I Z A

# The Effect of Private Health Insurance on Medical Care Utilization and Self-Assessed Health in Germany

Patrick Hullegie
Tobias J. Klein

June 2010

# The Effect of Private Health Insurance on Medical Care Utilization and Self-Assessed Health in Germany

**Patrick Hullegie**
*Netspar, CentER, Tilburg University*

**Tobias J. Klein**
*Netspar, CentER, Tilburg University
and IZA*

# ABSTRACT

# The Effect of Private Health Insurance on Medical Care Utilization and Self-Assessed Health in Germany[*]

In Germany, employees are generally obliged to participate in the public health insurance system, where coverage is universal, co-payments and deductibles are moderate, and premia are based on income. However, they may buy private insurance instead if their income exceeds the compulsory insurance threshold. Here, premia are based on age and health, individuals may choose to what extent they are covered, and deductibles and co-payments are common. In this paper, we estimate the effect of private insurance coverage on the number of doctor visits, the number of nights spent in a hospital and self-assessed health. Variation in income around the compulsory insurance threshold provides a natural experiment that we exploit to control for selection into private insurance. We document that income is measured with error and suggest an approach to take this into account. We find negative effects of private insurance coverage on the number of doctor visits, no effects on the number of nights spent in a hospital, and positive effects on health.

JEL Classification:    I11, I12, C31

Keywords:    private health insurance, medical care utilization, selection into insurance, natural experiment, regression discontinuity design, measurement error

Corresponding author:

Tobias Klein
Tilburg University
Department of Econometrics and OR
PO Box 90153
5000 LE Tilburg
The Netherlands
E-mail: T.J.Klein@uvt.nl

# 1 Introduction

In Germany, employees are generally obliged to participate in the public health insurance system, where coverage is universal, co-payments and deductibles are moderate, and premia are based on income. However, they may buy private insurance instead if their income exceeds the so-called compulsory insurance threshold.[1] Here, premia are based on age and health, individuals may choose to what extent they are covered, and deductibles and co-payments are common.[2] These differences in the incentive structure may affect both health behavior and the demand for medical care. In particular, because of the higher co-payments and deductibles, privately insured patients have stronger incentives to invest in prevention to decrease the likelihood of occurrence of an illness. Therefore, even in case the treatment provided to privately and publicly insured patients is exactly the same, we would expect privately insured patients to be less inclined to demand medical services.

An important difference affecting the supply of services is that for the same treatment the compensation doctors receive for privately insured patients is, on average, 2.3 times as high as the compensation for publicly insured patients (Walendzik et al., 2008). Therefore, doctors have an incentive to treat privately insured patients first, and more intensely, possibly providing better treatment (Jürges, 2009). For example, waiting times for privately insured patients are lower on average (Lungen et al., 2008). This may in turn affect the demand for medical care.

The combination of demand and supply side incentives determines whether the amount of services consumed is higher or lower for privately insured individuals, and which effect insurance type has on health. Ultimately, it is an empirical question whether more or less services are consumed and how health depends on insurance status.

In this paper, we study the effect of being privately insured on the number of doctor visits,

---

[1]About 90 percent of the German population is insured in the public health insurance system. Most remaining individuals buy private insurance (Colombo & Tapay, 2004).

[2]In our data (the sample also used for Tables 2 and 3 below), 70 percent of the privately insured individuals who answered the respective question have insurance contracts that involve deductibles or co-payments.

the number of nights spent in a hospital and self-assessed health. We do not estimate the effects of specific insurance characteristics but interpret the results in light of the fact that deductibles and co-payments are common features of private insurance contracts. An unusual feature of the German health insurance system allows us to control for selection into private insurance: as soon as income in the last year exceeds the so-called compulsory insurance threshold, individuals become eligible to opt out of the public health insurance system and may buy private insurance instead. Random variation in income around this compulsory insurance threshold generates a natural experiment that allows us to conduct a regression discontinuity (RD) analysis and estimate the effect of private insurance for those individuals who buy private insurance once becoming eligible.[3] This local average treatment effect is interesting to policymakers considering to increase the compulsory insurance threshold because such an increase would force exactly those individuals for whom we estimate the effect to be publicly insured.

We use survey data from the German Socio Economic Panel (GSOEP) for our analysis because German administrative data, that contain accurate income measures, do not contain health related information. In the data, we find direct evidence for measurement error in income. Moreover, we find that there is a sizable number of individuals who, according to their reported income, are not eligible to buy private insurance but at the same time report to be privately insured. The methodological contribution in this paper is to model the measurement error in the so-called forcing variable, income in our case, within the RD framework. This then allows us to estimate the effects of interest.

Controlling for selection into private insurance we find a significant negative effects of being privately insured on the number of doctor visits for those individuals who visit the doctor at least once in a three month period. At the same time, we find no significant effects on the number of

---

[3]The RD approach has been suggested by Thistlethwaite and Campbell (1960) and has recently been developed by Hahn et al. (2001). They show that under relatively mild assumptions the RD method can be interpreted as a local randomized experiment. This gives the results a strong internal validity. However, in general, a drawback is that the effect is only estimated for a small subset of the population of interest/the population that a social planner is concerned with. See also Imbens and Lemieux (2008), Lee and Lemieux (2009) and Van der Klaauw (2009) for recent discussions. Our setup is the same as in Battistin et al. (2009).

nights spent in a hospital, which can arguably be influenced less by the individual, and positive effects on self-assessed health. This suggests that privately insured patients receive better or more intense treatment each time they see a doctor, or that they invest more in prevention.

The remainder of this paper is organized as follows. Section 2 and 3 discuss related results and the institutional details, respectively. In Section 4, we provide information on the data and document that there is measurement error in income. Section 5 discusses the econometric approach, emphasizing our approach to modeling measurement error. Results are presented in Section 6, and a sensitivity analysis is performed in Section 7. Finally, Section 8 concludes.

## 2   Related Literature

The empirical literature on demand for health services dates back at least to the 1970s when the RAND Health Insurance Experiment (HIE) was conducted. One important finding is that the use of medical services responds negatively to changes in cost sharing, with a stronger effect for outpatient care than for inpatient care (Newhouse, 1974; Manning et al., 1987).

There are at least four studies for Germany that relate demand for medical services to insurance type. They all use the GSOEP data. Geil et al. (1997) estimate a count data model for hospital visits on data from 1984-1989, 1992, and 1994. They find no relationship between insurance coverage and the hospitalization decision. Riphahn et al. (2003) estimate a bivariate count data model for physician and hospital visits. They use data from 1984 through 1995 and find that neither hospital nights nor doctor visits depend on the insurance type of the individual. Pohlmeier and Ulrich (1995) and Jürges (2009) both estimate a negative binomial hurdle model. Pohlmeier and Ulrich (1995) use data from 1985 and find that privately insured individuals are less likely to contact a general practitioner but the number of visits once they do so is not significantly different from the one for publicly insured patients. Jürges (2009) uses data from 2002 and finds that privately insured individuals are less likely to visit a doctor at all, but given that

they do the number of doctor visits is significantly larger than that of patients covered by public health insurance. All four papers have in common that they do not control for selection into private insurance.[4]

# 3 Institutional details

In Germany, about 90% of the population is publicly insured (Colombo & Tapay, 2004). Buying public insurance is mandatory for dependent employees as long as their income does not exceed the so-called compulsory insurance threshold. The public insurance premium equals a certain percentage (nowadays about 15 percent that are equally shared between the employer and the employee) of gross income up to the so-called contribution ceiling, and equal to it thereafter.[5]

Table 1 shows the contribution ceilings and the compulsory insurance thresholds by the year in which the income was earned. To see how the system works consider an individual whose income, including all extra payments, in 2000 was 40,000 Euros. Then, he is eligible to buy private insurance in 2001 because his income exceeded 39,574 Euros, the compulsory insurance threshold. If his income stays the same or decreases in 2001, then he will have to join the public insurance system again in 2002 because the compulsory insurance threshold is 40,034 Euros for income earned in 2001. He can apply for an exemption if he loses eligibility *solely* due to the increase in the compulsory insurance threshold, i.e. if his income in 2001 is at least 39,574 Euros. This applies to very few individuals in our data, about a tenth of a percent of all individual-year observations, and we therefore abstract from this exemption in the remainder.[6]

---

[4]The first two papers allow for random effects. Until recently both the theoretical and the empirical literature on informational asymmetries focused on *adverse* selection and moral hazard (Akerlof, 1970; Rothschild & Stiglitz, 1976; Arrow, 1963). However, Finkelstein and McGarry (2006) and Fang et al. (2008) point out that there might be *advantageous* selection instead. Their explanation is that "good risks" select into insurance because they are more risk averse and therefore value insurance more than "bad risks" do.

[5]See Jürges (2009) and the references therein for more details on this and the following discussion.

[6]This is because income typically increases faster than the compulsory insurance threshold. Sozialgesetzbuch V §8 also defines two more situations in which exemptions are granted, namely a temporary reduction in working hours and a combination of part-time employment and paternity leave. We abstract also from these possibilities because they are not widely used. If, to the contrary, the number of individuals who would be granted an exemption was

Table 1: Contribution ceiling and compulsory insurance threshold

| Year | Contribution ceiling | Compulsory insurance threshold | Mean income |
|------|------|------|------|
| 1994 | 34,968 | 34,968 | 24,633 |
| 1995 | 35,892 | 35,892 | 25,126 |
| 1996 | 36,816 | 36,816 | 25,905 |
| 1997 | 37,728 | 37,728 | 26,423 |
| 1998 | 38,652 | 38,652 | 26,660 |
| 1999 | 39,108 | 39,108 | 27,060 |
| 2000 | 39,574 | 39,574 | 27,358 |
| 2001 | 40,034 | 40,034 | 27,741 |
| 2002 | 40,500 | 40,500 | 28,231 |
| 2003 | 41,400 | 45,900 | 28,626 |
| 2004 | 41,850 | 46,350 | 28,938 |
| 2005 | 42,300 | 46,800 | 29,060 |

Reported for West Germany by year in which the income was earned. Amounts are nominal amounts per year and in Euros. The contribution ceilings and the compulsory insurance thresholds are based on Sozialgesetzbuch V and own calculations. Mean income is taken from Sozialgesetzbuch VI, Anlage 1.

It is in general difficult to get back into the public health insurance system. Once an individual has bought private health insurance he can only get back into the public system when he becomes unemployed (provided that he is younger than 55) or when his income falls below the compulsory insurance threshold (Colombo & Tapay, 2004).

Due to a reform the compulsory insurance threshold increased substantially for income earned in 2003 and later. A special rule applied to individuals who actually bought private insurance in 2003, but who were not eligible for this anymore according to the new thresholds. They could still buy private insurance provided that their income is at least equal to the contribution ceiling, which increased only moderately.[7]

---

large, say 1 percent of all individuals, then we would overestimate the discontinuity in the probability to be privately insured at the compulsory insurance threshold by 1 percentage point because we assume that this probability is zero for individuals earning less than the compulsory insurance threshold, and we would therefore underestimate the local average treatment effects. It follows from equation (1) below that we would overestimate the effect by about 5 percent if the discontinuity in the denominator was, e.g., 19 instead of 20 percent.

[7]We excluded these individuals from the empirical analysis.

Contributions for private health insurance are mainly based on health and age, so buying private insurance is especially attractive for young individuals. As a consequence of this, and because of the fact that private insurers are allowed to reject individuals, the risk pool of the private insurers is much better than in the public system.

Coverage is universal in the public system. Deductibles and co-payments are limited. Privately insured individuals can buy better care, e.g. treatment by the head doctor in a hospital or a single room in a hospital, but this comes at a higher price. Deductibles and co-payments are much more common, and many insurers offer a rebate if an individual did not use medical services in the past calendar year. Unfortunately, specific characteristics of private insurance are not recorded in our data.

At this point it is worth noticing that there is a feature called family insurance in the German public health insurance system. A spouse is automatically insured if an individual is insured. For this it is mandatory that the spouse is not full time self-employed and that the spouse does not earn more than a rather low specified amount. If a married man is working then this system generates incentives against working for his wife because then she would have to pay contributions which amount to about 7.5 percent of her gross wage (the employer matches this and pays about the same amount to the system). The family insurance feature does not exist for private health insurance and therefore, individual insurance has to be purchased for each family member.

As already pointed out before insurance status has important consequences for the compensation of doctors. For a given treatment the compensation doctors receive for privately insured patients is, on average, 2.3 times as high as the compensation for publicly insured patients (Walendzik et al., 2008). Furthermore, there is indirect evidence that doctors face strong time constraints when treating patients. The consultation length for the average (publicly insured) individual is very low in Germany.[8] Deveugele et al. (2002, Table 4) compare the average consultation length for general practitioners in six countries and find that with 7.6 minutes it is lowest

---

[8]Recall that about 90 percent of the individuals are publicly insured. See footnote 1.

in Germany. It is highest in Switzerland, where it is equal to 15.6 minutes. Together with the differences in the compensation this suggests that doctors dedicate more time to privately insured patients.

# 4 Data

The GSOEP we use in this study contain information at the individual level on medical care utilization, self-assessed health, and background variables. We analyze data from West Germany for the period from 1995 to 2006.[9]

Our sample is constructed such that eligibility to opt out of the public insurance system is exclusively determined by income. Unemployed individuals who receive unemployment benefits are required to be in the public health insurance system. For them there is no way to opt out and therefore they are excluded. Self-employed, civil servants, soldiers, teachers in private schools and students are not required to be in the public system, even if their income is below the compulsory insurance threshold. Hence, eligibility does not depend on income and therefore they are excluded from the sample as well. Retired individuals, who receive a public pension, are required to have public health insurance. They may opt out if insurance was not mandatory in at least five years after the age of 55 and most of the time before that. Hence eligibility is only weakly related to income and therefore they are excluded. Individuals of age 55 and older are excluded for two reasons. First, because for them various ways to opt for (early) retirement exist. Second, because for them it is difficult to get back into the public health insurance system. Individuals under the age of 25 are excluded because a large fraction of them is covered by their parents' insurance.

---

[9]We do not use data before 1995 because the question on the number of doctor visits was phrased differently. We use data only up to 2006 because from 2007 onwards individuals had to earn more than the compulsory insurance threshold in three consecutive years in order to be eligible to buy private insurance. East German individuals have been excluded because it turned out that for them, even when we control for measurement error in income, there is no jump in the probability to be privately insured when income is equal to the compulsory insurance threshold.

To summarize, our study population consists of West German individuals, aged 25 to 55, with a regular employment contract for whom eligibility to opt out of the public health insurance system is exclusively determined by income.

Table 2 contains descriptive statistics for the variables we use in the analysis. The first set of rows contains the outcome variables.[10] Eligible (to buy private insurance) and privately insured individuals visit the doctor slightly less often, and report to be in slightly better health. They report to be less likely to stay in a hospital and to spend less nights in a hospital on average.

The second set of rows contains summary statistics for individual characteristics. Gross income is, by construction, on average higher for eligibles. In light of this it is not surprising that it is higher for privately insured (because only those with high enough incomes are eligible to buy private insurance). The remaining rows are informative about selection into private insurance. Given the characteristics of public and private insurance it is relatively more attractive to buy private insurance for individuals who are not married. This is because spouses whose income is relatively low are automatically covered by the insurance of the individual. This is reflected by the fact that privately insured individuals are less likely to be married. They are older and better educated.

One key variable in our analysis is gross yearly income. This is not reported by the GSOEP respondents but constructed from their reports on their average gross monthly income in the previous year and their reports on supplementary income such as 13th month salary, 14th month salary, Christmas bonus, vacation pay, profit share, premia, and bonuses. Using self-reported income and Table 1 we can compute the eligibility status for every individual.

Table 3 shows that there is a sizable number of individuals, 577, who, according to their reported income, are not eligible to buy private insurance, but at the same time report to have done so. These 577 individuals constitute 20 percent of the individuals with private health insurance.

---

[10]For the self-assessed health question, 'bad' is re-coded as a 1, 'poor' as 2, and so on, up to 'very good' as 5. Hence, a positive association between health and private insurance would be reflected in a positive coefficient on an indicator for private insurance in an ordinary least squares regression.

Table 2: Descriptive statistics

|  | (1)<br>Ineligible | (2)<br>Eligible | (3)<br>Public<br>insurance | (4)<br>Private<br>insurance | (5)<br>Total |
|---|---|---|---|---|---|
| At least 1 doctor visit | 0.619 | 0.594 | 0.611 | 0.521 | 0.613 |
|  | - | - | - | - | - |
| Doctor visits given at least 1 visit | 3.304 | 2.920 | 3.243 | 2.904 | 3.223 |
|  | (4.212) | (3.365) | (4.094) | (3.287) | (4.052) |
| Doctor visits | 2.045 | 1.733 | 1.999 | 1.651 | 1.977 |
|  | (3.682) | (2.963) | (3.581) | (2.866) | (3.541) |
| At least 1 night in hospital | 0.079 | 0.065 | 0.078 | 0.057 | 0.076 |
|  | - | - | - | - | - |
| Nights in hospital | 0.862 | 0.655 | 0.833 | 0.572 | 0.613 |
|  | (5.027) | (4.033) | (4.844) | (4.603) | (4.830) |
| Self-assessed health | 3.585 | 3.696 | 3.596 | 3.799 | 3.609 |
|  | (0.850) | (0.790) | (0.841) | (0.777) | (0.838) |
| Gross income | 23,914.80 | 61,249.00 | 29,879.10 | 63,515.70 | 31,998.60 |
|  | (9,693.90) | (27,755.60) | (18,005.30) | (41,082.40) | (21,837.50) |
| Years of education | 11.533 | 13.971 | 11.881 | 14.785 | 12.065 |
|  | (2.228) | (2.929) | (2.471) | (2.945) | (2.601) |
| Married | 0.654 | 0.746 | 0.676 | 0.649 | 0.674 |
|  | - | - | - | - | - |
| Male | 0.500 | 0.848 | 0.562 | 0.784 | 0.576 |
|  | - | - | - | - | - |
| Age | 39.393 | 42.161 | 39.872 | 41.775 | 39.992 |
|  | (8.338) | (7.206) | (8.229) | (7.274) | (8.186) |
| $N$ | 35,822 | 9,900 | 42,841 | 2,881 | 45,722 |

Means and standard deviations (in parentheses). For binary variables only proportions are shown. Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income. $t$-tests show that for all variables the difference in the mean between ineligible and eligible individuals and between publicly insured and privately insured individuals significantly different from zero.

Misreporting insurance status or measurement error in income may both be valid explanations for this.[11]

We consider it to be more plausible that income is measured with error because income is a real number, and may thus be recalled with errors, whereas insurance status is more easily known

---

[11]There is an extensive literature on measurement error in income, see for example Bound et al. (2001) for a survey. In order to study the accuracy of survey reports, they are typically compared with either employers' or administrative records. Some studies find that survey reports are highly correlated with record values, while others find much lower correlations. The mean of survey reports is found to be close to the mean of the record values. That is, under- or over-reporting, if present, is found to be moderate on average.

Table 3: Eligibility and health insurance type

|  | Public insurance | Private insurance | N |
|---|---|---|---|
| Ineligible | 35,245 | 577 | 35,822 |
| Eligible | 7,596 | 2,304 | 9,900 |
| N | 42,841 | 2,881 | 45,722 |

Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income.

because it is typically either public or private insurance. Moreover, there is direct evidence for measurement error in income because the GSOEP questionnaire asks respondents twice about their monthly income in a given year.[12] In particular, respondents are asked about the income they received in the preceding month (without extra payments) and about their average monthly income in the previous year. This provides us with two measures of monthly income for the same year. If both income reports would be reported without any error, and if the within year variance in monthly income is low, then both measures should be close to one another. That is, the data points in a scatter plot should be close to the 45 degree line. Such a scatter plot is shown in Figure 1. The deviations from the 45 degree line are substantial. This strongly suggests that there is measurement error in income.[13]

# 5   Econometric approach

Let $(y_i(0), y_i(1))$ be the pair of potential outcomes for each member $i$ of the study population. In our case $y_i(0)$ denotes the health outcome individual $i$ would experience in case public health insurance was assigned to him and $y_i(1)$ denotes the health outcome individual $i$ would experi-

---

[12]This is not the case for the total yearly income that we use to determine eligibility. Yearly income includes extra payments such as holiday pay. The fact that it is yearly income and not monthly income that determines eligibility is the reason that we do not exploit the availability of two monthly income measures in the main analysis.

[13]This is robust to controlling for working hours and job changes by means of a regression. The $R^2$ in this regression is 0.846, meaning that 15.4 percent of the variation in the report of previous year's average income remains unexplained.

Figure 1: Joint distribution of the two income measures
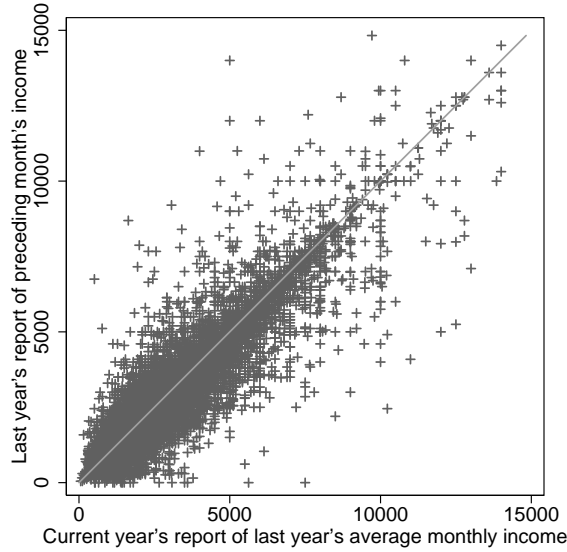


Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income. For this figure we use only income reports below $15,000$ Euros per month.

ence if private health insurance was assigned. That is, we consider private health insurance to be the "treatment."

An individual is eligible to buy private health insurance instead of public insurance if his income in the previous year exceeded the respective compulsory insurance threshold. That is, an individual is eligible when $z_i^* \geq 0$, where $z_i^*$ denotes the difference between income earned in the previous year and the corresponding compulsory insurance threshold. Buying private insurance is voluntary for eligible individuals so that some will buy it while others will not.

Following Hahn et al. (2001) we make three assumptions. First, we assume that the effect of private insurance is independent of $z_i^*$ around $z_i^* = 0$. This assumption is plausible as long as variation in income is independent of the effect of private insurance. It could be violated if individuals were to manipulate their income such that they become eligible to buy private insurance and the effect of private insurance was different for those individuals. To the best of our knowl-

edge there is no evidence for such manipulations in Germany. Second, we assume that the mean value of $y_i(0)$ conditional on $z_i^*$ is a continuous function of $z_i^*$ at $z_i^* = 0$. This assumption holds if the mean health outcome would be a smooth function in income around the compulsory insurance threshold once public insurance was exogenously assigned to everybody. This is highly plausible. Third, we assume that the decision to buy private insurance is monotone in eligibility. This is the monotonicity condition of Imbens and Angrist (1994). It holds by construction because ineligibles cannot buy private insurance. Under these assumptions the average treatment effect for those individuals that would buy private health insurance when becoming eligible is given by

$$\Delta^{LATE} \equiv \mathbb{E}\left(y_i(1) - y_i(0)\,|\,p_i = 1, z_i^* = 0\right) = \frac{\mathbb{E}\left(y_i|z_i^* = 0^+\right) - \mathbb{E}\left(y_i|z_i^* = 0^-\right)}{\mathbb{E}\left(p_i|z_i^* = 0^+\right)}, \tag{1}$$

where $y_i$ is the *observed* health outcome, $p_i$ is an indicator of private insurance, $\mathbb{E}\left(\cdot|z_i^* = 0^+\right) \equiv \lim_{\delta \downarrow 0} \mathbb{E}\left(\cdot|z_i^* = \delta\right)$, and $\mathbb{E}\left(\cdot|z_i^* = 0^-\right) \equiv \lim_{\delta \uparrow 0} \mathbb{E}\left(\cdot|z_i^* = \delta\right)$. This effect is of particular interest because it is directly related to the question what the effect of requiring all individuals with incomes slightly above the compulsory insurance threshold to buy public insurance would be.

Measurement error in income leads to misclassification of eligibility. Importantly, this misclassification is not independent of the true underlying income because if the true underlying income is below (above) the compulsory insurance threshold the classification error can only be that the individual is (not) eligible to buy private insurance. This precludes the use of an instrumental variables approach to estimating the unknown quantities in the numerator and denominator in equation (1).

The effect of the measurement error in income on estimates of these quantities is that no discontinuity in reported income is observed at the threshold (Battistin et al., 2009). In Figure 2, the dots are fractions of privately insured individuals, which we plot against the difference in income and the compulsory insurance threshold. The figure shows that these fractions are

13

Figure 2: Probability to be privately insured
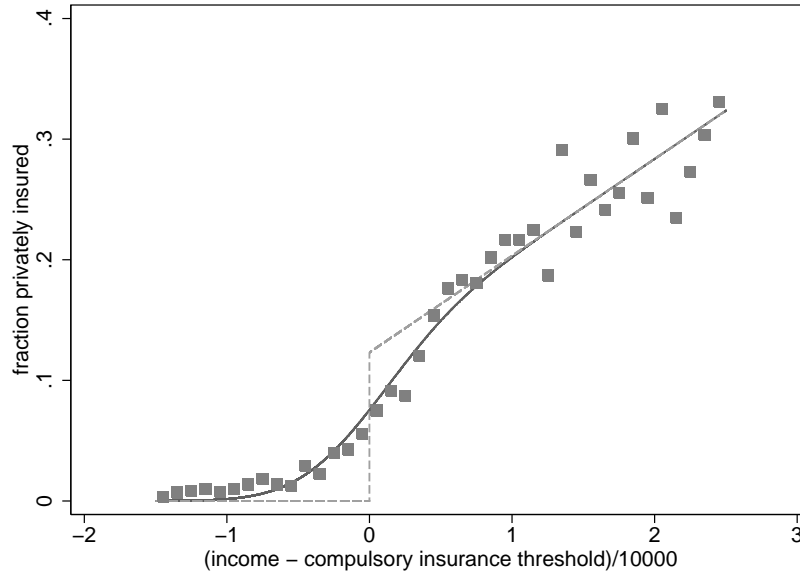
Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income and who earn between $15,000$ Euros less and $25,000$ Euros more than the compulsory insurance threshold ($N = 24,203$). The specification here imposes, for illustration, that the size of the discontinuity is the same in all years.

not zero if reported income is below the compulsory insurance threshold, i.e. if the value of the difference on the horizontal axis is negative, and that indeed there is no discontinuity in the fraction of privately insured at the threshold.

Towards estimating the local average treatment effect in the presence of measurement error we now develop an expression for the probability to be privately insured, which is equal to the conditional expectation of the indicator for being privately insured. Our approach is parametric and our main assumption is that $z_i^* = z_i - u_i$, where $u_i$ is normally distributed independent of $z_i$ and has mean zero and variance $\sigma_u^2$.[14] Furthermore, $u_i$ is assumed to be independent of private

---

[14]Notably, this is not classical measurement error. For classical measurement error we have $z_i = z_i^* + u_i$, which is equivalent, but we assume that $u_i$ is independent of $z_i^*$, and not of $z_i$. See Wooldridge (2002) for a discussion.

insurance status and the potential outcomes. We specify the (piecewise) linear probability model

$$
\mathbb{E}\left(p_i|z_i^*\right) = \begin{cases} 0 & \text{if } z_i^* < 0 \\ \alpha + \beta z_i^* & \text{if } z_i^* \geq 0. \end{cases}
$$

Recall that when true income is below the compulsory insurance threshold, i.e. when $z_i^* < 0$, then the probability of being privately insured is zero because ineligibles may not buy private insurance. Conversely, when true income exceeds the compulsory insurance threshold, i.e. when $z_i^* \geq 0$, individuals may buy private insurance.

We show in Appendix A that under these assumptions

$$
\mathbb{E}\left(p_i|z_i\right) = \Phi\left(\frac{z_i}{\sigma_u}\right) \cdot \left(\alpha + \beta z_i + \beta \sigma_u \frac{\phi\left(\frac{z_i}{\sigma_u}\right)}{\Phi\left(\frac{z_i}{\sigma_u}\right)}\right), \tag{2}
$$

where $\Phi\left(\cdot\right)$ is the standard normal cumulative distribution function and $\phi\left(\cdot\right)$ is the standard normal probability density function. Notably, this is the prediction for the relationship between the probability to be privately insured and the difference between reported income and the compulsory insurance threshold, $z_i$. The solid line in Figure 2 shows the estimated relationship for our data when we pool data across all years. The dots are sample fractions of privately insured. Comparing them to the solid line shows that the fit is reasonably good. Finally, the dashed line in this figure is the underlying relationship between the probability to be privately insured and the difference between actual (measured without error) yearly income and the compulsory insurance threshold, $z_i^*$.

A similar expression can be obtained for $\mathbb{E}\left(y_i|z_i\right)$. This involves specifying different linear functions to the left and right of the discontinuity,

$$
\mathbb{E}\left(y_i|z_i^*\right) = \begin{cases} \alpha_0 + \beta_0 z_i^* & \text{if } z_i^* < 0 \\ \alpha_1 + \beta_1 z_i^* & \text{if } z_i^* \geq 0, \end{cases}
$$

15

so that, under our assumptions,

$$\mathbb{E}\left(y_i|z_i\right) = \left(1 - \Phi\left(\frac{z_i}{\sigma_u}\right)\right)\left(\alpha_0 + \beta_0 z_i - \beta_0 \sigma_u \frac{\phi\left(\frac{z_i}{\sigma_u}\right)}{1 - \Phi\left(\frac{z_i}{\sigma_u}\right)}\right)$$
$$+ \Phi\left(\frac{z_i}{\sigma_u}\right)\left(\alpha_1 + \beta_1 z_i + \beta_1 \sigma_u \frac{\phi\left(\frac{z_i}{\sigma_u}\right)}{\Phi\left(\frac{z_i}{\sigma_u}\right)}\right). \tag{3}$$

The parameters for both $\mathbb{E}\left(p_i|z_i\right)$ and $\mathbb{E}\left(y_i|z_i\right)$ are jointly estimated using the feasible generalized nonlinear least squares estimator for nonlinear systems of equations. From these parameter estimates we then calculate the local average treatment effect. For this observe that $\alpha$, $\alpha_0$, and $\alpha_1$ are equal to $\mathbb{E}\left(p_i|z_i^* = 0^+\right)$, $\mathbb{E}\left(y_i|z_i^* = 0^-\right)$, and $\mathbb{E}\left(y_i|z_i^* = 0^+\right)$, respectively. Hence, it follows from equation (1) that the local average treatment effect is given by

$$\Delta^{LATE} = \frac{\alpha_1 - \alpha_0}{\alpha}. \tag{4}$$

# 6   Results

We jointly estimate the equation for the probability to be privately insured conditional on reported income, equation (2), and the equation for medical care utilization conditional on reported income, equation (3). Throughout, we allow the probability to be privately insured to have year specific jumps at the compulsory insurance threshold. This is reasonable since the compulsory insurance threshold changed over time (see Table 1). We impose that the local average treatment effect is the same in all years, i.e. we impose that $\Delta^{LATE}$, our parameter of main interest, is not only locally independent of $z_i^*$, but over a whole range of values. Then, it follows from equation (4) that we can replace $\alpha_1$ by $\alpha_0 + \Delta^{LATE} \cdot \alpha$. Notice that the size of both the numerator and the denominator in equation (1) is still allowed to vary across years, but we impose that the relative change in both is the same. Finally, we impose that expected health outcomes do not depend on

16

income, i.e. $\beta_0 = \beta_1 = 0$.[15]

We first estimate equation (2) alone. Results are reported in Table 4.[16] Coefficient estimates are marginal effects because the underlying model is a linear probability model. The probability is zero for negative $z_i^*$ and for positive $z_i^*$ it is linear in it. The results indicate that for all years there is a discontinuous jump in the probability to buy private insurance at $z_i^* = 0$. In 1995, the size of the jump is 9 percentage points, in 1996 it is 6 percentage points. From 1997 to 2001 the jump is about 10 percentage points. In 2002 and 2003, the jump increases slightly, and between 2004 and 2006 the jump substantially increases to approximately 18 percentage points. Supposedly, this is due to the increase in the compulsory insurance threshold for income earned in 2003, which affects the jump in the probability to be privately insured in 2004. For all individuals in our estimation sample the predicted value for the probability to be privately insured is between 0 and 1.

Table 5 presents the estimates of $\Delta^{LATE}$ for doctor visits in the past three months, the number of nights spent in a hospital, and self-assessed health. The respective baseline outcome is the average outcome for publicly insured individuals for whom true income is equal to the compulsory insurance threshold.

In specification (1), we use an indicator for at least one doctor visit as the dependent variable. This is a linear probability model since the expected outcome is a probability. 60.6 percent of the publicly insured individuals see a doctor at least once within a three month period. We find no significant effect of private insurance on this. In specification (2), we estimate the effect of private insurance on the number of doctor visits for those individuals who visit a doctor at least once. The baseline outcome is 3.329 doctor visits. The effect of private insurance on this is estimated to be negative and significant at the 1 percent level. The estimated magnitude of the

---

[15]We conducted several robustness checks. By jointly estimating more general models (involving non-zero slopes that were allowed to differ across years, e.g.) and our baseline specification we could check, respectively, whether treatment effect estimates were significantly different from the ones obtained using the baseline specification, and in general they were not.

[16]Estimates are very similar when we estimate equations (2) and (3) together.

Table 4: Probability to be privately insured

| | |
|---|---|
| (Gross income - threshold)/10000 | 0.075*** |
| | (0.005) |
| Discontinuity 1995 | 0.089*** |
| | (0.013) |
| Discontinuity 1996 | 0.064*** |
| | (0.013) |
| Discontinuity 1997 | 0.099** |
| | (0.041) |
| Discontinuity 1998 | 0.098*** |
| | (0.014) |
| Discontinuity 1999 | 0.107*** |
| | (0.013) |
| Discontinuity 2000 | 0.101*** |
| | (0.010) |
| Discontinuity 2001 | 0.109*** |
| | (0.011) |
| Discontinuity 2002 | 0.132*** |
| | (0.010) |
| Discontinuity 2003 | 0.114*** |
| | (0.010) |
| Discontinuity 2004 | 0.193*** |
| | (0.011) |
| Discontinuity 2005 | 0.191*** |
| | (0.012) |
| Discontinuity 2006 | 0.178*** |
| | (0.011) |
| $\sigma_u$ | 0.463*** |
| | (0.034) |
| $R^2$ | 0.184 |
| $N$ | 24,203 |

Standard errors are clustered at the individual level and shown in parentheses. $*, **, ***$ denote significance at the 10, 5, and 1 % level, respectively. Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income and who earn between $15,000$ Euros less and $25,000$ Euros more than the compulsory insurance threshold.

effect, however, seems to be too big. Specification (3) is for the number of doctor visits in the entire sample. This is a combination of the two effects we discussed above. The mean baseline

Table 5: Baseline specification

| | (1) At least one doctor visits | (2) Doctor visits for subsample | (3) Doctor visits | (4) At least one night in hospital | (5) Nights in hospital | (6) Self-assessed health |
|---|---|---|---|---|---|---|
| $\Delta^{LATE}$ | -0.079 | -3.746*** | -2.137*** | -0.063* | -1.084* | 0.449*** |
| | (0.076) | (0.945) | (0.546) | (0.035) | (0.572) | (0.160) |
| Baseline outcome | 0.606*** | 3.329*** | 2.013*** | 0.074*** | 0.783*** | 3.614*** |
| | (0.005) | (0.054) | (0.039) | (0.002) | (0.039) | (0.011) |
| $N$ | 24,203 | 14,579 | 24,203 | 24,203 | 24,203 | 24,203 |

Standard errors are clustered at the individual level and shown in parentheses. $*, **, ***$ denote significance at the 10, 5, and 1 % level, respectively. Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income and who earn between $15,000$ Euros less and $25,000$ Euros more than the compulsory insurance threshold.

outcome is estimated to be 2.013. The estimated effect is negative and significant at the 1 percent level, but again the magnitude of the point estimate is too big as it exceeds the baseline in terms of the magnitude.

Manning, Morris, and Newhouse (1981) argue that the decision to visit a doctor at all, the so-called contact decision, is made by the individual, whereas the number of visits is mainly determined by the doctor. However, it could also be that the patient and the doctor jointly determine the number of visits, or that fewer visits are needed for privately insured patients because they have invested in prevention. Furthermore, it could be that privately insured patients are treated more intensely so that less doctor visits are necessary. This is sensible because doctors are paid based on the number of treatments, not on the number of visits itself, and receive a higher compensation when they treat privately insured patients. They are time constrained and may thus focus on treating privately insured patients first (Lungen et al., 2008; Jürges, 2009), while spending relatively little time on publicly insured patients (Deveugele et al., 2002).

In specification (4) we use an indicator for at least one night spent in a hospital as the dependent variable. This is also a linear probability model. 7.4 percent of the publicly insured spend at least one night in a hospital. The results indicate that there is no significant effect of private insurance on this (at the 5 percent level). Specification (5) is for the number of nights

spent in a hospital and also here we find no significant effect of private insurance (also at the 5 percent level). These findings for hospital nights are in line with those of Geil et al. (1997) and Riphahn et al. (2003), and is intuitively plausible as the number of nights spent in a hospital can be influenced less by the individual than the the number of doctor visits is. Finally, we find that private insurance has a positive effect on health. Again, the size of the effect seems to be too big.

# 7   Sensitivity analysis

Generally, we do not have to control for covariates when performing an RD analysis unless the distribution of the covariates changes when we move from the left to the right of the discontinuity (Imbens & Lemieux, 2008). The measurement error in the forcing variable, however, prevents us from performing the usual tests. However, it is still feasible to perform the analysis incorporating a dependence of the baseline outcome and the probability to be privately insured on additional covariates. Table 6 reports the results. They are similar to our baseline results.

Some of the studies that use GSOEP data additionally condition on health when estimating the relationship between private insurance coverage and the health outcomes (Jürges, 2009, e.g.). For two reasons we consider it reasonable to condition on previous period's health instead of current health. First, one of the outcomes in this study is current period's health so that conditioning on current health is not sensible, at least for this outcome. Second, current period's health is likely to be endogenous. We condition on previous period's health by re-estimating the model for individuals who report in the previous period that their health is "satisfactory." Table 7 contains the results.[17]

As a further robustness check, it is interesting to estimate the difference in the respective expected outcome between individuals with reported values of $z_i$ slightly above zero and slightly below zero. Battistin et al. (2009) show that under the assumption that at least some individuals

---

[17]These results were obtained using a two-step procedure to achieve convergence. This procedure is described in the Online Appendix.

Table 6: Specification with covariates

|  | (1) At least one doctor visits | (2) Doctor visits for subsample | (3) Doctor visits | (4) At least one night in hospital | (5) Nights in hospital | (6) Self-assessed health |
|---|---|---|---|---|---|---|
| $\Delta^{LATE}$ | 0.104 | -2.480*** | -0.819 | -0.020 | -0.687 | 0.499*** |
|  | (0.083) | (0.875) | (0.533) | (0.039) | (0.621) | (0.171) |
| Baseline outcome | 0.598*** | 3.263*** | 1.964*** | 0.073*** | 0.770*** | 3.612*** |
|  | (0.005) | (0.053) | (0.039) | (0.002) | (0.040) | (0.011) |
| Years of education | 0.000 | -0.075*** | -0.048*** | -0.003*** | -0.048*** | 0.026*** |
|  | (0.002) | (0.016) | (0.012) | (0.001) | (0.011) | (0.004) |
| Married | 0.011 | -0.119 | -0.053 | 0.003 | -0.084 | 0.011 |
|  | (0.009) | (0.095) | (0.068) | (0.004) | (0.068) | (0.019) |
| Gender (male) | -0.188*** | -0.395*** | -0.861*** | -0.018*** | -0.126* | 0.032 |
|  | (0.009) | (0.098) | (0.078) | (0.005) | (0.069) | (0.022) |
| Age | 0.002*** | 0.036*** | 0.029*** | 0.001*** | 0.031*** | -0.025*** |
|  | (0.001) | (0.006) | (0.004) | (0.000) | (0.004) | (0.001) |
| $N$ | 23,830 | 14,360 | 23,830 | 23,830 | 23,830 | 23,830 |

Standard errors are clustered at the individual level and shown in parentheses. $*, **, ***$ denote significance at the 10, 5, and 1 % level, respectively. Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income and who earn between 15,000 Euros less and 25,000 Euros more than the compulsory insurance threshold.

Table 7: Baseline specification for subsample of individuals whose health in the previous period was "satisfactory"

|  | (1) At least one doctor visits | (2) Doctor visits for subsample | (3) Doctor visits | (4) At least one night in hospital | (5) Nights in hospital | (6) Self-assessed health |
|---|---|---|---|---|---|---|
| $\Delta^{LATE}$ | -0.257 | -4.421** | -4.284*** | -0.083 | -3.220*** | 0.053 |
|  | (0.174) | (1.869) | (1.447) | (0.093) | (1.113) | (0.234) |
| Baseline outcome | 0.682*** | 3.552*** | 2.467*** | 0.089*** | 0.927*** | 3.200*** |
|  | (0.010) | (0.107) | (0.089) | (0.006) | (0.092) | (0.014) |
| $N$ | 4,071 | 2,742 | 4,071 | 4,071 | 4,071 | 4.071 |

Standard errors are clustered at the individual level and shown in parentheses. $*, **, ***$ denote significance at the 10, 5, and 1 % level, respectively.

report their income accurately these estimates are lower bounds on the magnitude of the numerator in (1). Moreover, and more importantly, they also show that then the sign is equal to the sign of the local average treatment effect.

For this we perform separate local linear regressions to the left and to the right of 0, using

Table 8: Local linear regression estimates of the discontinuity at the threshold

| | ROT | half ROT | twice ROT |
|---|---|---|---|
| At least 1 doctor visit | 0.021 | 0.021 | 0.006 |
| | (0.027) | (0.034) | (0.018) |
| Doctor visits given at least 1 visit | -0.120 | -0.230 | -0.162 |
| | (0.214) | (0.318) | (0.147) |
| Doctor visits | -0.033 | -0.105 | -0.059 |
| | (0.154) | (0.235) | (0.101) |
| At least 1 night in hosp. | -0.002 | 0.007 | -0.001 |
| | (0.010) | (0.014) | (0.007) |
| Nights in hospital | 0.048 | 0.073 | 0.171 |
| | (0.140) | (0.232) | (0.115) |
| Self-assessed health | -0.002 | -0.025 | 0.017 |
| | (0.035) | (0.055) | (0.029) |

Estimates of the discontinuity that were obtained using Silverman's rule-of-thumb (ROT) bandwidth, as well as half and twice that bandwidth. We use a different bandwidth to the left and to the right of the discontinuity, respectively. Standard errors are bootstrapped, clustered at the individual level, and shown in parentheses. Sample consists of dependent employees for whom eligibility to opt out of the public health insurance system is exclusively determined by income and who earn between $15,000$ Euros less and $25,000$ Euros more than the compulsory insurance threshold.

Silverman's rule-of-thumb (ROT) bandwidth, of the respective outcome on $z_i$. Table 8 reports the results. In line with our baseline estimates it shows that private insurance has a big effect on the number of doctor visits.[18]

Results of further robustness checks are reported in the Online Appendix. In particular, we relax the assumption that expected outcomes are not related to income ($\beta_0 = \beta_1 = 0$), estimate the variance of the measurement error using the two available income measures, conduct the analysis for subsamples of individuals whose two income reports are close to one another, and obtain estimates using an estimation sample that is more narrowly defined in terms of income.

---

[18]Notably, the size of the discontinuity that is estimated here is not the local average treatment effect.

# 8  Conclusions

In this paper we estimate the effect of private health insurance on the number of doctor visits, the number of nights spent in a hospital, and self-assessed health in Germany. Variation in income around the compulsory insurance threshold generates a natural experiment which allows us to control for selection into private insurance and estimate respective average treatment effects for individuals who buy private insurance once they become eligible by earning enough.

We show that it is important to account for measurement error in income and suggest a way to do so. We find a significant negative effect of private insurance on the number of doctor visits for those individuals who see the doctor at least once. At the same time, we find no effect of private health insurance on the number of nights spent in a hospital, and a positive effect on self-assessed health. This suggests that private health insurance either has a positive effect on investment in prevention, because of the monetary incentives provided to the insured, or that privately insured patients receive more intense or better treatment each time they visit a doctor.

# Appendix A: Derivations

In this appendix we derive an expression for $\mathbb{E}(p_i|z_i^*) = \Pr(p_i = 1|z_i^*)$. Recall that $z_i^* = z_i - u_i$, where $u_i$ is normally distributed with mean 0 and variance $\sigma_u^2$, statistically independent of $z_i$, $p_i$ and the potential outcomes. For $z_i^* < 0$ we have that $\mathbb{E}(p_i|z_i^*) = 0$ by definition. For $z_i^* \geq 0$ we specify $\mathbb{E}(p_i|z_i^*)$ to be a linear function in $z_i^*$, a linear probability model. That is,

$$\mathbb{E}(p_i|z_i^*) = \begin{cases} 0 & \text{if } z_i^* < 0 \\ \alpha + \beta z_i^* & \text{if } z_i^* \geq 0. \end{cases}$$

By the law of total probability,

$$\mathbb{E}(p_i|z_i) = \Pr(z_i^* < 0|z_i) \cdot 0 + \Pr(z_i^* \geq 0|z_i) \cdot \mathbb{E}(p_i|z_i, z_i^* \geq 0).$$

The assumptions about the measurement error imply that this is equivalent to

$$\mathbb{E}(p_i|z_i) = \Pr(u_i \leq z_i) \cdot (\alpha + \beta \mathbb{E}(z_i - u_i|z_i, u_i \leq z_i)). \tag{5}$$

Recall that if $v$ is standard normally distributed then $\mathbb{E}(v|v < c) = -\phi(c)/\Phi(c)$, which is known as the inverse Mills ratio, where $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative distribution function and the probability density function, respectively. Using this equation (5) can be rewritten as

$$\mathbb{E}(p_i|z_i) = \Phi\left(\frac{z_i}{\sigma_u}\right) \cdot \left(\alpha + \beta z_i + \beta \sigma_u \frac{\phi\left(\frac{z_i}{\sigma_u}\right)}{\Phi\left(\frac{z_i}{\sigma_u}\right)}\right).$$

# References

Akerlof, G. (1970). The market for 'lemons': Quality uncertainty and the market mechanism. *Quarterly Journal of Economics*, *84*, 488-500.

Arrow, K. (1963). Uncertainty and the welfare economics of medical care. *American Economic Review*, *53*, 941-973.

Battistin, E., Brugiavini, A., Rettore, E., & Weber, G. (2009). The retirement consumption puzzle: Evidence from a regression discontinuity approach. *American Economic Review*, 2209 - 2226.

Bound, J., Brown, C., & Mathiowetz, N. (2001). Measurement error in survey data. In J. Heckman & E. Leamer (Eds.), *Handbook of Econometrics, volume 5* (p. 3705 - 3843). Elsevier Science.

Colombo, F., & Tapay, N. (2004). *Private health insurance in OECD countries: The benefits and costs for individuals and health systems.* OECD Health Working Paper No 15.

Deveugele, M., Derese, A., van den Brink-Muinen, A., Bensing, J., & De Maeseneer, J. (2002). Consultation length in general practice: Cross sectional study in six European countries. *British Medical Journal*, *325*, 472-477.

Fang, H., Keane, M., & 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 demensions of private information: Evidence from the long-term care insurance market. *American Economic Review*, *96*, 938 - 958.

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

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica*, *69*, 201 - 209.

Imbens, G., & Angrist, J. (1994). Identification and estimation of local average treatment effects. *Econometrica*, *62*, 467-475.

Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: A guide to pratice. *Journal of Econometrics*, *142*, 615 - 635.

Jürges, H. (2009). Health insurance status and physician behavior in Germany. In *Schmoller's Jahrbuch* (Vol. 129, p. 297 - 307).

Lee, D., & Lemieux, T. (2009). Regression discontinuity designs in economics. *Working Paper No. 14723 NBER*.

Lungen, M., Stollenwerk, B., Messner, P., Lauterbach, K. W., & Gerber, A. (2008). Waiting times for elective treatments according to insurance status: A randomized empirical study in Germany. *International Journal for Equity in Health*, *7*(1), 1-7.

Manning, W., Morris, C., & Newhouse, J. (1981). A two-part model of the demand for medical care: Preliminary results from the health insurance study. In J. van der Gaag & M. Perlman (Eds.), *Economics and health economics.* Amsterdam: North Holland.

Manning, W., Newhouse, J., Duan, N., Keeler, E., Leibowitz, A., & Marquis, M. (1987). Health insurance and the demand for medical care: Evidence from a randomized experiment. *American Economic Review*, *77*, 251 - 277.

Newhouse, J. (1974). A design for a health insurance experiment. *Inquiry*, *11*, 5 - 27.

Pohlmeier, W., & Ulrich, V. (1995). An econometric model of the two-part decisionmaking process in the demand for health care. *Journal of Human Resources*, *30*(2), 339-361.

Riphahn, R., Wambach, A., & 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., & Stiglitz, J. (1976). Equilibrium in competitive insurance markets: An essay in the economics of incomplete information. *Quarterly Journal of Economics*, *90*, 624-649.

Thistlethwaite, D., & Campbell, D. (1960). Regression discontinuity analysis: An alternative to the ex post facto experiment. *Journal of Educational Psychology*, *51*, 309 - 317.

Van der Klaauw, W. (2009). Regression discontinuity anlaysis: A survey of recent developments in economics. *Labour*, *22*, 219 - 245.

Walendzik, A., Gress, S., Manouguian, M., & Wasem, J. (2008). Vergütungsunterschiede im ärztlichen Bereich zwischen PKV und GKV auf Basis des standardisierten Leistungsniveaus der GKV und Modelle der Vergütungsangleichung. *Diskussionsbeitrag No. 165, University of Duisburg-Essen*.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data.* Cambridge, MA: MIT Press.