Deutsches Institut für
Wirtschaftsforschung

DIW BERLIN

www.diw.de

SOEP The German Socio-Economic Panel Study

SOEPpapers

on Multidisciplinary Panel Data Research

141

Hendrik Schmitz

Do Optional Deductibles Reduce the Number of Doctor Visits? Empirical Evidence with German Data

Berlin, October 2008

SOEPpapers on Multidisciplinary Panel Data Research
at DIW Berlin

This series presents research findings based either directly on data from the German Socio-Economic Panel Study (SOEP) or using SOEP data as part of an internationally comparable data set (e.g. CNEF, ECHP, LIS, LWS, CHER/PACO). SOEP is a truly multidisciplinary household panel study covering a wide range of social and behavioral sciences: economics, sociology, psychology, survey methodology, econometrics and applied statistics, educational science, political science, public health, behavioral genetics, demography, geography, and sport science.

The decision to publish a submission in SOEPpapers is made by a board of editors chosen by the DIW Berlin to represent the wide range of disciplines covered by SOEP. There is no external referee process and papers are either accepted or rejected without revision. Papers appear in this series as works in progress and may also appear elsewhere. They often represent preliminary studies and are circulated to encourage discussion. Citation of such a paper should account for its provisional character. A revised version may be requested from the author directly.

Any opinions expressed in this series are those of the author(s) and not those of DIW Berlin. Research disseminated by DIW Berlin may include views on public policy issues, but the institute itself takes no institutional policy positions.

The SOEPpapers are available at
http://www.diw.de/soeppapers

Editors:

Georg **Meran** (Vice President DIW Berlin)
Gert G. **Wagner** (Social Sciences)
Joachim R. **Frick** (Empirical Economics)
Jürgen **Schupp** (Sociology)

Conchita **D'Ambrosio** (Public Economics)
Christoph **Breuer** (Sport Science, DIW Research Professor)
Anita I. **Drever** (Geography)
Elke **Holst** (Gender Studies)
Frieder R. **Lang** (Psychology, DIW Research Professor)
Jörg-Peter **Schräpler** (Survey Methodology)
C. Katharina **Spieß** (Educational Science)
Martin **Spieß** (Survey Methodology)
Alan S. **Zuckerman** (Political Science, DIW Research Professor)

ISSN: 1864-6689 (online)

German Socio-Economic Panel Study (SOEP)
DIW Berlin
Mohrenstrasse 58
10117 Berlin, Germany

Contact: Uta Rahmann | urahmann@diw.de

Do Optional Deductibles Reduce the Number of Doctor Visits? Empirical Evidence with German Data

Hendrik Schmitz*
Ruhr Graduate School in Economics (RGS Econon)

October 29, 2008

Abstract

Deductibles in health insurance are often regarded as a means to contain health care costs when individuals exhibit moral hazard. However, in the absence of moral hazard, voluntarily chosen deductibles may instead lead to self-selection into different insurance contracts. We use a set of new variables in the German Socioeconomic Panel for the years 2002, 2004, and 2006 that measure individual health more accurately and include risk-attitudes towards health in order to determine the price elasticity of demand for health care. A latent class approach that takes into account the panel structure of the data reveals that the effect of deductibles on the number of doctor visits is negligible. Private add-on insurance increases the number of doctor visits. However, altogether the effects of the insurance state on the demand for doctor visits are small in magnitude.

JEL Classification: I11, I18, G22

Keywords: health insurance, deductibles, add-on insurance, count data, latent class panel model

*I am grateful to Reinhold Schnabel, Daniel Baumgarten, Hendrik Jürges, Alfredo Paloyo, Annika Meng, Volker Clausen, Wolfram Richter, and Nicolas Ziebarth for valuable comments. This paper benefited from discussion at the EEA annual meeting 2008 in Milan, the 2nd RGS doctoral conference in Essen, and the annual meeting of Verein für Socialpolitik 2008 in Graz. Financial support by the Leibniz association is gratefully acknowledged. All errors are my own. All correspondence to Hendrik Schmitz, Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI Essen), Hohenzollernstr. 1–3, 45128 Essen, Germany, Fax: +49-201-8149-500, e-mail: hendrik.schmitz@rwi-essen.de

1 Introduction

Increasing health care expenditures have been an issue for several decades in most industrialised countries. While demographic change and technological progress can be seen as the main driving forces behind the increase, both factors can hardly be tackled in order to contain costs. Moreover, they might lead to even higher expenditures in the future. In order to lower costs, it seems more promising to detect and to reduce inefficiencies in the health care sector.

One inefficiency from the demand side is the problem of moral hazard induced by insurance, defined as "the change in health behavior and health care consumption caused by insurance" (Zweifel and Manning, 2000). This excess demand for (supposedly trivial) health services could be suppressed via cost-sharing, e.g., deductibles or co-payments for doctor or hospital visits. However, the optimal amount of cost-sharing is a priori unclear. In a world with risk-averse individuals and without moral hazard, full cover insurance (i.e., no deductibles) is a first-best solution. Less than full cover insurance (e.g., with optional deductibles) could instead lead to adverse selection. Yet, when moral hazard is present, the first-best solution is no longer feasible and the introduction of mandatory deductibles can lead to a second-best solution (see Breyer et al., 2004). Hence, finding out the price elasticity of demand for health services (i.e., the amount of moral hazard) is an important empirical task in order to design an optimal health insurance system.

There are several studies that analyse the impact of the state of insurance on the number of doctor visits or hospitalisations with German data, e.g., Pohlmeier and Ulrich (1995), Geil et al. (1997), and Riphahn et al. (2002). They compare the behaviour of individuals covered by private health insurance with those covered by public insurance. In the German system, the former can choose between having a deductible or not whereas there are basically no deductibles for the latter.[1] The results are mixed. Pohlmeier and Ulrich (1995) find that the probability of visiting a general practitioner (GP) is higher for the publicly insured, that is, implicitly, for those with more insurance cover. Geil et al. (1997) show that, for females, being covered by public insurance has a positive (but moderate) effect on the number of hospitalisations, while no significant effect was found for males. Finally, Riphahn et al., (2002) find that private add-on insurance raises the number of hospitalisations of males while other variables that indicate the insurance state are not significant.

By exploiting the panel structure of the data, Riphahn et al. (2002) and Geil et al. (1997) account for unobserved individual effects that affect doctor visits, such as individual frailty, using a random effects model. However, none of the mentioned studies account

[1] A brief introduction to the German health insurance system will be given in chapter 2.

for potential endogeneity problems that come with the insurance choice. That is, finding positive effects of full cover insurance on the demand for doctor visits should not necessarily be interpreted as a causal effect and, thus, not as moral hazard. When making a decision about the insurance plan, the individual takes into account her expected demand for health services in the future. A person who knows that she needs many visits to the doctor in the future might not buy private insurance with deductibles. Instead, she would prefer either private insurance without a deductible or even public insurance. Thus, the health status both affects the demand for doctor visits and the insurance type and an incompletely observed health status by the researcher might lead to biased results. Another problem are unobserved preferences like risk-aversion. Risk-averse individuals tend to demand more doctor visits and at the same time prefer full cover insurance.

Altogether, because there are no experimental data available for Germany (as, e.g., the well known RAND health insurance experiment) and deductibles are in general not mandatory but can be chosen optionally, finding positive effects of public insurance on the number of doctor visits (or hospital stays) might merely reflect a selection effect. One way to deal with the endogeneity problem is an instrumental variable approach (see, e.g., Windmeijer and Santos Silva, 1997 and Schellhorn, 2001, for an application with Swiss data), although, in practice, it is hard to find good instruments, i.e. instruments which are valid and not weak.

This paper addresses possible endogeneity as an omitted variable bias problem and uses better information on the health status of individuals than is normally available in survey data and direct measures of individual risk-aversion. Using these additional pieces of information, the tests performed cannot reject the hypothesis that the insurance state is exogenous. In order to account for unobserved heterogeneity (which remains important even after inclusion of the newly available information on health status and risk-aversion), this paper follows the approach of Bago d'Uva (2006) who uses a latent class hurdle model for panel count data.

It turns out that while the insured with deductibles have less doctor visits than the insured without deductibles, no causal effect of the deductible itself can be found. This leads to the conclusion that self-selection instead accounts for the lower number of doctor visits of the insured with deductibles. The positive effect of add-on insurance on the number of doctor visits, however, can be found even after accounting for unobserved heterogeneity.

The paper is structured as follows. Section 2 gives a short introduction into the German health insurance system. Section 3 describes the data and the empirical strategy, while section 4 presents the estimation results. Section 5 concludes.

2 Institutional Background

The German health insurance system consists of two parts. Roughly 88 per cent of the population are insured by public health insurance (also called the statutory health insurance, SHI). It is statutory for all individuals with earnings below a certain income threshold (3,975 Euro per month in 2007) and financed by payroll taxes. Therefore, it is independent of age, gender, or health status of the insured. Since 1997, the publicly insured are allowed to choose between different insurance companies yet the benefit package is heavily regulated and does not vary much between companies.[2] The insurance companies are not allowed to reject applicants and non-working family members are covered without an extra premium. Until the most recent reform which took effect in April 2007, there were basically no (optional) deductibles in the public scheme. However, in 2004, a co-payment of 10 Euro for the first visit in a quarter was introduced.[3] Furthermore, there were low co-payments for hospitalisations (10 Euro per calendar day up to a maximum of 28 days) and for prescription drugs (Winkelmann, 2004). As of April 2007, the insured can choose between different contracts which include contracts with and without deductibles. The publicly insured can additionally purchase private add-on insurance that either increases quality (e.g. double rooms in hospitals) or covers co-payments on denture or glasses.

Individuals who earn more than the income threshold, the self-employed, and civil servants are allowed to opt out of the public insurance system and instead buy private insurance (which accounts for the remaining 12 per cent of the German population, abstracting from a small group without any insurance coverage). They are usually called the "voluntarily insured". The private insurance premium does not depend on income but is instead a risk-equivalent contribution (depending on age, gender, and health status). Private insurance companies can reject bad risks. Furthermore, privately insured individuals have to pay higher premia in order to cover non-working family members. Thus, being a bad risk or having dependents might be two of the reasons for staying voluntarily in the public system for about 50 per cent of all the individuals who are allowed to opt out although private insurance often is perceived to lead to higher quality of health care. Private insurance companies usually offer a set of different contracts, including the choice of deductibles. As opposed to the public insurance scheme, the privately insured first have to pay the costs they incur and then get reimbursed at the end of the year. The insured who do not send in claims get a reimbursement of a part of their insurance premium. Individuals

[2]About 95 per cent of the package that all insurers have to provide is stated in Social Code Book V (see Buchner and Wasem, 2003).

[3]See Augurzky et al. (2006) for an evaluation of this reform. The authors do not find reductions in the number of doctor visits due to the reform and conclude that 10 Euro is too low to lead to significant effects.

who opt out of the public insurance system are in general not allowed to re-enter later as long as they do not fall below the income threshold. Furthermore, contracts with a deductible cannot be transformed into full cover contracts without proof of good health. Hence, the decision to buy private insurance and about the deductible are practically lifetime decisions.

Regarding incentive schemes faced by the insured, four groups can be compared. The privately insured with a deductible can be expected to have the strongest incentive to contain costs, followed by the privately insured without deductibles. The publicly insured without add-on insurance might have a lower incentive to contain costs than privately insured but a higher one than those who purchase add-on insurance.

3 Empirical Model

3.1 Data and Variable Description

The database for the empirical analysis is the German Socioeconomic Panel (GSOEP), which started in 1984 in West Germany and was extended to include East Germany in June 1990. There were several refreshments resulting in a sample size of more than 20,000 adult individuals living in more than 12,000 households that participated in the GSOEP survey in 2006 (see, e.g., Wagner et al., 2007). The GSOEP includes questions about the number of doctor visits within the last three months prior to the interview and the number of hospital trips in the previous year. Due to the generally low price elasticity of hospital trips, we focus on the number of doctor visits to detect moral hazard.[4]

The number of doctor visits depends, to a large extent, on the health status of the individual. However, direct measurement of the health status is somewhat complicated and, especially in general surveys such as the GSOEP, often restricted to the self-assessed health status (SAH, on a 5-point scale from very good to bad). As often argued, this measure is prone to measurement error and reporting bias, possibly leading to biased coefficient estimates (see, e.g., Bound, 1991 or Crossley and Kennedy, 2002). Furthermore, it is unlikely that a complete picture of individual health is captured by the SAH. Two newly available and more objective measures are used here to alleviate this problem: the Physical Component Summary Scale (PCS), a measure of physical health, and the Mental Component Summary Scale (MCS) a measure of mental health (see Nübling et al., 2007

[4]The data used in this paper were extracted using the Add-On package PanelWhiz v2.0 (Nov 2007) for Stata. PanelWhiz was written by Dr. John P. Haisken-DeNew (john@panelwhiz.eu).
The PanelWhiz generated DO file to retrieve the SOEP data used here and any Panelwhiz Plugins are available upon request. Any data or computational errors in this paper are my own. Haisken-DeNew and Hahn (2006) describe PanelWhiz in detail.

for a description). These measures are based on the SF12-questionnaire in the GSOEP that includes several questions about health quality and satisfaction of the individuals.[5] Both variables are calculated using explorative factor analysis and lie between 0 and 100, with a higher value indicating a better health status. The mean value of the GSOEP 2004 population is 50 points with a standard deviation of 10 points. Although both measures are also based on self-reported information, they give a much more detailed picture of the true health status.

As an objective measure of individual health we use the body mass index (BMI) and the number of hospital stays in the previous year. The three variables PCS, MCS, and BMI are only available in the waves 2002, 2004, and 2006, hence, these three waves of the GSOEP are used for the analysis. It turns out that after controlling for these health measures, the self assessed health status still contains information to explain the number of doctor visits and should therefore not be left out.

One further variable that is newly available in the GSOEP is a self-assessed attitude towards risk concerning health matters on an 11-point scale from 0 (very risk-averse) to 10 (not at all risk-averse).[6] The risk-attitude can be expected to have an important impact on the demand for health care. It was only asked in 2004. However, it seems reasonable to assume that this preference did not change in the span of five years. Although the attitude towards risks are self-assessed, Dohmen et al. (2005) show in an experimental setting with a subgroup of the GSOEP households that it is reliable. Another variable that reflects preferences of the individual is the assessment about worries concerning the own health (on a 3-point scale between 1= very concerned and 3= not at all concerned). It is on the one hand informative about the health status, on the other hand, conditional on the health status, it captures a part of the individual doctor visiting behaviour.

Using these new variables (together with a set of other socio economic variables which are not the focus of this study), we can control for a lot of important heterogeneity that has an effect on the number of doctor visits and the insurance state at the same time and which remained unobserved in previous studies.

Only the privately insured can have contracts with deductibles. In our dataset the information on deductibles of the privately insured is restricted to a binary variable indicating the existence or absence of a deductible. Although the publicly insured can indicate what kind of private add-on insurance they hold, we collapse these into a binary variable indicating whether or not an individual holds some kind of add-on insurance due to the low coverage of the publicly insured with add-on insurance (only about 14 per cent of all

[5]Questions concern, among others, bodily pain, stress, ability to carry out everyday tasks, phases of melancholy, etc. Using all the 12 variables as regressors instead of the two combined measures did not improve the estimations.

[6]The question in the GSOEP is: "How would you rate your willingness to take risks with your health?"

publicly insured hold some kind of private add-on insurance). However, add-on insurance that cover hospital stays or medical costs abroad are not included here. More precisely, the binary variable *Add-on* states whether an individual holds supplementary insurance that covers dentures, corrective devices, some kinds of therapeutic measures, or others.

We restrict the sample of individuals to those older than 25 years[7] and also exclude civil servants due to their special insurance status. In general, the employer of a civil servant covers 50 per cent (or more) of the health care costs while civil servants have to insure only the remaining 50 per cent (either privately or publicly). Treating a civil servant with private insurance and deductible similar to other privately insured would certainly bias the results. All together, we use information from 18,024 individuals leading to 46,440 observations in person-year form after exclusion of observations with missing values in any of the variables used for the regression analysis.

Table 1: Doctor Visits of Subgroups

	Average # of visits	sd	Probability of one visit	sd	Average # of visits if > 0	sd	Number of obs.
Whole Sample	2,53	4,05	0,69	0,46	3,64	4,42	46440
Public Insurance	2,56	4,02	0,70	0,46	3,64	4,36	41345
– voluntary	2,03	3,48	0,64	0,48	3,16	3,91	4540
– with add-on	2,61	4,28	0,71	0,45	3,66	4,67	3311
– volunt. with add-on	2,28	4,00	0,71	0,46	3,22	4,43	645
Private Insurance	2,26	4,27	0,62	0,48	3,63	4,93	5079
– with deductible	1,97	3,85	0,58	0,49	3,38	4,54	2818
– without deductible	2,62	4,72	0,67	0,47	3,91	5,31	2261

Source: GSOEP, pooled years 2002, 2004, 2006; individuals older than 25, no civil servants.

The number of doctor visits in the previous three months of several groups with different insurance states in the pooled sample is given in Table 1. The overall mean is 2.53, with 31 per cent of all the individuals having no doctor visit at all. Conditional on having at least one visit, the average number of doctor visits in the whole sample is 3.64. The group of publicly insured has a higher number of doctor visits than the group of privately insured (2.56 vs. 2.26). However, both groups are hardly comparable because, first of all, the group of privately insured consists of individuals with higher income and better education - characteristics that are known to be correlated with better health (see table 4 in the appendix for means of the covariates for different subgroups). Furthermore, this group has the better risk-pool because bad risks are either rejected by private health

[7]This excludes individuals who are either covered by their parents' insurance.

insurance companies or would have to pay high contributions that preclude them from buying private health insurance. The groups of voluntarily publicly insured and privately insured are more suitable for comparison. Although the first group can be expected to have the worse risk-pool (due to the abovementioned reasons), it exhibits a lower number of doctor visits. This comes as a surprise and might point to physician-induced demand.[8] The comparison of the privately insured with and without deductible shows a clear picture: the privately insured with a deductible have much less doctor visits (1.97 vs. 2.62), a lower probability of visiting a doctor (58 per cent vs. 67 per cent), and even fewer doctor visits conditional on having at least one visit (3.38 vs. 3.91). Likewise, individuals with add-on insurance exhibit slightly higher numbers of doctor visits. However, these are only unconditional numbers which do not account for different health statuses across groups or different attitudes towards visiting a doctor. That is, in order to establish causation, a more detailed analysis that controls for important observable and unobservable factors is necessary.

3.2 Hurdle Model and Endogeneity of the Insurance Choice

It is often argued that the observed number of doctor visits is a result of two different (and probably independent) decision-making processes. First, the patient decides whether or not to see a doctor in case of an illness. Once a doctor is seen, however, the doctor determines the length of the treatment. Hence, a hurdle model seems to be the most appropriate formulation in order to explain the number of doctor visits (Mullahy, 1986; Pohlmeier and Ulrich, 1995). Let

$$g(y_{it}) = \begin{cases} f_1(0|x_{it}) & \text{if } y_{it} = 0 \\ (1 - f_1(0|x_{it}))f_2(y_{it}|x_{it}, y_{it} > 0) & \text{if } y_{it} > 0 \end{cases} \tag{1}$$

where $f_2(y_{it}|x_{it}, y_{it} > 0) = f_2(y_{it}|x_{it})[1 - f_2(0|x_{it})]^{-1}$, y_{it} is the number of doctor visits of individual i at time t, and x_{it} is a vector of covariates. $f_1(0|x_{it}) = P(y_{it} = 0|x_{it})$ is a binary function that determines the probability of having no doctor visit at all in a given period. Given that the number of doctor visits exceeds zero (with probability $1 - f_1(0|x_{it})$), a truncated-at-zero function $f_2(y_{it}|x_{it}, y_{it} > 0)$ determines the exact number of visits.

When analysing the impact of health insurance on the demand for health care services, there is possibly an endogeneity problem. This mainly stems from omitted variable bias in two cases. First, due to the risk-equivalised insurance premia of private insurance

[8]Physicians can charge treatments of privately insured at least 2.3 times higher than of publicly insured, which clearly gives them an incentive to focus more on privately insured. See Jürges, 2007, for a recent analysis of supplier-induced demand with data from the GSOEP.

companies, individuals with better health tend to buy private insurance (and a deductible) and have less doctor visits at the same time. In survey data, however, the health status is typically poorly observed. Thus, a lot of information about the true health status remains unobserved, which probably leads to biased estimates. The second source of endogeneity might be preferences like risk-aversion that, on one hand, affect the number of doctor visits (risk-averse are likely to have more doctor visits given a fixed health status) and, on the other hand, the insurance state (risk-averse tend to buy full cover insurance). However, including the new set of health and risk variables might reduce the endogeneity problem.

Given the economic interpretation of the hurdle model, one can argue that possible endogeneity of the insurance choice should mainly play a role in the first stage, namely when the patient has full control. Thus, to determine the degree of endogeneity, we first focus on that stage and model the decision to see a doctor together with the decision about a certain insurance contract. Consider the following bivariate model

$$y_1 = 1[x_1\beta_1 + h\delta_1 + \alpha_1 y_2 + \mu_1 > 0] \tag{2}$$

$$y_2 = 1[x_2\beta_2 + h\delta_2 + \mu_2 > 0] \tag{3}$$

where y_1 is the binary decision to see a doctor and y_2 the binary decision about the insurance contract (e.g. deductible yes/no, or add-on insurance yes/no), h is the observed health status, and x_1 and x_2 are socio-economic variables like age, education, and income. Finally, μ_1 and μ_2 capture unobserved effects like unobserved health and preferences towards visiting a doctor. We assume that the number of doctor visits in one period depends on the insurance status whereas the insurance status does not depend on the number of doctor visits in that same period (given a certain health status). The correlation between μ_1 and μ_2, $\rho = corr(\mu_1, \mu_2)$, does not equal 0 if the unobserved effects that affect the number of doctor visits and the insurance choice are correlated. Assuming a bivariate normal distribution of the error terms, the parameters of this model (and the correlation ρ can be estimated by a bivariate probit.

We fit two different regressions, one where the insurance variable is *deductible* and one where it is *add-on*. As discussed in chapter 2, only the voluntarily insured can opt for an insurance with a deductible. The private add-on insurance, on the other hand, is only interesting for the publicly insured. Hence, we only include the voluntarily insured (precisely, these are the voluntary publicly insured together with the privately insured) in the first regression and the publicly insured in the second one.

In the first regression, ρ can be expected to be negative if the insured who expect to have fewer doctor visits in the future (for unobserved reasons) tend to buy insurance with a deductible. In the group of publicly insured, the insurance-state variable is add-on insurance. Unlike in the deductible case, add-on insurance leads to more insurance coverage, hence, a positive ρ can be expected.

Although the nonlinear model is identified by functional form, we add variables to x_2 that are assumed to have no influence on the decision to visit a doctor and are thus excluded from x_1. These are risk-aversion concerning financial matters and the general attitude towards co-payments for health care services.[9] The bivariate probit model allows for endogeneity of the insurance choice in the doctor-visits equation via correlation of the error terms. However, as table 2 shows, in neither of the two cases is the estimated correlation coefficient significantly different from zero. We follow (Knapp and Seaks, 1998) in using a t-test on the correlation coefficient as a test for endogeneity of the dummy regressors. Here, the hypothesis of $\rho = 0$ (no endogeneity) cannot be rejected in either case.

Table 2: Estimated correlation in bivariate probit

Equation	$\hat{\rho}$	$\hat{se}(\hat{\rho})$	Observations
Deductible	0.050	(0.18)	9246
Add-on	-0.053	(0.11)	39698

Estimations done by Stata program *biprobit*, standard errors clustered by individuals. Full estimation results in the appendix, see table 5.

In both equations even the sign is different from what is expected, however, the values are very close to zero. The set of new variables is highly jointly significant in the doctor-visits equation in both cases. It is furthermore jointly significant in the insurance-equations (only at the 10%-level in the add-on-equation; test-statistics not reported here). Thus, it can be argued that capturing information from the new health variables and the degree of risk-aversion (plus health worries) reduces the endogeneity problem by a substantial amount. But one can think of even more reasons originating in the insurance system that make it possible for endogeneity to be less of a problem than possibly expected. As discussed in chapter 2, opting out of private insurance and deciding on deductibles are practically lifetime decisions. While the (partly unobserved) health status should have a high impact on the decision about health insurance in a given year, it only affects the number of doctor visits in the following years but not the decision about the insurance state (at least as far as private insurance and deductibles are concerned). Moreover, Grabka (2006) gives another important reason for the privately insured to choose a contract with a deductible that is independent of changes in the health status of the insured. Unlike in

[9]Risk-aversion concerning financial matters is asked in the same way as the risk-aversion concerning health. The question about the attitudes towards co-payments was asked only in 2002. As in the case of risk-aversion, we assume that this is a time-invariant preference.

the case of public insurance, cost containment and the stability of contribution rates have not been a big issue in the past decades in the private insurance system. This has led to a much higher proportional increase in costs than in the public sector and, thus, in steadily increasing contribution rates for the privately insured. One way for an insured to stop an increase in the contribution rate in a given year is to transform a contract without a deductible into one with a deductible to keep the basic insurance premium stable. In this case, the decision about buying insurance with a deductible is not affected by a change in the health status but by other reasons.

In the next section we leave aside endogeneity concerns and turn back to the hurdle model.

3.3 Latent Class Hurdle Model

Given that the dependent variable (number of doctor visits in the previous three months) is an integer, it is appropriate to use a count data model in order to specify the two underlying functions in the hurdle model. While the Poisson model is a good starting point for count data, it is often seen to be too restrictive due to its assumption of the equality of mean and variance of the dependent variable, which is clearly not the case here (see table 1). In order to allow for over-dispersion, it is common to introduce a gamma-distributed error term, ending up with the negative binomial distribution ("negbin"; see, e.g., Cameron and Trivedi, 2005, ch. 20 for a derivation) with the following probability density function:

$$f(y_{it}|\mu,\alpha) = \frac{\Gamma(\alpha^{-1}+y_{it})}{\Gamma(\alpha^{-1})(\Gamma(y_{it}+1))}\left(\frac{\alpha^{-1}}{\alpha^{-1}+\mu}\right)^{\alpha^{-1}}\left(\frac{\mu}{\mu+\alpha^{-1}}\right)^{y_{it}} \tag{5}$$

where μ =exp$(x'_{it}\beta)$ and α is the over-dispersion parameter.

Combining the negative binomial distribution with the hurdle structure in (1), f_1 becomes

$$f_1(0|x_{it}) = P(y_{it}=0|x_{it},\beta_1) = (\mu_1+1)^{-1} \tag{6}$$

where μ_1 =exp$(x'_{it}\beta_1)$.[10] The truncated part in (1) becomes

$$f_2(y_{it}|x_{it},\beta_2;y_{it}>0) = \frac{\Gamma(\alpha^{-1}+y_{it})}{\Gamma(\alpha^{-1})(\Gamma(y_{it}+1))((1+\alpha\mu_2)^{\alpha^{-1}}-1)}\left(\frac{\mu_2}{\mu_2+\alpha^{-1}}\right)^{y_{it}} \tag{7}$$

where μ_2 =exp$(x'_{it}\beta_2)$.

While the x_{it}'s capture a lot of observable heterogeneity between individuals across time (especially the health status, insurance status, age, sex, and education), there might

[10] Here, α is set to one because it would not be identified from β_1 in this binary information case (see Bago d'Uva, 2006)

still be great a deal of unobservable heterogeneity left. This could be general unmeasured frailty or preferences towards visiting a doctor. These factors clearly affect the demand for health care and can be considered time-invariant. In order to account for this unobserved heterogeneity, we use the latent class hurdle model derived by Bago d'Uva (2006) for panel count data in which the time-invariant individual effect follows an arbitrary discrete distribution that takes on a small number of masspoints. The latent class hurdle model combines the basic hurdle model (that groups individuals into "users" and "non-users") with a finite mixture model (where the latent classes can be given interpretations such as "high users" and "low users" (Deb and Trivedi, 2002)).

As in Clark and Etile (2006), the individual probability of belonging to one of C latent classes is specified as a multinomial logit.

$$\pi_{ij} = \frac{\exp(z_i'\gamma_j)}{\sum_{g=1}^{C}\exp(z_i'\gamma_g)}, j = 1, \ldots, C \tag{8}$$

This ensures that $0 < \pi_{ij} < 1$ and $\sum_{j=1}^{C}\pi_{ij} = 1$. In order to guarantee that each individual belongs to the same latent class over all time periods, we choose the z_i as time-invariant characteristics. We follow Bago d'Uva (2005) in using the individual averages of the x_{it}, defining $z_i = \bar{x}_i$. Note that this specification allows for correlation between the observable characteristics and the unobserved individual heterogeneity (that is, here, the latent class).

The likelihood function is finally given by

$$L = \prod_{i=1}^{N}\sum_{j=1}^{C}\pi_{ij}\prod_{t=1}^{T_i} g_j(y_{it}|x_{it}, \theta_j) \tag{9}$$

where $\theta_j = (\beta_{1j}, \beta_{2j}, \alpha_j)$ and equation (1), (5), (6), and (7) are plugged into equation (8). The most flexible formulation allows for different slope parameters in every latent class ($\beta_{1j} \neq \beta_{1k}$ and $\beta_{2j} \neq \beta_{2k}$ for $j \neq k$) and different parameters in the two hurdle parts ($\beta_{1j} \neq \beta_{2j}$). That is, belonging to a certain latent class does not only alter the intercept but is allowed to affect each slope parameter. This, however, requires estimation of many parameters. For instance, a fully flexible hurdle model with three latent classes and, say, 20 regressors and a constant would include 171 coefficients[11] that have to be estimated. This flexible specification is very data-demanding. Because we only have three waves and, furthermore, carry out different regressions for publicly and voluntarily insured, we restrict the model to the same slope parameters across latent classes and allow only for intercept heterogeneity. This still requires estimation of 91 different parameters in the

[11] These are three classes times two sets of parameters for the first and second hurdle part times 21 parameters (incl. constant) plus three different over-dispersion parameters plus two times 21 parameters for the probability of belonging to a certain latent class.

above example.[12] Thus, the use of a finite mixture model is motivated from a statistical point of view in this case, namely the possibility to introduce a random-effect without imposing too strong distributional assumptions, instead of trying to find different effects of observable variables for different latent classes. The likelihood function is maximized with respect to the vectors $\theta_1, \ldots, \theta_C, \gamma_1, \ldots, \gamma_{C-1}$ using the Broyden-Fletcher-Goldfarb-Shanno quasi-Newton algorithm.[13] Note, finally, that in this specification, the two parts of the hurdle model are not assumed to be independent. This restrictive assumption is relaxed due to the latent class specification.

4 Estimation Results

The latent class panel model captures unobserved individual effects. Based on the Akaike criterion (AIC), it outperforms a standard hurdle model with a logit as the first part and a truncated negbin as a second part (results not shown here). As with the bivariate probit, we carry out separate regressions for both groups of insured (public and voluntary insurance). According to the AIC, models with four latent classes dominate models with two respectively three classes in both cases.[14] Due to the nonlinearity of the model, the interpretation of the estimated parameters is somewhat difficult. Here we focus on the calculated marginal effects of the most interesting variables in order to interpret the results (table 3). The full regression results can be found in tables 6, 7, and 8 in the appendix.

Before turning to the coefficients related to the insurance state, we briefly summarize the effects of other important variables, particularly the newly available variables. All three variables indicating the health status show a highly significant effect on the demand for health care. Both higher PCS and higher MCS (indicating a healthier individual) lead to less doctor visits, where the effect of physical health is stronger than the one of mental health. However, even after controlling for both more comprehensive health measures, most of the dummy variables indicating self-assessed general health status remain significant. A step from one category to a better one increases the expected number of doctor visits, everything else being constant. Most of the variables are highly significant in both stages and, thus, turn out to have the most important impact on the demand for health services.

[12] $2\times(20+3)+1+2\times21=91$. It turned out, that even allowing for different over-dispersion parameters lead to problems with maximising the likelihood function. Restricting the over-dispersion to one parameter made the maximisation much more feasible. Furthermore, to increase the degrees of freedom, we implicitly assume that males and females differ only by a constant term. However, fitting different regressions for males and females did not lead to different results concerning the insurance state variables in earlier versions of the paper.

[13] We used the *ml* command in Stata and drawed heavily on the code provided in Jones et al. (2006, ch. 11). Different starting values were used to rule out local maxima of the likelihood function.

[14] A model with five latent classes did not converge in either case.

Table 3: Marginal effects in latent class hurdle model

	Voluntary Insured				Publicly Insured			
	$P(Y > 0 \mid X)$		$E[Y \mid Y > 0, X]$		$P(Y > 0 \mid X)$		$E[Y \mid Y > 0, X]$	
Deductible	0.007	0.016	-0.037	0.096				
Private	-0.032	0.020	-0.097	0.103				
Addon					0.030**	0.010	0.185**	0.074
PCS	-0.007***	0.001	-0.050***	0.006	-0.007***	0.001	-0.061***	0.003
MCS	-0.003***	0.001	-0.031***	0.004	-0.001***	0.000	-0.026***	0.002
Risk attitude health	0.001	0.003	-0.051***	0.017	-0.001	0.002	-0.002	0.013
Worries Health	-0.040***	0.011	-0.116	0.067	-0.036***	0.005	-0.322***	0.031
SAH very good	-0.010	0.027	-0.560***	0.130	-0.079***	0.014	-0.526***	0.079
SAH good	0.001	0.015	-0.265***	0.091	-0.042***	0.007	-0.272***	0.045
SAH bad or poor	0.065***	0.023	0.394***	0.129	0.037***	0.011	0.286***	0.051
BMI high	0.005	0.014	0.049	0.085	-0.001	0.007	0.004	0.045
BMI very high	0.073***	0.021	0.276	0.146	-0.004	0.010	-0.032	0.062

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. Marginal effects of continuous variables are calculated numerically, all marginal effects evaluated at sample means.

The body-mass index dummies are not significant (except for *BMI very high* in the first stage in the group of the publicly insured). It can be argued that the effect of the BMI is already well captured by the previously discussed measures. The self-assessed risk attitude towards health matters has a significant influence on the number of doctor visits in only in the second stage in group of publicly insured. As could be expected, more risk-averse individuals (with a lower value of risk-aversion on the 11-point scale) have a higher number of doctor visits.[15] Finally, worries about the own health are associated with more doctor visits. Being very concerned about the own health especially influences the decision to see a doctor in the first stage. Thus, this variable captures not only more information about the true health status but also behavioural differences between individuals.

The insurance status variables also show a clear picture. Holding add-on insurance raises the probability of visiting a doctor (the probability of visiting a doctor is 3%-points higher for an individual with add-on insurance than for one without). The marginal effect of *deductible* is about zero. Holding private insurance decreases the probability of doctor visits, however, it is not significantly different from zero.[16] Thus, it can be concluded that only add-on insurance holders show the expected incentive effect. Conditional on the health status and risk preferences, holding private insurance with a deductible does not seem to lower the probability of visiting a doctor.

Given that the patient has visited the doctor at least once in the previous three months, the expected number of doctor visits shows a similar behaviour. The marginal effect

[15] Without controlling for the health status, the effect of risk-aversion concerning health matters could be positive, meaning that risk-averse individuals care more for their health and, thus, need less doctor visits. Controlling for the health status, however, this variable captures the pure effect of risk-aversion on the decision of visiting a doctor.

[16] *Deductible* and *Private* are also not jointly significant in the first stage.

of *Deductible* in the second stage is again very small and not statistically significant. Conditional on at least one visit, the expected number of visits in the last three months is 0.037 lower for individuals with a deductible. A similar result holds for private insurance, although the marginal effect is somewhat higher in absolute terms.

Instead, add-on insurance even increases the number of doctor visits in the second stage (with a marginal effect of 0.185). This (as well as the results on risk-aversion) shows that the second stage is not entirely determined by the physician's behaviour but also by the patient.

The results indicate the absence of a causal effect of deductibles on the demand for health care. Hence, the lower number of doctor visits of the insured with a deductible is instead a sign of self-selection rather than of moral hazard. In contrast, the strong effect of add-on insurance, even after controlling for health status, observable preferences, and unobserved heterogeneity, indicates an incentive effect of health care consumers.

5 Conclusion

Finding out the price elasticity of demand for health care is an important empirical task. A low elasticity (i.e., here, absence of different behaviour caused by different insurance contracts) in a world with risk-averse individuals implies an insurance system with full cover insurance as a best solution. Dealing with moral hazard, however, less than full cover insurance could help to contain inefficiently high demand for health care.

Empirically, one can compare the behaviour of individuals with full cover insurance with the one of individuals with less than full cover insurance in order to find out the elasticity. However, if the choice of deductibles is voluntary (and deductibles are not randomly assigned), there will be self-selection into different insurance schemes. In Germany, the insured with deductibles have a lower number of doctor visits than insured without deductibles. However, using a set of newly available variables in the GSOEP (especially more comprehensive health measures and measures of risk-aversion) and a latent class panel model that accounts for unobserved heterogeneity, no causal effect of the deductibles was found.

The results might be interpreted as absence of moral hazard. However, the highly regulated German health care system only allows patients to make own decisions in rare occasions (for instance, whether they want to make a doctor visit or not). Not to find a significant effect here does not necessarily mean that individuals do not show moral hazard behaviour in general. Furthermore, the results do not necessarily imply that deductibles as such are not suitable to contain health care costs in Germany in general.

First, especially individuals with a high income are allowed to chose a deductible. It is by all means possible that the price elasticity in this special group is not representative for all households in the population. The significant positive impact of private add-on insurance on the number of doctor visits is an indication of this point because this insurance type can be purchased by the whole population. Second, and most important, deductibles which are not voluntary but mandatory (as, e.g., in the Swiss health system) might have a strong effect on the demand for health care.

Finally, although the estimated effects for add-on insurance are statistically significant, they are rather small in magnitude, compared to the baseline effect (the different constant for the different latent classes) and other effects like health, gender, and socio-economic characteristics (see the full regression results in table 6). Thus, this study generally confirms the notion that the German population only responds slowly and weakly on financial incentives given by the health care system.

References

[1] Augurzky, B., T. Bauer and S. Schaffner (2006), “Copayments in the German Health System. Does It Work?” RWI Discussion Paper No. 43, Essen.

[2] Bago d’Uva, T. (2005): “Latent class models for use of primary care: evidence from a British panel”, Health Economics 15, 873-892.

[3] Bago d’Uva, T. (2006): “Latent class models for health care utilisation.”, Health Economics 15, 329-343.

[4] Bound, J. (1991): Self-reported versus objective measures of health in retirement models. Journal of Human Resources 26, 106-138.

[5] Breyer, F., P. Zweifel und M. Kifmann (2004): “Gesundheitsökonomik”. 5. Auflage. Springer Verlag.

[6] Buchner F, Wasem J., (2003): Needs for further improvement: risk adjustment in the German health insurance system. Health Policy 65, 21-35.

[7] Cameron, A. C., P. K. Trivedi, Microeconometrics.

[8] Clark A, F. Etilé (2006): “Don’t give up on me baby: spousal correlation in smoking behaviour”, Journal of Health Economics 25, 958-978.

[9] Crossley, T. F. and S. Kennedy (2002): The reliability of self-assessed health status, Journal of Health Economics 21, 643-658.

[10] Deb P., PK. Trivedi, (2002): “The structure of demand for health care: latent class versus two-part models”, Journal of Health Economics 21, 601-625.

[11] Dohmen T., A .Falk, D. Huffman, U. Sunde, J. Schupp and IZA Bonn G. G. Wagner (2005): "Individual Risk Attitudes: New Evidence from a Large, Representative, Experimentally-Validated Survey", IZA Discussion Paper 1730, September.

[12] Geil, P., A. Million, R. Rotte, K. F. Zimmermann (1997), “Economic Incentives and Hospitalization in Germany”, Journal of Applied Econometrics, Vol. 12, No. 3, Special Issue: Econometric Models of Event Counts. (May - Jun., 1997), pp. 295-311.

[13] Grabka, M. (2006): “Prämien in der PKV: Deutlich stärkerer Anstieg als in der gesetzlichen Krankenversicherung”, DIW Wochenbericht Nr. 46/2006.

[14] Haisken-DeNew, John P. and Markus Hahn (2006) “PanelWhiz: A Flexible Modularized Stata Interface for Accessing Large Scale Panel Data Sets”, mimeo, http://www.panelwhiz.eu

[15] Jones, A., N. Rice, T. Bago d'Uva and S. Balia (2007): Applied Health Economics.

[16] Jürges, H. (2007):"Health Insurance Status and physician-induced demand for medical services in Germany: new evidence from combined district and individual level data", SOEPpapers on Multidisciplinary Panel Data Research No. 8, Berlin

[17] Knapp, L.G. and Seaks, T.G. (1998):"A Hausman test for a dummy variable in probit", Applied Economics Letters 5, 321-323

[18] Mullahy, J. (1986): "Specification and Testing of Some Modified Count Data Models." Journal of Econometrics 33, 341-65.

[19] Nübling M., H. H. Andersen, A. Mühlbacher, J. Schupp, and G. G. Wagner (2007): Computation of Standard Values for Physical and Mental Health Scale Scores Using the SOEP Version of SF12v2. Schmollers Jahrbuch, Vol. 127(1),

[20] Pohlmeier, W. and V. Ulrich (1995), "An econometric model of the two-part decision-making process in the demand for medical care", Journal of Human Resources, 30, 339-61.

[21] Riphahn, R.T., 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.

[22] Schellhorn M., (2001): "The effect of variable health insurance deductibles on the demand for physician visits", Health Economics 10, 441-456.

[23] Wagner, G. G., J. R. Frick, and J. Schupp (2007): "The German Socio- Economic Panel study (SOEP) - evolution, scope and enhancements". Journal of Applied Social Science (Schmollers Jahrbuch) 127 (1), 139-169

[24] Windmeijer, F. A. G. and J. M. C. Santos Silva (1997): "Endogeneity in Count Data Models: An Application to Demand for Health Care" Journal of Applied Econometrics, Vol. 12, No. 3, Special Issue: Econometric Models of Event Counts. (May - Jun., 1997), pp. 281-294.

[25] Zweifel, P. und W. G. Manning (2000): "Moral Hazard and Consumer Incentives in Health Care", in: Anthony J. Culyer and Joseph P. Newhouse, Handbook of Health Economics, Vol. 1A, 409-459. North-Holland.

Appendix

Table 4: Sample means by subgroups

	SHI without add-on	SHI with add-on	Voluntary with deductible	Voluntary without deductible
PCS	48.34	50.43	52.22	51.47
MCS	49.74	50.34	51.68	50.49
SAH very good	0.05	0.07	0.11	0.10
SAH good	0.37	0.43	0.47	0.46
SAH bad or poor	0.18	0.12	0.10	0.11
BMI high	0.38	0.34	0.38	0.38
BMI very high	0.17	0.15	0.10	0.12
Hospital Stays t-1	0.17	0.15	0.10	0.12
Degree of Handicap	7.96	5.38	3.5	4.98
Smoker	0.28	0.31	0.28	0.27
Worries Health	2.06	2.20	2.31	2.25
Risk attitude health	2.74	3.16	3.46	3.27
Female	0.54	0.56	0.33	0.36
Age	51.1	47.20	49.52	49.09
Foreign	0.07	0.02	0.05	0.02
Married	0.68	0.70	0.70	0.73
Children under 16	0.30	0.34	0.34	0.34
West Germany	0.72	0.86	0.87	0.85
Full-time employed	0.37	0.50	0.68	0.62
Self-employed	0.03	0.07	0.40	0.20
Blue collar worker	0.18	0.15	0.01	0.05
White collar worker	0.30	0.49	0.34	0.43
Health job	0.02	0.02	0.05	0.03
Net-household inc./1000	2.64	3.49	5.19	4.13
Net-labour inc./1000	0.76	1.23	2.38	1.86
Years of schooling	11.6	12.59	14.32	13.85
2002	0.34	0.25	0.31	0.36
2004	0.36	0.31	0.36	0.34
Observations	38034	3311	2818	6801

Source: GSOEP, pooled years 2002, 2004, 2006; individuals older than 25, no civil servants.

Table 5: Bivariate probit results

	Voluntary Insured				Public Insured			
	Doctor Binary		Deductible		Doctor Binary		Addon	
Deductible	-0.125	(0.30)						
Private	-0.060	(0.04)						
Addon					0.248	(0.21)		
PCS	-0.024***	(0.00)	0.006	(0.00)	-0.023***	(0.00)	-0.003	(0.00)
MCS	-0.007***	(0.00)	0.007**	(0.00)	-0.005***	(0.00)	0.002	(0.00)
SAH very good	-0.198**	(0.07)	-0.003	(0.07)	-0.247***	(0.04)	0.007	(0.05)
SAH good	-0.093*	(0.04)	-0.028	(0.04)	-0.124***	(0.02)	-0.019	(0.03)
SAH poor or bad	0.162*	(0.07)	0.165*	(0.07)	0.192***	(0.03)	-0.084*	(0.04)
BMI high	-0.024	(0.03)	-0.012	(0.04)	0.026	(0.02)	-0.004	(0.03)
BMI very high	0.065	(0.05)	-0.056	(0.06)	0.042	(0.02)	0.043	(0.04)
Hospital Stays t-1	0.301***	(0.05)	0.013	(0.03)	0.343***	(0.03)	0.034*	(0.01)
Degree Disability	0.008***	(0.00)	-0.002	(0.00)	0.006***	(0.00)	-0.000	(0.00)
Smoker	-0.188***	(0.04)	0.036	(0.04)	-0.202***	(0.02)	0.056	(0.03)
Worries health	-0.200***	(0.03)	-0.002	(0.03)	-0.167***	(0.01)	0.022	(0.02)
Risk attitude health	-0.013	(0.01)	-0.004	(0.01)	-0.003	(0.00)	0.008	(0.01)
Female	0.246***	(0.04)	-0.010	(0.05)	0.283***	(0.02)	0.147***	(0.03)
Age	-0.037***	(0.01)	0.033**	(0.01)	-0.029***	(0.00)	0.015*	(0.01)
Age squared	0.000***	(0.00)	-0.000**	(0.00)	0.000***	(0.00)	-0.000**	(0.00)
Foreign	-0.092	(0.08)	0.205*	(0.09)	0.072*	(0.03)	-0.445***	(0.06)
Married	0.078	(0.04)	-0.161**	(0.05)	0.030	(0.02)	0.029	(0.03)
Children under 16	-0.100**	(0.04)	0.038	(0.05)	-0.068**	(0.02)	0.001	(0.03)
West Germany	0.091	(0.05)	0.124*	(0.06)	-0.081***	(0.02)	0.413***	(0.03)
Full-time employed	-0.230***	(0.06)	0.106	(0.07)	-0.076**	(0.03)	0.054	(0.04)
Self-employed	0.027	(0.08)	0.353***	(0.08)	-0.165***	(0.04)	0.204**	(0.07)
Blue collar worker	-0.041	(0.10)	-0.454***	(0.13)	-0.104***	(0.03)	-0.011	(0.05)
White collar worker	0.229**	(0.07)	-0.221**	(0.08)	0.043	(0.03)	0.140***	(0.04)
Health job	-0.336***	(0.08)	0.059	(0.09)	-0.184***	(0.05)	-0.003	(0.07)
Net household inc./1000	0.007	(0.01)	0.026***	(0.01)	0.005	(0.01)	0.039***	(0.01)
Net labour inc./1000	-0.002	(0.01)	-0.014	(0.01)	0.009	(0.01)	0.035	(0.02)
Years of schooling	0.020***	(0.01)	0.009	(0.01)	0.032***	(0.00)	0.035***	(0.01)
2002	0.058	(0.04)	-0.126***	(0.03)	0.080***	(0.02)	-0.336***	(0.02)
2004	0.012	(0.03)	-0.032	(0.02)	0.054**	(0.02)	-0.240***	(0.02)
Risk attitude finance			0.024*	(0.01)			0.025***	(0.01)
Attitdue cost-sharing			-0.173***	(0.02)			-0.084***	(0.02)
Constant	2.959***	(0.35)	-1.878***	(0.37)	2.370***	(0.16)	-2.316***	(0.24)
ρ	0.050	(0.18)			-0.053	(0.11)		
Log-pseudolikelihood	-10574.217				-30738.545			
Observations	9246				39698			

*p<0.05, **p<0.01, ***p<0.001; Standard errors in parentheses

Estimations done by Stata program biprobit, standard errors clustered by individuals, less observations than in table 4 because of missing values in the instruments.

Table 6: Estimation results of latent class hurdle model

	Voluntary Insured				Publicly Insured			
	$\hat{\beta}_{stage1}$		$\hat{\beta}_{stage2}$		$\hat{\beta}_{stage1}$		$\hat{\beta}_{stage2}$	
Deductible	0.044	(0.10)	-0.021	(0.06)				
Private	-0.204	(0.13)	-0.055	(0.06)				
Addon					0.192**	(0.07)	0.088**	(0.03)
PCS	-0.046***	(0.01)	-0.029***	(0.00)	-0.041***	(0.00)	-0.030***	(0.00)
MCS	-0.018***	(0.01)	-0.018***	(0.00)	-0.009***	(0.00)	-0.013***	(0.00)
SAH very good	-0.063	(0.17)	-0.372***	(0.10)	-0.470***	(0.08)	-0.297***	(0.05)
SAH good	0.002	(0.09)	-0.151**	(0.05)	-0.265***	(0.05)	-0.136***	(0.02)
SAH bad or poor	0.427**	(0.16)	0.205***	(0.06)	0.239***	(0.07)	0.134***	(0.02)
BMI high	0.034	(0.09)	0.028	(0.05)	-0.008	(0.04)	0.002	(0.02)
BMI very high	0.487**	(0.15)	0.148*	(0.07)	-0.028	(0.06)	-0.016	(0.03)
Hospital Stays t-1	0.574***	(0.11)	0.174***	(0.03)	0.456***	(0.05)	0.088***	(0.01)
Degree of Handicap	0.015***	(0.00)	0.002*	(0.00)	0.007***	(0.00)	-0.000	(0.00)
Smoker	-0.146	(0.11)	0.056	(0.06)	-0.284***	(0.05)	-0.136***	(0.03)
Risk attitude health	0.002	(0.02)	-0.029**	(0.01)	-0.226***	(0.03)	-0.159***	(0.02)
Worries Health	-0.252***	(0.07)	-0.066	(0.04)	-0.004	(0.01)	-0.001	(0.01)
Female	0.743***	(0.15)	0.176**	(0.07)	0.330***	(0.06)	-0.045	(0.03)
Age	-0.079**	(0.03)	0.001	(0.01)	-0.026	(0.01)	0.002	(0.01)
Age squared	0.001**	(0.00)	-0.000	(0.00)	0.000*	(0.00)	-0.000	(0.00)
Foreign	0.078	(0.25)	-1.160***	(0.20)	0.142	(0.10)	0.119	(0.07)
Married	0.215	(0.12)	-0.105	(0.06)	0.047	(0.05)	0.041	(0.03)
Children under 16	-0.001	(0.12)	0.063	(0.07)	-0.071	(0.05)	-0.072*	(0.03)
West Germany	0.436*	(0.19)	0.389***	(0.08)	0.102	(0.06)	0.073*	(0.03)
Full-time employed	-0.325	(0.17)	-0.111	(0.08)	-0.001	(0.07)	0.024	(0.04)
Self-employed	0.096	(0.18)	0.060	(0.09)	-0.145	(0.11)	-0.016	(0.06)
Blue collar worker	0.431	(0.28)	-0.151	(0.14)	-0.073	(0.07)	-0.154***	(0.04)
White collar worker	0.123	(0.21)	-0.099	(0.10)	0.059	(0.07)	-0.084*	(0.04)
Health job	0.155	(0.22)	0.078	(0.13)	-0.033	(0.10)	0.039	(0.06)
Net-household inc./1000	0.017	(0.02)	0.011*	(0.01)	-0.022*	(0.01)	0.005	(0.00)
Net-labour inc./1000	-0.019	(0.02)	-0.023	(0.01)	0.012	(0.03)	-0.081***	(0.02)
Years of schooling	-0.011	(0.02)	-0.002	(0.01)	0.060***	(0.01)	0.031***	(0.01)
2002	0.080	(0.07)	-0.009	(0.04)	0.048	(0.04)	0.104***	(0.02)
2004	-0.046	(0.07)	-0.062	(0.04)	0.056	(0.03)	-0.003	(0.01)
Constant (Laten class 1)	3.704***	(1.00)	2.657***	(0.47)	2.553***	(0.44)	2.205***	(0.24)
Constant (Laten class 2)	5.056***	(1.04)	4.572***	(0.46)	3.777***	(0.44)	4.287***	(0.23)
Constant (Laten class 3)	6.600***	(0.99)	3.348***	(0.46)	6.101***	(0.53)	2.569***	(0.24)
Constant (Laten class 4)	4.973***	(0.96)	2.345***	(0.47)	4.293***	(0.44)	3.255***	(0.23)
Alpha			0.434***	(0.03)			0.328***	(0.01)
Log-pseudolikelihood	-16766.063				-77597.446			
Akaike	33876.126				156765.660		.	
Observations	9619				141345			

Standard errors in parentheses; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. The hypothesis of equality of all parameters across the two stages could be rejected in favor of the hurdle model.

Table 7: Probabilities of latent class membership - Deductible

	π_1		π_2		π_3	
	coefficient	std. error	coefficient	std. error	coefficient	std. error
Deductible	-0.378	(0.52)	-0.162	(0.63)	-0.837	(0.43)
Private	0.802	(0.52)	1.267*	(0.59)	1.007*	(0.42)
PCS	-0.138**	(0.05)	-0.109*	(0.05)	-0.108**	(0.04)
MCS	-0.028	(0.03)	0.011	(0.03)	-0.009	(0.02)
SAH very good	4.020***	(1.13)	3.739**	(1.15)	1.604*	(0.79)
SAH good	1.539*	(0.69)	2.058**	(0.80)	0.612	(0.45)
SAH bad or poor	0.961	(0.86)	1.092	(0.92)	0.451	(0.65)
BMI high	0.347	(0.39)	0.150	(0.47)	-0.004	(0.31)
BMI very high	0.931	(0.54)	-0.574	(0.69)	-0.292	(0.43)
Hospital Stays in t-1	3.358*	(1.35)	3.130*	(1.37)	3.079*	(1.36)
Degree Disability	-0.024	(0.01)	0.004	(0.01)	-0.014	(0.01)
Smoker	0.684	(0.49)	-1.278*	(0.62)	-0.262	(0.35)
Risk attitude health	0.119	(0.08)	0.212**	(0.08)	0.039	(0.06)
Worries health	0.043	(0.41)	-0.398	(0.42)	-0.429	(0.29)
Female	2.018***	(0.54)	0.624	(0.62)	0.873	(0.49)
Age	0.062	(0.12)	0.096	(0.14)	0.004	(0.11)
Age squared	0.000	(0.00)	-0.001	(0.00)	0.000	(0.00)
Foreign	0.076	(1.67)	4.072***	(1.11)	-1.074	(1.79)
Married	0.982*	(0.42)	1.124*	(0.53)	0.517	(0.36)
Children under 16	0.103	(0.57)	-0.728	(0.54)	-0.520	(0.42)
West Germany	-0.425	(0.51)	-1.422*	(0.66)	-0.960	(0.53)
Full-time employed	0.603	(0.74)	0.780	(0.80)	-0.049	(0.54)
Self-employed	-1.155	(0.85)	-1.619	(0.93)	-1.154	(0.66)
Blue collar worker	2.208	(1.57)	0.470	(1.35)	0.550	(1.20)
White collar worker	-1.330	(1.00)	-0.509	(0.96)	-0.061	(0.71)
Health job	2.720*	(1.20)	-0.502	(1.47)	-0.452	(1.04)
Net household inc./1000	-0.086	(0.11)	-0.079	(0.05)	-0.077	(0.05)
Net labour inc./1000	-0.122	(0.17)	0.057	(0.09)	0.023	(0.08)
Years of schooling	0.103	(0.06)	0.025	(0.08)	0.187**	(0.06)
2002	-0.235	(0.73)	0.705	(0.69)	0.217	(0.57)
2004	-0.215	(0.83)	-0.110	(0.81)	0.189	(0.58)
Constant	0.971	(5.49)	-0.194	(4.29)	5.065	(3.15)

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Coefficients have to be interpreted relative to the base categrory

Table 8: Probabilities of latent class membership - Addon

	π_1		π_2		π_3	
	coefficient	std. error	coefficient	std. error	coefficient	std. error
Addon	-0.252	(0.18)	0.312	(0.27)	-0.059	(0.28)
PCS	-0.008	(0.01)	-0.029	(0.02)	-0.032*	(0.01)
MCS	-0.007	(0.01)	-0.003	(0.01)	-0.017*	(0.01)
SAH very good	-0.124	(0.26)	1.127**	(0.43)	-1.584*	(0.76)
SAH good	0.060	(0.15)	0.347	(0.28)	0.239	(0.21)
SAH poor or bad	-1.050***	(0.20)	-1.053**	(0.34)	-0.659**	(0.25)
BMI high	-0.081	(0.11)	-0.077	(0.18)	0.051	(0.16)
BMI very high	-0.114	(0.16)	0.290	(0.20)	0.160	(0.20)
Hospital Stays in t-1	-1.988***	(0.20)	0.073	(0.07)	-0.791***	(0.15)
Degree Disability	-0.016***	(0.00)	0.008*	(0.00)	-0.008**	(0.00)
Smoker	0.016	(0.12)	0.248	(0.16)	-0.623***	(0.18)
Worries health	-0.018	(0.10)	-0.073	(0.18)	-0.576***	(0.14)
Risk attitude health	0.006	(0.03)	-0.025	(0.04)	0.003	(0.03)
Female	-0.735***	(0.15)	-0.012	(0.18)	-0.502**	(0.17)
Age	0.081**	(0.03)	-0.022	(0.04)	0.124***	(0.04)
Age squared	-0.001*	(0.00)	0.000	(0.00)	-0.001*	(0.00)
Foreign	0.272	(0.27)	-0.027	(0.25)	0.513	(0.31)
Married	0.175	(0.13)	0.313	(0.19)	0.271	(0.16)
Children under 16	-0.019	(0.13)	0.198	(0.20)	-0.351	(0.23)
West Germany	0.019	(0.15)	0.190	(0.19)	-0.813***	(0.17)
Full-time employed	0.303	(0.18)	0.045	(0.28)	-0.065	(0.30)
Self-employed	0.341	(0.31)	0.279	(0.49)	0.029	(0.44)
Blue collar worker	-0.051	(0.20)	0.089	(0.29)	-0.612	(0.33)
White collar worker	-0.396*	(0.19)	-0.267	(0.32)	-0.604*	(0.30)
Health job	0.971**	(0.30)	0.695	(0.51)	-0.342	(0.69)
Net household inc./1000	-0.051	(0.04)	-0.068	(0.06)	0.102**	(0.04)
Net labour inc./1000	-0.136	(0.10)	0.229	(0.13)	-0.209	(0.15)
Years of schooling	0.038	(0.03)	-0.093*	(0.04)	0.101**	(0.03)
2002	-0.248	(0.25)	-0.562	(0.48)	0.344	(0.39)
2004	0.157	(0.26)	-0.393	(0.51)	-0.321	(0.40)
Constant	-0.631	(1.06)	1.293	(1.51)	-1.212	(1.41)

* $p<0.05$, ** $p<0.01$, *** $p<0.001$

Coefficients have to be interpreted relative to the base category