.

The model presented in this paper shows that the public plan is not necessarily adversely selected under public health insurance with opt-out. However, the evidence for Germany suggests that private insurers have a competitive advantage over public health insurance providers under current legislation. This imbalance affects adversely the low-income earners, who are mandatorily insured in public insurance and have to bear the increase in contributions due to the outflow of good risks into private insurance. To conclude this paper, I briefly discuss two policies which aim at levelling the playing field between the providers of public and private health insurance in Germany.

A widely endorsed policy are risk-adjusted transfers between public and private insurers (Kifmann and Nell 2013, Grunow and Nuscheler 2014), which would eliminate the selection on observables. The implementation of a risk adjustment mechanism would be relatively inexpensive, given that Germany could build on an existing mechanism. The existing risk adjustment mechanism compensates public health insurance providers for differences in their risk pool and has been found to be generally effective (Nuscheler and Knaus 2005). However, there is some evidence that insurers select risks based on geography, which is not included in the risk adjustment formula (Bauhoff 2012).

In 2007, Germany introduced a reform that allows public health insurance providers to offer contracts with varying degrees of cost sharing (so-called *choice policies*). Theoretically, this reform should mitigate the selection on unobservables, but, to the best of my knowledge, there is no evidence on how this regulatory change has affected selection. Assessing the impact of the 2007 health insurance reform could prove to be an interesting topic for future research.

References

Akerlof, G. A. (1970) The market for "lemons": quality uncertainty and the market mechanism, *Quarterly Journal of Economics*, **84**, 488–500.

Andersen, H. H., Nübling, M., Mühlbacher, A., Schupp, J. and Wagner, G. G. (2007) Computation of standard values for physical and mental health scale scores using the SOEP version of SF12v2, *Journal of Applied Social Science Studies*, **127**, 171–182.

Bauhoff, S. (2012) Do health plans risk-select? An audit study on Germany's Social Health Insurance, *Journal of Public Economics*, **96**, 750–759.

Brown, J., Duggan, M., Kuziemko, I. and Woolston, W. (2014) How does risk selection respond to risk adjustment? New evidence from the Medicare Advantage program, *American Economic Review*, **104**, 3335–3364.

Buchmueller, T. C., Fiebig, D. G., Jones, G. and Savage, E. (2013) Preference heterogeneity and selection in private health insurance: The case of Australia, *Journal of Health Economics*, **32**, 757–767.

Bünnings, C. and Tauchmann, H. (2014) Who opts out of the statutory health insurance? A discrete time hazard model for Germany, *Health Economics*, pp. n/a–n/a.

Cao, Z. and McGuire, T. G. (2003) Service-level selection by HMOs in Medicare, *Journal of Health Economics*, **22**, 915–931.

Chiappori, P.-A. and Salanié, B. (2000) Testing for asymmetric information in insurance markets, *Journal of Political Economy*, **108**, 56–78.

Cohen, A. and Siegelman, P. (2010) Testing for adverse selection in insurance markets, *Journal of Risk and Insurance*, **77**, 39–84.

Colombo, F. and Tapay, N. (2004) Private health insurance in OECD countries: the benefits and costs for individuals and health systems, OECD Health Working Paper 15.

Dardanoni, V. and Li Donni, P. (2012) Incentive and selection effects of Medigap insurance on inpatient care, *Journal of Health Economics*, **31**, 457–470.

de la Mata, D., Machado, M. P. and Valdés, M. N. (2014) Hide and seek: Asymmetric information in the Chilean private health insurance market, unpublished manuscript.

Dohmen, T., Falk, A., Huffman, D., Sunde, U., Schupp, J. and Wagner, G. G. (2011) Individual risk attitudes: measurement, determinants, and behavioral consequences, *Journal of the European Economic Association*, **9**, 522–550.

Einav, L., Finkelstein, A. and Levin, J. (2010) Beyond testing: empirical models of insurance markets, *Annual Review of Economics*, **2**, 311–336.

Finkelstein, A. and McGarry, K. (2006) Multiple dimensions of private information: evidence from the long-term care insurance market, *American Economic Review*, **96**, 938–958.

Finkelstein, A. and Poterba, J. (2014) Testing for asymmetric information using "unused observables" in insurance markets: evidence from the U.K. annuity market, *Journal of Risk and Insurance*, **81**, 709–734.

Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J. P., Allen, H., Baicker, K. and Oregon Health Study Group (2012) The Oregon health insurance experiment: evidence from the first year, *Quarterly Journal of Economics*, **127**, 1057–1106.

Geil, P., Million, A., Rotte, R. and Zimmermann, K. F. (1997) Economic incentives and hospitalization in Germany, *Journal of Applied Econometrics*, **12**, 295–311.

Grunow, M. and Nuscheler, R. (2014) Public and private health insurance in Germany: the ignored risk selection problem, *Health Economics*, **23**, 670–687.

Hullegie, P. and Klein, T. J. (2010) The effect of private health insurance on medical care utilization and self-assessed health in Germany, *Health Economics*, **19**, 1048–1062.

Kifmann, M. and Nell, M. (2013) Fairer Systemwettbewerb zwischen gesetzlicher und privater Krankenversicherung, Hamburg Center for Health Economics Research Paper.

Manning, W. G., Newhouse, J. P., Duan, N., Keeler, E. B., Leibowitz, A. and Marquis, M. S. (1987) Health insurance and the demand for medical care: Evidence from a randomized experiment, *American Economic Review*, **77**, 251–77.

Newhouse, J. P., Price, M., Huang, J., McWilliams, J. M. and Hsu, J. (2012) Steps to reduce favorable risk selection in Medicare Advantage largely succeeded, boding well for health insurance exchanges, *Health Affairs*, **31**, 2618–2628.

Nuscheler, R. and Knaus, T. (2005) Risk selection in the German public health insurance system, *Health Economics*, **14**, 1253–1271.

Olivella, P. and Vera-Hernández, M. (2013) Testing for asymmetric information in private health insurance, *Economic Journal*, **123**, 96–130.

Riley, J. G. (1985) Competition with hidden knowledge, *Journal of Political Economy*, **93**, 958–976.

Riphahn, R. T., Wambach, A. and Million, A. (2003) Incentive effects in the demand for health care: a bivariate panel count data estimation, *Journal of Applied Econometrics*, **18**, 387–405.

Rothschild, M. and Stiglitz, J. (1976) Equilibrium in competitive insurance markets: an essay on the economics of imperfect information, *Quarterly Journal of Economics*, **90**, 629–649.

Sapelli, C. and Vial, B. (2003) Self-selection and moral hazard in Chilean health insurance, *Journal of Health Economics*, **22**, 459–476.

Schmitz, H. (2011) Direct evidence of risk aversion as a source of advantageous selection in health insurance, *Economics Letters*, **113**, 180–182.

Appendix

Variables whose coefficients are of the opposite (same) sign in the two columns of Table 5 and statistically significant in both columns constitute a source of adverse (advantageous) selection against (in favour of) the public sector based on observables.

Table 5. *Sources of Selection on Observables*

Dependent variable:	OPTOUT		HOSPITAL	
Disability status	-0.006	(0.014)	0.034**	(0.013)
Chronic illness	-0.012**	(0.005)	0.029***	(0.007)
Hospitalised last 5 years	-0.008	(0.010)	0.041*	(0.022)
Doctor visits last 3 years	-0.004**	(0.002)	0.006***	(0.001)
BMI >30	0.001	(0.007)	0.004	(0.009)
Sick leave >6 weeks last 3 years	0.015	(0.018)	0.018	(0.025)
Children	-0.011***	(0.003)	-0.009**	(0.003)
Married	-0.012**	(0.006)	0.025**	(0.010)
Female	0.006	(0.007)	-0.011	(0.010)
Female 20-39	-0.011	(0.010)	0.050***	(0.015)
Age 20-29 (omitted)	—	—	—	—
Age 30-39	-0.033***	(0.008)	0.017	(0.017)
Age 40-49	-0.066***	(0.009)	0.010	(0.018)
Age 50-60	-0.082***	(0.010)	0.026	(0.019)
Self-employed	-0.061	(0.043)	0.018	(0.063)
Self-employed · Log income	0.008*	(0.004)	-0.002	(0.006)
Year 1999 (omitted)	—	—	—	—
Year 2000	-0.008	(0.010)	0.012	(0.016)
Year 2001	0.001	(0.011)	0.007	(0.020)
Year 2002	0.031**	(0.016)	0.000	(0.023)
Year 2003	0.019	(0.015)	-0.005	(0.023)
Year 2004	0.017	(0.015)	0.017	(0.025)
Year 2005	0.005	(0.014)	-0.002	(0.023)
Year 2006	0.006	(0.015)	0.006	(0.024)
Year 2007	0.037**	(0.016)	-0.008	(0.023)
Year 2008	-0.003	(0.014)	0.014	(0.025)
Year 2009	0.027	(0.016)	0.010	(0.025)
Year 2010	0.005	(0.015)	0.001	(0.025)
N	9,017		9,017	

Notes: Table reports average marginal effects from a bivariate probit estimation of equations (2) and (3), where $c = 0$ is imposed. Means of OPTOUT and HOSPITAL are 0.041 and 0.085, respectively. Bootstrapped standard errors (1000 replications) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

The following three tables document that the empirical results are qualitatively unaffected by the possibly confounding factors which are discussed in Section 5.4 of the main text. Table 6 shows that there is robust evidence of adverse selection against the public plan. Table 7 demonstrates that the self-selection based on self-assessed health and risk aversion is robust. Table 8 illustrates that there is robust, albeit imprecisely estimated, evidence of asymmetric information.

Table 6. *Robustness Checks: Magnitude of Selection*

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: No controls						
$\hat{\rho}$	-0.234***	-0.084	-0.079**	-0.089***	-0.139***	-0.124**
LR test of $\rho = 0$ (p-value)	0.004	0.143	0.050	0.006	0.008	0.024
Panel B: Control for observables						
$\hat{\rho}$	-0.171*	-0.050	-0.041	-0.036	-0.089	-0.075
LR test of $\rho = 0$ (p-value)	0.066	0.431	0.349	0.323	0.123	0.207
Panel C: Control for observables and unused observables						
$\hat{\rho}$	-0.172*	-0.052	-0.035	-0.029	-0.085	-0.069
LR test of $\rho = 0$ (p-value)	0.070	0.419	0.423	0.419	0.148	0.254
N	1,874	5,842	7,621	8,427	6,635	5,660

Notes: Table reports the correlation of the residuals from bivariate probit estimations of equations (2) and (3). Panel A reports the correlations for the model with $b = c = 0$. Panel B reports the correlations for the model with $c = 0$. Panel C reports the correlations for the model with five *unused observables*. Column 1 shows estimates for the subsample of individuals who are in the first or second year of being eligible to choose private insurance. Column 2 shows estimates after controlling for two additional insuree characteristics: a mental component scale and a physical component scale. The dependent variable HOSPITAL in the bivariate probit model whose estimates are reported in column 3 is an indicator which equals one if a person is hospitalised within in two years after his opt out decision. The dependent variable HOSPITAL in the bivariate probit model whose estimates are reported in column 4 is an indicator which equals one if a person is hospitalised within the period from $t + 1$ onwards until the end of the sample period. Column 5 (6) shows estimates for the subsample of individuals who are self-employed and/or have incomes of 5,000 (10,000) Euros above the *compulsory insurance threshold*. * p<0.1, ** p<0.05, *** p<0.01.

Table 7. *Robustness Checks: Sources of Selection*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: OPTOUT						
Self-assessed health	0.022**	0.012**	0.011***	0.010***	0.011***	0.011***
	(0.010)	(0.005)	(0.003)	(0.003)	(0.004)	(0.004)
Risk aversion	-0.004	-0.003**	-0.004***	-0.003**	-0.004**	-0.004**
	(0.004)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)
Smoking	0.051***	0.014**	0.017***	0.014***	0.017***	0.016**
	(0.014)	(0.006)	(0.005)	(0.005)	(0.006)	(0.007)
Weekly hours of work	-0.000	0.001	-0.000	0.000	0.000	0.001
	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Log income	-0.030***	-0.024***	-0.021***	-0.021***	-0.025***	-0.027***
	(0.012)	(0.005)	(0.005)	(0.004)	(0.005)	(0.005)
Dependent variable: HOSPITAL						
Self-assessed health	-0.008	-0.007	-0.027***	-0.027***	-0.020***	-0.021***
	(0.010)	(0.007)	(0.007)	(0.010)	(0.005)	(0.006)
Risk aversion	-0.004	-0.007***	-0.008**	-0.007	-0.005**	-0.006**
	(0.004)	(0.002)	(0.004)	(0.006)	(0.002)	(0.002)
Smoking	-0.007	-0.008	0.001	-0.019	-0.006	-0.008
	(0.015)	(0.009)	(0.012)	(0.021)	(0.009)	(0.009)
Weekly hours of work	-0.003	0.001	0.004*	0.002	0.004**	0.003*
	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.002)
Log income	0.003	0.002	-0.006	0.002	0.000	-0.001
	(0.012)	(0.008)	(0.010)	(0.018)	(0.008)	(0.008)
N	1,874	5,842	7,621	8,427	6,635	5,660

Notes: Table reports average marginal effects of the *unused observables* from bivariate probit estimations of equations (2) and (3). Column 1 shows estimates for the subsample of individuals who are in the first or second year of being eligible to choose private insurance. Column 2 shows estimates after controlling for two additional insuree characteristics: a mental component scale and a physical component scale. The dependent variable HOSPITAL in the bivariate probit model whose estimates are reported in column 3 is an indicator which equals one if a person is hospitalised within in two years after his opt out decision. The dependent variable HOSPITAL in the bivariate probit model whose estimates are reported in column 4 is an indicator which equals one if a person is hospitalised within the period from $t + 1$ onwards until the end of the sample period. Column 5 (6) shows estimates for the subsample of individuals who are self-employed and/or have incomes of 5,000 (10,000) Euros above the *compulsory insurance threshold*. Means of OPTOUT and HOSPITAL vary across columns. Bootstrapped standard errors (1000 replications) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

Table 8. *Robustness Checks: Testing for Asymmetric Information*

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable: OPTOUT						
SAH=1	0.012	-0.015	0.006	0.004	0.013	0.013
	(0.343)	(0.174)	(0.126)	(0.118)	(0.142)	(0.147)
SAH=2	-0.015	-0.016	-0.007	-0.006	-0.008	-0.008
	(0.063)	(0.014)	(0.011)	(0.011)	(0.014)	(0.014)
SAH=3 (omitted)	—	—	—	—	—	—
SAH=4	0.004	0.013*	0.005	0.006	0.005	0.006
	(0.015)	(0.007)	(0.005)	(0.005)	(0.006)	(0.007)
SAH=5	0.028	0.022*	0.020***	0.018**	0.019**	0.020**
	(0.019)	(0.012)	(0.008)	(0.007)	(0.009)	(0.009)
Dependent variable: HOSPITAL						
SAH=1	-0.073	-0.014	-0.013	-0.070	0.039	0.016
	(0.329)	(0.043)	(0.044)	(0.054)	(0.028)	(0.031)
SAH=2	-0.000	0.014	0.036**	0.043*	0.022*	0.012
	(0.022)	(0.014)	(0.017)	(0.023)	(0.012)	(0.014)
SAH=3 (omitted)	—	—	—	—	—	—
SAH=4	-0.018	-0.023**	-0.036***	-0.033**	-0.027***	-0.031***
	(0.014)	(0.010)	(0.011)	(0.016)	(0.008)	(0.009)
SAH=5	-0.037*	0.014	-0.044**	-0.040	-0.026*	-0.036**
	(0.022)	(0.016)	(0.018)	(0.026)	(0.014)	(0.015)
N	2,295	6,108	8,272	9,154	7,207	6,139

Notes: Table reports average marginal effects of the five categories of self-assessed health from bivariate probit estimations of equations (2) and (3). 1=bad, 2=poor, 3=satisfactory, 4=good, 5=excellent. Column 1 shows estimates for the subsample of individuals who are in the first or second year of being eligible to choose private insurance. Column 2 shows estimates after controlling for two additional insuree characteristics: a mental component scale and a physical component scale. The dependent variable HOSPITAL in the bivariate probit model whose estimates are reported in column 3 is an indicator which equals one if a person is hospitalised within in two years after his opt out decision. The dependent variable HOSPITAL in the bivariate probit model whose estimates are reported in column 4 is an indicator which equals one if a person is hospitalised within the period from $t + 1$ onwards until the end of the sample period. Column 5 (6) shows estimates for the subsample of individuals who are self-employed and/or have incomes of 5,000 (10,000) Euros above the *compulsory insurance threshold*. Means of OPTOUT and HOSPITAL vary across columns. Bootstrapped standard errors (1000 replications) in parentheses. * p<0.1, ** p<0.05, *** p<0.01.