

Thönnies, Stefanie

Conference Paper

Do Deductibles reduce Moral Hazard in the German Statutory Health Insurance? - Empirical Evidence

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Health Economics 1, No. A19-V4

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Thönnies, Stefanie (2015) : Do Deductibles reduce Moral Hazard in the German Statutory Health Insurance? - Empirical Evidence, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Health Economics 1, No. A19-V4, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/112912>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Do Deductibles reduce Moral Hazard in the German Statutory Health Insurance? – Empirical Evidence

February 2015

Abstract

This paper estimates the effect of participating in the optional tariff “Premium Refund”, an implicit optional deductible tariff, on different measures of medical demand. Specifically, it analyzes whether participating in the tariff can reduce ex-post moral hazard. Therefore, I use panel data from a German company health insurance fund covering the years 2008 to 2012. In order to remove potential selection bias of healthy individuals selecting into the tariff, I match the group of participants with the group of nonparticipants by socioeconomic characteristics and baseline medical characteristics, amongst others. In addition, I combine matching with regression to make results more robust. I find that participating in the premium refund tariff significantly reduces the probability of visiting a general practitioner by 6 percentage points. However, the probability of visiting a doctor because of a trivial disease such as a common cold is not reduced. I conclude that there is evidence that participation in the tariff reduces medical demand. It remains unclear, though, whether this is due to a reduction in moral hazard or whether it is driven through some other channel.

1 Introduction

In the German statutory health insurance (SHI) system, one can observe a trend that has been stable for many years: Medical expenditures grow faster than the wage level does (Die Bundesregierung, 2010). Since contributions to the SHI are wage-dependent, as a consequence the gap between revenues and expenditures of sickness funds grows. This problem could be solved by increasing the contribution rates, but this in turn would mean higher labor costs and might lead to higher unemployment, and it is an additional burden for employees. Therefore, increasing contribution rates are not desired by the German government (Augurzky et al., 2006). There are several forces behind growing medical expenditures. Firstly, technological progress provides additional and often more expensive possibilities in treating patients. Secondly, demographic change leads to an ageing German population and to a higher demand of treatment. Finally, there are inefficiencies in the system, one of them being moral hazard (Schmitz, 2012). The latter means that the demand for insured benefits rises if not an individual but the insurance has to cover its costs (Breyer and Buchholz, 2009). This is either because individuals do not invest enough to prevent the event of damage (ex-ante moral hazard), or because individuals do not try to minimize costs in case of damage if these are taken over by the insurance company (ex-post moral hazard). While it is difficult for the insurance company to control whether individuals invest enough to prevent the event of damage (e.g. healthy food, no smoking, sports), there are devices how ex-post moral hazard can be reduced.

Generally, co-payment could be a device to reduce moral hazard. Economic theory predicts that co-payment may under some circumstances reduce the individual demand for health care services by increasing the price paid by the consumer at the time of consumption (Arrow, 1963; Pauly, 1968; cited after Kiil and Houlberg, 2014). The magnitude of the effect depends on the price elasticity of demand. Imposed on health care services for which the demand is price elastic, co-payment may be shown to reduce the demand. Indeed, there is empirical evidence for this, although with mixed results.

First of all, in the well-known RAND experiment it could be shown for the US that the number of doctor visits became smaller when the deductible increased (Manning et al., 1987). But there is also evidence - albeit non-experimental - for Germany: Pütz and Hagist (2006) show that an optional deductible is able to reduce the number of doctor visits as well as the probability to visit a specialist. Also expenditures for doctor visits, hospital stays and drug prescriptions are decreased. Schmitz (2012) analyzes whether more insurance coverage leads to more demand of medical services. He finds that effects are rather small and that mostly low users are affected, while high users, i.e. individuals with a lot of doctor visits, are not affected. Augurzky et al. (2006) as well as Schreyögg and Grabka (2010) study the effects of the 2004 introduced co-payment of 10 Euro that

had to be paid for every first doctor visit in each quarter. Both find that this type of deductible does not reduce the probability of visiting a doctor.

In most studies that find a reduction, it is not clear whether the reduction in demand is due to decreased moral hazard. In many cases, there is only evidence for a negative association between the amount of co-payment and the extent of demand for medical services. Anyway, compulsory co-payment is not applicable to the German SHI which is characterized by the solidarity principle. The latter means that all insured individuals have access to comprehensive coverage irrespective of how much each person has paid into the social insurance system. Thereby, there should be a reallocation between the rich and the poor, the sick and the healthy, as well as between singles and families with children. This reallocation would be disturbed because individuals who are sick, poor, or who have children would have to pay relatively more.

Since 2007 social health insurance companies have been allowed to offer the optional tariff "Premium refund"¹, which is basically an implicit co-payment. It was argued that it could motivate individuals to invest more in health and to avoid unnecessary doctor visits (Schulze Ehrling and Weber, 2007). In fact, it can be observed that participants cause lower medical expenditures compared to non-participants (cp. Chapter 6) but it is unclear whether this is due to selection. Therefore, I want to analyze whether the premium refund tariff is a device to reduce ex-post moral hazard, i.e. to avoid over-consumption of medical treatment. While treatment in hospital or drug prescriptions are primarily decided on by the doctor, the decision to visit a practitioner lies with the patient. This applies especially to general practitioners and to a lesser extent to specialists. Finally, one could imagine that participants have an incentive to avoid visiting a general practitioner in case of a common cold which can also be cured by staying at home and being at rest.

In this paper I estimate the causal effect of participating in the optional tariff "Premium Refund" in the year 2012 on the probability of visiting a general practitioner or a specialist as well as the probability of visiting a doctor due to a common cold. In addition, I estimate the effect on the sickness fund's medical expenditures for general practitioners and specialists although one has to keep in mind that in most cases the patient only can decide on visiting a doctor or not, but not necessarily on the type and extent of treatment. Using propensity score matching as well as a rich set of control variables, I want to eliminate bias which could be caused by self-selection. I use administrative data from a German company health insurance fund. Since in the year 2012 there were

¹This became possible through the law "Strengthening Competition Conditions within the Statutory Health Insurance System" (GKV-Wettbewerbsstärkungsgesetz). The general regulations have been modified in the context of the "Act for the Sustainable and Socially Balanced Financing of Statutory Health Insurance" (GKVFinanzierungsgesetz) in 2011 as well as the law "Gesetz zur Beseitigung sozialer Überforderung bei Beitragsschulden in der Krankenversicherung" (KVBeitrSchG) in 2013.

170,189 participants in the whole SHI system and until today the number of participants has further risen (GBE, 2014), this is a relevant analysis.

In the literature the premium refund tariff is evaluated to have negative consequences. Regorz (2007) analyzes the existing premium refund tariffs in the SHI in 2007. He concludes that they are not sustainable because the effect of keeping good risks that would otherwise have left the insurance is smaller than the deadweight effect. The latter means good risks that would have stayed with the insurance anyway take the premium without reducing their medical expenditures. However, Regorz does not use any data to verify his assumption but does only argue theoretically. In addition, Malin and Schmidt (1995) evaluate the premium refund tariff which was implemented from 1989 to 1994 as a pilot project with six health insurance funds in Germany. They arrive at the conclusion that participation in the optional tariff does not influence the utilization of medical services and that the tariff therefore cannot reduce expenditures of the sickness fund (cited after Schulze Ehring and Weber, 2007). Also Schulze Ehring and Weber (2007) state that it is discussed controversially whether optional tariffs can cause a reduction in medical demands. A survey among publically insured has shown that potential participants analyze offered optional tariffs very critically and that they only consider it if it redounds to their advantage (Paquet and Stein, 2007). In general, the public opinion on optional tariffs is rather negative. It has been feared that they weaken the solidarity system (Schulze Ehring and Weber, 2007).

This paper contributes to the existing literature in three parts. Firstly, to the best of my knowledge, it is the first paper that analyzes the premium refund tariff that was implemented in the 2007 reform using data. Secondly, since they do not stem from a survey, the data are free of recall bias. Thirdly, due to detailed information on the individuals' diagnoses, I can specifically analyze ex-post moral hazard.

The paper is organized as follows. Section 2 gives some background information on the legal regulations, Section 3 describes the model. In Section 4 and Section 5, the data and the identification strategy are presented. Afterwards descriptive statistics (Section 6) and results (Section 7) are shown. Section 8 discusses the results and concludes.

2 Background

The German health insurance system is characterized by the coexistence of a statutory and a private health insurance system. The vast majority (88.15 % in 2011) is publicly insured (GBE, 2015). Everybody is obligated to insure himself. However, not everyone can decide between the two systems. Only high earners with an income above the social security ceiling, self-employed, and civil servants can (but do not have to) choose a private

health insurance company. Within the system of the statutory health insurance, there were 132 sickness funds in 2014 (GKV head organization, 2015). Most of them can be chosen by all individuals, while some are only open for special groups (e.g. with respect to place of residence or employer). The sickness fund which is analyzed in this paper belongs to the former group. Each sickness fund has to accept an applicant, irrespective of his health status or income. It is characteristic of the statutory health insurance system that children and non-working spouses can be co-insured with the working spouse's sickness fund free of charge. The contribution to the statutory health insurance depends on the insurant's gross wage and is borne jointly by the employer and the employee. The employer pays a fixed percentage share of the employee's labor income. In contrast, the employee pays a potentially varying percentage share of his labor income. This means if the insurance fund raises the contribution rate, the additional pressure lies on the employee only. In contrast, in the private health insurance sector the contribution depends on the insurant's risk (i.e. age and health status, amongst others). Unlike private insurance funds that refund patients' expenses, benefits in the statutory health insurance are provided as benefits-in-kind.

While deductibles are common in the private health insurance system, they are rather new to the statutory health insurance system. After some pilot projects, e.g. in the 1990s (Malin and Schmidt, 1995) and in 2003 (Pütz and Hagist, 2006), they were introduced to the whole system on 01.04.2007 within the framework of optional tariffs. Optional tariffs are intended to increase competition within the statutory health insurance system as well as between public and private health insurance companies. There are seven different types of optional tariffs. Two types have to be offered by all statutory health insurance funds and five types can be offered voluntarily by the sickness funds. Here, the focus is on the optional tariff "Premium Refund" which is one of the latter. It can be offered to all members that have been insured with the sickness fund for at least three months. If a member and his co-insured family members do not cause expenditures within one calendar year, he receives a refund of 1/12 of his yearly insurance contribution at most. Otherwise, the refund is cut to zero. Since the insurance contribution depends on the income, the refund does so, too. However, some types of treatment do not cut the premium. These are treatment of children under 18 years, early diagnosis examinations, and prevention, amongst others. If a member chooses the premium refund tariff, he is bound to it as well as to the respective sickness fund for at least one year. However, there is an extraordinary right to terminate the membership under special circumstances.

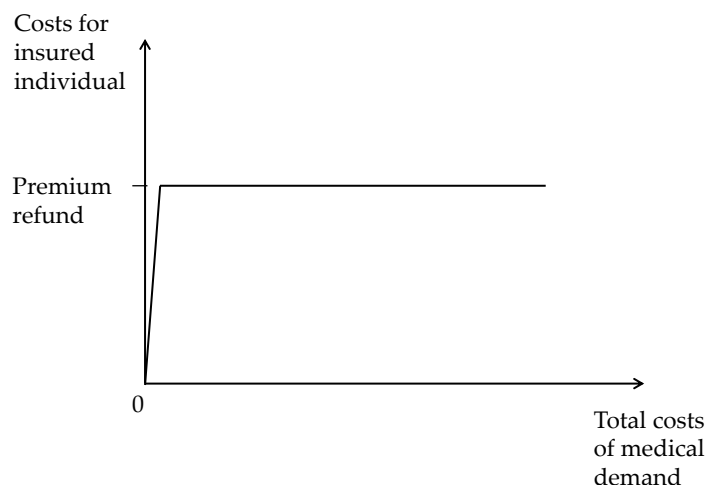
It has to be noticed that the sickness fund can further specify the tariff's regulations. Like many other company health insurance funds, also the analyzed one allows the insurant to enroll retrospectively, at the latest until the end of the respective year. This leads to the problem that new participants are not necessarily affected by the tariff. Instead,

one has to assume that a considerable share enrolls in the tariff by the end of the year if they discover they did not cause any insurance claims. I aim at removing this so-called lock-in effect by eliminating all new participants of the year 2012 that have not already participated in 2011 from the sample.

3 Model

The regulations say that the insured individual receives the premium refund only if he (and his co-insured family members) did not cause any expenses for the insurance company within one year. This means in case they make use of the health insurance, individuals implicitly face a deductible as high as the premium refund, irrespective of the amount of their costs. Afterwards, they face marginal costs equal to zero (cf. Figure 1).

Figure 1: Marginal costs



Therefore, what can be affected by participating in the tariff is the decision to visit a doctor for the first time in the respective year, or not to visit him. Clearly, this only holds for price-elastic demand (Jacobs et al., 2010). While there are severe sicknesses where the individual does not have a choice and demand is price-inelastic, or sicknesses where he is uncertain if he needs to see a doctor, there are also trivial sicknesses that clearly do not demand a doctor, e.g. a common cold. With such trivial sicknesses, demand is supposed to be price-elastic. For this reason, I analyze whether participants have a lower probability of visiting a doctor due to a common cold. Assuming that nearly everybody gets a common cold once in a while, everybody is affected by this sickness in a similar way. If someone visits a doctor due to a common cold this does not mean that he is sicker than someone with a common cold who does not visit a doctor. Instead, it only reveals his behavior. For trivial diseases, individuals mostly visit a general practitioner

whereas specialists are usually visited in the case of a more severe illness. Therefore, I also differentiate between the probability of visiting a general practitioner and that of visiting a specialist. In addition, I estimate the costs that incurred due to doctor visits. However, one has to be aware that the patient himself, because of the principal-agent relationship between doctor and patient, mostly cannot fully decide on the extent of the treatment, and is therefore not fully responsible for the amount of costs.

In principal, it could be a problem if the patient cannot pay the deductible and has therefore no chance to decide for or against the doctor visit (Jacobs et al., 2010). Then, he does not have access to important medical treatment and the disease can even aggravate. Here, this is not the case because the deductible is only paid implicitly, i.e. the individual does receive no premium and he does not have to pay extra money.

4 Data

The panel data cover the years 2008 to 2012 and result from billing processes of a German company health insurance fund. It covers yearly costs per insurance member, including his co-insured family members but excluding those under the age of 18. The yearly costs were standardized according to the number of persons counting to the specific family as well as to the number of days the family was insured with this fund in the respective year. The sample is limited to individuals that have been insured for at least 150 days in the year 2012 as well as 150 days in sum of the years 2008 to 2010. This is done in order to avoid an underreporting of costs. Furthermore, observations with missing values in any of the used covariates were dropped, just as participants of the year 2012 that did not already participate in the tariff in 2011.

The yearly total costs contain costs for hospitalization, visits at the doctor's, drugs, sickness payments, as well as other costs which include payments for rehabilitation, prevention courses, preventive examinations, home healthcare products, and administration costs, amongst others. Furthermore, the number of doctor visits, by general practitioners and by specialists, and the number of drug prescriptions can be observed. In addition, the number of days employees were absent from work due to illness is known. One problem in Germany is that in many employment contracts employees can be on sick-leave without attestation for the first two days. Hence the data will underreport the real number of days individuals were absent from work due to illness.

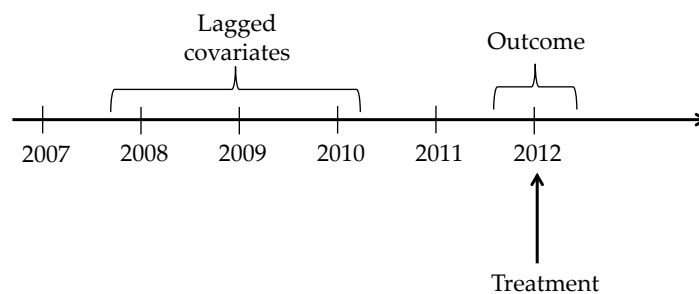
Moreover, some socioeconomic information on the insured person is available. The data provide information on the person's age, gender, insurance status, the number of co-insured family members, education, occupational status, as well as the federal state he lives in. In addition, it is known whether the person has participated in the bonus program

of the sickness fund during at least one of the years 2008 to 2012. One could argue that one should additionally control for having participated in the optional tariff in earlier years. However, since – due to the sample definition – all participants of the year 2012 already participated in 2011, this would destroy the common support. Therefore, it is not possible to control for this.

To condition on lagged outcomes and on some lagged covariates, I use the average of the years 2008 to 2010, if available. Averaging the three years is done in order to smooth unusual shocks that can be caused by an accident, for example. Observations of the year 2011 are not used as control variables, however, because individuals could anticipate their participation and choose to antedate medical demands. Furthermore, since usually individuals participate many years in a row, it is beneficial to control on observations of rather earlier years which increases the probability that they were not affected by having already participated then.

Covariates that cannot be affected by participation in the tariff are measured in 2012 (age, gender, insurance status, number of co-insured family members, education, occupational status, federal state). Other covariates are used as lags because they could be influenced by participation in the tariff. For these covariates, Figure 2 graphically shows from which years types of medical demand are used as covariates. Participation in the bonus program is an exception as it is measured in all years between 2008 and 2012. However, as the bonus program has been existent years before the optional tariff was implemented, it is unlikely that participation in the optional tariff would affect participation in the bonus program.

Figure 2: Timeline



5 Identification Strategy

I observe two groups. The treatment group consists of individuals that participate in the optional tariff, whereas the control group does not participate. I am interested in

the average treatment effect on the treated (ATT) which is the difference in demand for medical treatment between persons in the treatment group that have been treated on the one hand and persons in the treatment group had they not been treated on the other hand:

$$\tau_{ATT} = E[Y(1)|D = 1] - E[Y(0)|D = 1]$$

where $D = 1$ indicates that an individual belongs to the group that will choose the optional tariff. $Y(1)$ is the demand of individuals that actually choose the optional tariff, whereas $Y(0)$ is the demand of these individuals had they not chosen the optional tariff.

Naturally, the counterfactual $E[Y(0)|D = 1]$ is not known. What one finds, if one compares the observable outcomes between the two groups, is the ATT plus a selection bias. Presumably, individuals who choose to participate in the optional tariff differ from persons who do not, even if there was no treatment. Therefore, I use the large group of non-participants to find individuals that are similar to the participants in all relevant pre-treatment characteristics and employ matching estimation based on the propensity score. In deciding which variables to include and which functional form to be chosen for calculating the propensity score, I follow the recommendations by Imbens (2014)².

There are four conditions that have to be fulfilled for identification in matching. Firstly, the conditional independence assumption (CIA) says that conditional on a sufficient set of covariates, the potential outcomes are independent of treatment for given values of the confounding variables. Participation in the tariff can be endogenous if there are differences between individuals with respect to their health status and with respect to their risk attitudes towards health (Schmitz, 2012). Individuals that expect high expenses in the future might not choose a deductible, presumably also not if it is implicit in the form of a premium (non-)refund. Since the health status is associated with both the demand for health services and the choice of participating in the tariff, it is important to control for it. I use lagged values of medical costs, of the number of doctor visits, days of sickness absence, and drug prescriptions as a proxy for the real health status previous to the year 2012. Furthermore, risk-averse individuals might prefer full insurance coverage and therefore be less likely to participate in the tariff. Simultaneously, given a certain health status, they might demand more medical treatment, or they might show higher

²Imbens proposes to pre-select variables that are assessed being important according to economic theory. In addition, a set of variables is selected where it is not clear whether they should be included in the model. Each of these variables is tested by comparing a logistic regression of the treatment dummy on pre-selected variables with a logistic regression of the treatment dummy on pre-selected variables and one of the variables that are to be tested. The variable where the likelihood ratio test statistic is the highest is included in the model. This procedure is repeated until the test statistic falls below some threshold value. In line with Imbens (2014), I use 1.00 for linear terms and 2.71 for quadratic/interaction terms. However, instead of logistic regression, I apply probit regression.

preventive effort (Schmitz, 2012). As a consequence, the direction of the association is not clear, but it is necessary to control for the association between risk attitudes, demand for medical treatment, and tariff choice. As a proxy for the real risk attitude towards health, I use the information whether the individual has participated in the sickness fund's bonus program in any of the years 2008 to 2012. In bonus programs, insurants are rewarded with cash or non-cash benefits by the sickness fund if they can prove a healthy lifestyle and preventive actions. Individuals that participate in a bonus program are more concerned with their health status and therefore suspected to be risk averse. I assume that risk attitude is rather stable over a period of five years, and that it is a valid signal if the individual participated in the bonus program in any of these years. In addition, I use a rich set of covariates to make the two groups as similar as possible with regard to relevant characteristics. Besides age and gender, I control for the insurance status, number of co-insured family members, education, occupational status, federal state of residence, and whether the individual has visited a doctor due a common cold. While some covariates were chosen due to economic theory, some were selected by the Imbens algorithm (Imbens, 2014). In addition, the algorithm selected a large set of quadratic and interaction terms. I argue that the set of covariates in this work is rich enough to take the CIA as given. To support this argumentation, I will test the effect of the participation on a pseudo outcome, namely on the probability of visiting a hospital.

Secondly, there has to be common support. This means for any given value of the confounding variables, a unit could potentially be observed in both the treatment group and the control group (Heckman et al., 1999). I trim the data to ensure the common support. Thirdly, the stable unit treatment value assumption (SUTVA) should be fulfilled. This assumption states that the treatment status of any unit does not affect the potential outcomes of other units, and that there is no variation in treatment between units. I expect the SUTVA to be fulfilled because one family's participation in the tariff should not affect another family's behavior with regard to medical demands. In addition, the treatment is equal for all participants. Finally, exogeneity of confounders should hold, i.e. the confounding variables are not influenced by the treatment in a way that is related to the outcome variables. As illustrated in Chapter 4, this assumption is fulfilled.

I use the kernel estimator for the matching procedure. Instead of using only few controls that are the closest to each treated individual, kernel matching uses many more controls. The number of used controls depends on the specific function and on the bandwidth. Here, I use the epanechnikov kernel with a bandwidth of 0.0005. This is a very small bandwidth which may lead to increased variance of the estimator but which is supposed to reduce the estimator's bias (Caliendo and Kopeinig, 2008). It ensures that only similar observations are matched and should make the two groups as similar as possible. In order to ensure common support, I drop those observations with a propensity score smaller than

the smallest or higher than the highest propensity score in the other group. Subsequently, I repeat the matching procedure and decide by means of a graphical analysis, in which region the common support is sufficient. Observations outside this region are removed. Now that the common support is fulfilled, I conduct the final, proper matching procedure. Standard errors are calculated by means of bootstrapping.

To make results more robust, I additionally use the combination of kernel matching and regression adjustment. This combination has the doubly-robustness property, i.e. results are consistent if either the propensity score or the regression function is correctly specified (Imbens and Wooldridge, 2009). Therefore, I repeat the matching procedure as described before and regress the matched sample, i.e. the treated and the weighted controls, on participation in 2012 and on all control variables that also have been used for the matching procedure:

$$w_i demand_i = \beta_0 + \beta_1 w_i participation_i + \beta_2 w_i age_i + \beta_3 w_i male_i + \dots + u_i$$

where w_i is the individual weight of each person and u_i is the error term. *demand* is representative for each of the analyzed outcomes.

6 Descriptive Statistics

As Table A (appendix) shows, the share of men between 15 and 54 years is slightly over-represented in the sample whereas the share of men aged 65 and older is a bit under-represented. For women, the situation is similar. The share of female insurance members aged 15 to 34 is a bit higher than in the whole SHI and the share of women aged 45 and older is a bit lower than in the SHI. This shows that the sickness fund has relatively young insurance members. The share of men is significantly higher than in the SHI. However, all in all the insurant's members' structure in the sample is sufficiently similar to that in the whole SHI in order to be able to make statements about the SHI in general.

In 2012, the final sample contains 195,014 insurance members whereof 704 participate in the premium refund tariff. Thereof, 331 members receive a premium refund. Once a person has chosen the optional tariff, he often stays with it for many years. Of the 704 participants in 2012, 521, 341, 132, and 44 individuals have already participated in 2010, 2009, 2008, or 2007, respectively.

The variables age, gender, number of co-insured family members, total expenditures before, participation in the bonus program, number of doctor visits before, common cold before, insurance status, and federal state have been pre-selected for the algorithm sug-

gested by Imbens (2014). The rest of the variables which have been used for matching (cf. Table 1) were then selected by the algorithm³.

Participants of the optional tariff differ significantly from unmatched non-participants (Table 1, treatment group and control group). This becomes obvious through the standardized bias which lies far above 5% for almost every variable that is included. On average, the treatment group is slightly younger, and the share of men is higher. In the treatment group, there are slightly fewer co-insured family members per insurance member. Furthermore, twice as many participants of the optional tariff, compared to non-participants, also participate or have participated in the bonus program. All measures of medical costs and of contacts with the health insurance system are significantly lower for participants. In both groups, the biggest share consists of compulsorily insured persons followed by the group of voluntarily insured persons, but in the treatment group the share of voluntarily insured individuals is higher than in the control group. Finally, in both groups most individuals have finished an apprenticeship and do not have a university-entrance diploma. In the treatment group, the share of individuals holding a university degree is considerably larger than in the control group.

³In addition to the variables in Table 1, the following variables also have been selected but are not shown in Table 1: age x age; age x no. co-insured family members; age x sex; bonus program x expenditures on drug prescriptions, before; no. co-insured family members x no. drug prescriptions, before; drug prescription, before x expenditures on doctor visits, before; age x no. doctor visits, before; sex x no. days sickness absence, before; common cold, before x drug prescription, before; no. co-insured family members x expenditures on doctor visits, before; total expenditures, before x total expenditures, before; drug prescription, before x doctor visit, before; total expenditures, before x no. drug prescriptions, before; no. drug prescriptions, before x no. drug prescriptions, before; total expenditures, before x expenditures on drug prescriptions, before; no. drug prescriptions, before x expenditures on drug prescriptions, before; no. drug prescriptions, before x sickness benefits, before; no. days sickness absence, before x doctor visit, before.

Table 1: Descriptive Statistics: Mean Values before and after Matching

	Treatment group Mean	Control group Mean	Standardized bias	Matched control group Mean	Standardized bias
Age	39.18 (9.89)	41.08 (11.41)	-17.8	39.20 (9.94)	-0.2
Male	0.78 (0.42)	0.65 (0.48)	27.8	0.78 (0.42)	0.3
Number of co-insured family members	1.49 (0.95)	1.63 (1.05)	-14.0	1.49 (0.95)	-0.1
Bonus program in 2008 to 2012	0.47 (0.50)	0.21 (0.41)	56.2	0.48 (0.50)	-1.7
Compulsorily insured	0.63 (0.48)	0.80 (0.40)	-37.7	0.63 (0.48)	1.4
Voluntarily insured	0.26 (0.44)	0.15 (0.35)	29.1	0.26 (0.44)	0.5
Pensioner	0.01 (0.08)	0.02 (0.15)	-12.2	0.01 (0.09)	-0.6
Apprenticeship, no university-entrance diploma	0.50 (0.50)	0.67 (0.47)	-34.7	0.50 (0.50)	0.4
No apprenticeship, university-entrance diploma	0.15 (0.36)	0.11 (0.31)	12.3	0.15 (0.36)	0.0
Apprenticeship, university-entrance diploma	0.13 (0.34)	0.10 (0.30)	10.5	0.13 (0.34)	-0.7
Degree of university of applied science	0.08 (0.28)	0.04 (0.20)	16.4	0.08 (0.28)	-0.8
University degree	0.13 (0.33)	0.05 (0.22)	26.5	0.13 (0.33)	0.8
Non-skilled worker	0.18 (0.38)	0.24 (0.42)	-14.2	0.17 (0.38)	1.3
Skilled worker, master, overseer	0.11 (0.31)	0.17 (0.38)	-17.1	0.11 (0.32)	-0.5
Employee	0.20 (0.40)	0.20 (0.40)	-1.0	0.20 (0.40)	0.2
Homeworker	0.00 (0.04)	0.00 (0.02)	3.5	0.00 (0.04)	0.6
Part-time worker	0.06 (0.24)	0.11 (0.31)	-16.8	0.06 (0.25)	-0.1

	Treatment group	Control group		Matched control group	
	Mean	Mean	Standardized bias	Mean	Standardized bias
Expenditures	381.08	659.28	-35.5	394.48	-1.7
total, before	(635.27)	(907.84)		(696.27)	
Expenditures on doctor	130.65	233.94	-54.1	134.97	-2.3
visits, before	(159.63)	(217.82)		(155.02)	
Doctor visit,	0.83	0.93	-34.0	0.83	-0.2
before	(0.37)	(0.24)		(0.37)	
No. of doctor visits,	4.72	8.39	-56.5	4.85	-2.0
before	(5.47)	(7.39)		(5.42)	
No. of visits at	1.98	3.57	-59.3	2.02	-1.4
general practitioner, before	(2.20)	(3.08)		(2.23)	
Expenditures on drug	18.09	50.94	-33.8	19.77	-1.7
prescriptions, before	(84.82)	(108.34)		(104.75)	
Drug prescription,	0.63	0.86	-54.1	0.63	0.6
before	(0.48)	(0.35)		(0.48)	
No. of drug	1.09	2.60	-65.3	1.15	-2.7
prescriptions, before	(1.84)	(2.69)		(2.13)	
Expenditures on sickness	17.8	39.14	-8.6	19.08	-0.5
benefit, before	(225.92)	(267.58)		(222.35)	
No. Of days on	3.52	7.26	-29.3	3.63	-0.9
sickness absence, before	(10.19)	(14.86)		(9.94)	
Common cold,	0.04	0.06	-6.7	0.04	-1
before	(0.20)	(0.23)		(0.20)	
N	704	232,093		194,310	

Standard errors in parentheses.

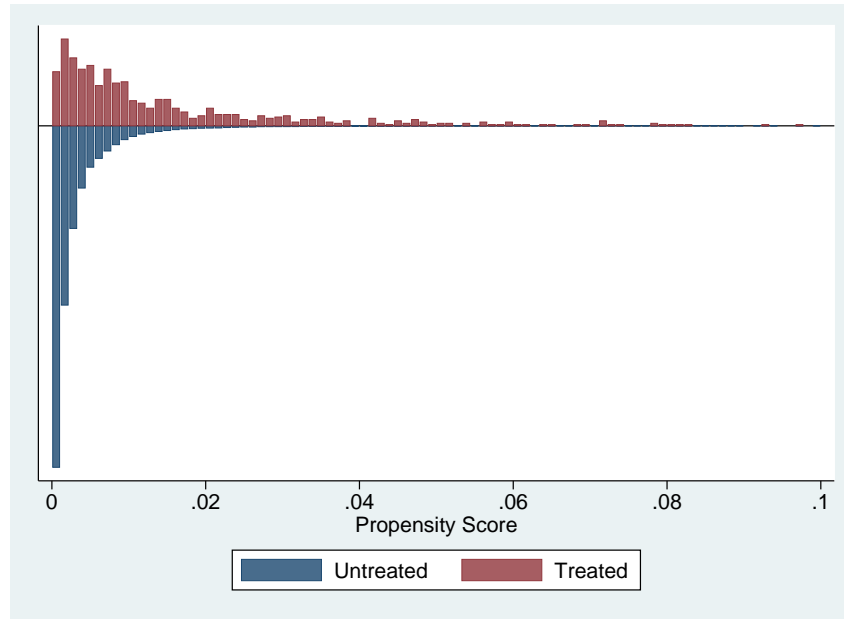
7 Results

7.1 Matching Quality

After the matching procedure, the average values of all relevant variables have converged between the treatment and the matched control group (Table 1, treatment group and matched control group). The standardized bias is less than 4% for all variables that have been used for matching, also for those that are not shown in Table 1. According to Caliendo and Kopeinig (2008), with a standardized bias below 5% the matching procedure can be assessed being successful.

Another condition that has to be fulfilled in order to have good matches is common support (see above). Figure 3 shows that this condition is fulfilled, after carrying out the trimming procedure described in chapter 5: For each value of the propensity score, individuals of the treatment group are closely surrounded by individuals of the control group.

Figure 3: Common Support



7.2 Estimation Results

Table 2 (column 1) presents the estimation results after having constructed a control group by using kernel matching. The probability of visiting a general practitioner is significantly reduced by 6 percentage points. In contrast, there is no effect on the probability of visiting a specialist. These findings are in line with theory. However, I also do not find any effect on the probability of visiting a doctor due to a common cold which contradicts my hypothesis. This means there is evidence that the implicit deductible is able to reduce the probability of visiting a doctor, but there is no clear evidence that this is due to a reduction in ex-post moral hazard.

Expenditures on doctor visits decrease for specialists whereas there is no effect for general practitioners. This confirms my hypothesis that the patient can decide whether he visits a doctor or not, but the extent of treatment is usually decided on by the doctor.

Next, the matching results are compared to results from a combination of matching and regression (column 2). This combination of methods has the advantage of being more robust than either matching or regression alone. The results show that matching alone already has lead to estimates that are very similar to those from the combination of the two methods. The differences in the estimation results between the two specifications are negligible, and the kernel matching estimator (column 1) is assessed being an appropriate evaluation method.

Table 2: Estimation Results

Outcome	(1) Kernel matching	(2) Kernel matching with regression
Doctor visit (general practitioner)	-0.06*** (0.02)	-0.05*** (0.02)
Doctor visit (specialist)	-0.02 (0.02)	-0.02 (0.02)
Common cold	-0.00 (0.00)	-0.00 (0.00)
Expenditures on doctor visits (general practitioner)	0.18 (10.57)	0.57 (9.89)
Expenditures on doctor visits (specialist)	-26.68*** (8.78)	-25.37*** (8.87)
N treated	704	704
N controls	194,310	194,053

Standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.10$. Each line contains coefficients of separate regressions.

7.3 Sensitivity Analysis

In order to test whether the results are stable, I carry out some robustness checks. I use a logit model instead of a probit model to predict the propensity score. This is done because there are many small values of the propensity score, and especially in the tails of the distributions there may be some differences between using a logit and a probit. There are nearly no differences between the two models (results not shown in the paper).

In addition, in order to find support that the CIA is fulfilled, I run a regression on a pseudo outcome, i.e. an outcome that is known to be unaffected by the treatment (Imbens and Wooldridge, 2009). As the pseudo outcome, I choose the probability of visiting a hospital in 2012. There is no significant effect of participating in 2012 on the probability of visiting a hospital in the same year (coefficient = 0.00, bootstrapped standard error = 0.01, p-value = 0.68). Since the CIA cannot be directly tested, this is not a proof, but it supports the CIA. It implies that the treated observations are not distinct from the controls in that the distribution of $Y(0)$ for the treated units is comparable to the distribution of $Y(0)$ for the controls.

8 Conclusion

I find that participation in the premium refund tariff significantly reduces the probability of visiting a general practitioner by 6 percentage points. However, the probability of visiting a doctor due to a common cold is not affected. This means that there is no clear evidence that the reduction in demand for medical service is due to reduced ex-post moral hazard.

I suspect that many individuals face some illness over the year which is either severe such that they have to visit a doctor, or the individual is not able to judge whether the illness is severe or not. Then he also has to see the doctor to be sure. In either case, during the rest of the calendar year there is no more incentive to forgo an unnecessary doctor visit. Another problem could be in the definition of trivial sicknesses. Perhaps it is not enough to focus on one single ICD-10 code (J00). Instead, one could add other trivial diseases to the analysis. Anyway, the amount of the premium should be high enough to encourage individuals to forgo unnecessary doctor visits.

Because the insurants' structure is sufficiently similar to that of the whole SHI with respect to age and gender, the presented results can all in all give a usable estimate of the analyzed research question, and it can be concluded that participation in the premium refund tariff leads to a reduction in medical expenditures. It has to be analyzed, though, whether this

is due to reduced ex-post moral hazard or whether the effect is driven through a different channel.

REFERENCES

Augurzky, B., Bauer, T.K., and Schaffer, S. (2006) Copayment in the German Health System: Does It Work? IZA Discussion Paper No. 2290.

Breyer, F. and Buchholz, W. (2009) *Ökonomie des Sozialstaats*. Berlin Heidelberg: Springer.

Caliendo M, Kopeinig S (2008) Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, Vol. 22 (1), 31-72.

Die Bundesregierung (2010) (<http://www.bundesregierung.de/Content/DE/Magazine/Magazin/SozialesFamilieBildung/093/093.html?context=Inhalt%2C4>, 14.03.2013).

GBE, Gesundheitsberichterstattung des Bundes (2014) (<http://www.gbe-bund.de/gbe10/i?i=856:14442317D>, 20.09.2014).

GBE, Gesundheitsberichterstattung des Bundes (2015) (<https://www.gbe-bund.de>, 08.02.2015).

Heckman J, LaLonde R, Smith J (1999) The economics and econometrics of active labor market programs. In O. Ashenfelter and D. Card (eds), *Handbook of Labor Economics*, Vol. 2, 1865-2097. Amsterdam: Elsevier.

Imbens, G.W. (2014) Matching Methods in Practice: Three Examples. IZA Discussion Paper No. 8049.

Imbens, G.W. and Wooldridge, J.M. (2009) Recent Developments in the Econometrics of Program Evaluation. *Journal of Economic Literature*, Vol. 47 (1), 5-86.

Jacobs, K., Kip, C., Schulze, S. (2010) Mehr Kostenerstattung in der GKV? Auswirkungen auf Kostentransparenz und Inanspruchnahmeverhalten – Wunsch und Wirklichkeit. *GGW*, Vol. 10(4), 17-26.

Kiil, A. and Houlberg, K. (2014) How does copayment for health care services affect demand, health and redistribution? A systematic review of the empirical evidence from

1990 to 2011. *European Journal of Health Economics*, Vol. 15 (8), 813-828.

Leuven, E. and Sianesi, B. (2003) PSMATCH2: Stata module to perform full Mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing. (<http://ideas.repec.org/c/boc/bocode/s432001.html>, version 4.0.10).

Malin, E.M. and Schmidt, E.M. (1995) Beitragsrückzahlung in der GKV: ein Instrument zur Kostendämpfung? *Die Betriebskrankenkasse* 12/1995, 759-763.

Manning, W.G., Newhouse, J.P., Duan, N., Keeler, E., Leibowitz (1987): Health insurance and the demand for health care. Evidence from a randomized experiment. *American Economic Review*, 77 (3), 251-277.

Paquet, R. and Stein, M. (2007) Keine reine Freude an den Wahlтарifen – Umfrage zeigt kritische Verbraucher. *Die Betriebskrankenkasse* 05/2007, 192-196.

Pütz, C. and Hagist, C. (2006) Optional deductibles in social health insurance systems. Findings from Germany. *European Journal of Health Economics*, Vol. 7 (4), 225-230.

Regorz, A. (2007) GKV-Wahlтарife: Wie riskant für Krankenkassen sind Tarife zur Beitragsrückerstattung? (http://www.regorz.de/analyse_beitragsrueckerstattung.pdf, 15.03.2014).

Rosenbaum, P.R. (2002) Covariate Adjustment in Randomized Experiments and Observational Studies. *Statistical Science*, Vol. 17 (3), 286-327.

Schmitz, H. (2012) More health care utilization with more insurance coverage? Evidence from a latent class model with German data. *Applied Economics*, 44, 4455-4468.

Schulze Ehring, F. and Weber, C. (2007) Wahlтарife in der GKV – Nutzen oder Schaden für die Versichertengemeinschaft? WIP-Diskussionspapier 4/07.

Appendix

Table A: Comparison of the sample and the SHI population in 2010

	Sample: Share in %		SHI: Share in %	
	Men	Women	Men	Women
14 and younger	0.12	0.11	0.13	0.12
15 - 10	2.54	1.68	0.78	0.57
20 - 24	4.93	3.81	2.85	2.62
25 - 29	5.24	4.37	3.88	3.65
30 - 34	5.58	4.02	3.78	3.36
35 - 39	5.72	3.42	3.69	3.18
40 - 44	7.30	4.14	4.92	4.27
45 - 49	7.55	4.29	5.39	4.73
50 - 54	6.07	3.31	4.69	4.19
55 - 59	4.37	2.20	4.01	3.61
60 - 64	3.44	1.78	3.42	3.22
65 - 69	2.78	1.85	3.37	3.80
70 - 74	2.61	1.78	3.85	4.54
75 - 79	1.41	1.03	2.47	3.20
80 and older	1.18	1.38	2.43	5.29
Total	60.83	39.17	49.65	50.35