CES ifo

Working Papers

Heterogeneous Effects of a Nonlinear Price Schedule for Outpatient Care

Helmut Farbmacher
Peter Ihle
Ingrid Schubert
Joachim Winter
Amelie Wuppermann

CESifo

Center for Economic Studies & Ifo Institute

Heterogeneous Effects of a Nonlinear Price Schedule for Outpatient Care

Abstract

Theoretical considerations suggest that nonlinear health care price schedules have heterogeneous effects on health care demand. In this paper, we develop and apply a finite mixture bivariate probit model to analyze whether there are heterogeneous reactions to the introduction of a nonlinear price schedule in the German statutory health insurance system. In administrative insurance claims data from the largest German health insurance plan, we find that some individuals strongly react to the new price schedule while a second group of individuals does not react. Post-estimation analyses reveal that the group of the individuals who do not react to the reform includes the relatively sick. These results are in line with forward-looking behavior: Individuals who are already sick expect that they will hit the kink in the price schedule and thus are less sensitive to the co-payment.

Helmut Farbmacher
Munich Center for the Economics of
Aging (MEA) / Max Planck Society
Amalienstraße 33
Germany – 80799 Munich
farbmacher@mea.mpisoc.mpg.de

Peter Ihle
University Hospital of Cologne
PMV forschungsgruppe
Herderstraße 52
Germany – 50931 Cologne
peter.ihle@uk-koeln.de

Ingrid Schubert
University Hospital of Cologne
PMV forschungsgruppe
Herderstraße 52
Germany – 50931 Cologne
ingrid.schubert@uk-koeln.de

Joachim Winter
University of Munich
Department of Economics
Ludwigstraße 33
Germany – 80539 Munich
winter@lmu.de

Amelie Wuppermann (corresponding author)
University of Munich
Department of Economics
Ludwigstraße 33
Germany – 80539 Munich
amelie.wuppermann@lmu.de

November 21, 2013

1   Introduction

Nonlinear price schedules are a common feature of many health insurance systems. Non-linearities often arise due to deductibles or combinations of co-payments and maximum out-of-pocket amounts. Economic theory predicts that these nonlinearities lead to heterogeneity in the effect of out-of-pocket expenditures on health care demand even among individuals with identical preferences. In the presence of a nonlinear price schedule, heterogeneity could arise as the out-of-pocket price is not the relevant price for forward-looking individuals. Instead, the relevant price depends on individuals' expectations of their future health care use. For example, in a price schedule in which costs drop to zero once out-of-pocket expenditures exceed a certain amount, individuals who expect that their out-of-pocket expenditures will exceed the maximum amount have little incentive to reduce current health care use. They expect to pay the same total amount independent of their health care use today. Conversely, individuals who expect that their expenditures will not exceed the maximum amount face the full copayment for their current health care use. For forward-looking individuals, the sensitivity with respect to the current price of care is thus predicted to depend on expected future health care use, which naturally varies between individuals (Keeler *et al.*, 1977; Ellis, 1986). An average measure of price sensitivity to copayments in a nonlinear price schedule might thus hide substantial heterogeneity. A similar argument has been made by Bitler *et al.* (2006) concerning the effects of welfare reforms on labor market outcomes.

In this paper, we develop a finite mixture model to analyze the changes in health care demand induced by an introduction of a nonlinearity in the price schedule in the German statutory health insurance (SHI). Our model adapts and extends the model of Atella *et al.* (2004) to accommodate a bivariate binary outcome (indicators for visits to general practitioners and specialists) and correlated error terms in a panel data setting. The finite mixture model allows us to estimate the sensitivity to changes in copayments for different classes of individuals without having to specify *a priori* which individual belongs to which class. We determine the number of different classes according to the normalized entropy criterion (NEC), a criterion which ensures that the classes are well-separated and nonoverlapping. This criterion is rarely applied in economics although it provides a clear clustering structure, which enhances interpretation. For each latent class and combination of the bivariate binary outcomes, we calculate average marginal effects and their standard errors. We further calculate posterior probabilities of class memberships, based on which we assign individuals to latent classes. We then examine the composition of these classes

to describe the sources of heterogeneity.

We employ this approach to analyze the effects of the introduction of a nonlinearity in the price schedule of the German SHI. The SHI is the public health insurance system in Germany that is mandatory for most employees and covers around 90% of the German population. The specific reform that we analyze introduced a per-quarter fee for doctor visits. Before 2004, the publicly insured in Germany did not have to copay for doctor visits. Since 2004, they have to pay a fee of € 10 for the first visit to a doctor in each calendar quarter. Additional visits in the same quarter are free of charge. The price an individual pays thus drops from € 10 to € 0 after the first doctor visit in a quarter. Only individuals who do not visit any doctor in a quarter can avoid paying the fee.

In order to detect whether the reform affected individual behavior, we focus on the effect of the per-quarter fee on the probability of not visiting any doctor (henceforth, "no doctor visit") in a quarter. If individuals are forward looking, their reactions to the German per-quarter fee should depend on their expected probability of having to visit a doctor. For individuals who expect that they will likely have to visit a doctor within the next three months, behavior should not change. Other individuals, however, might expect that they can avoid going to the doctor through the entire quarter and thus avoid paying the fee. For these individuals, the probability of no doctor visit should increase. We therefore expect that there will be at least two groups with different reactions to the newly introduced per-quarter fee. Standard approaches to the evaluation of health reforms which focus on average effects (conditional on observable covariates) would combine these heterogeneous effects and might yield misleading policy conclusions.[1]

Our analysis is based on administrative claims data from the largest German health insurance plan for the year 2002 to 2005. These data allow us to reliably observe doctor visits within quarters and are thus particularly suited for our purposes. However, they only contain information on health through diagnosis codes or claims for prescribed drugs. The latter two measures are, of course, only observed if individuals have visited a doctor. If the introduction of the copayment affected the probability of doctor visits, it would also

---

[1]Earlier studies of this reform have abstracted from potentially heterogeneous treatment effects (Augurzky *et al.*, 2006; Schreyögg and Grabka, 2010). Farbmacher and Winter (2013) address specific measurement problems and nonlinearities that arise when using survey data to study the effects of a quarterly copayment, but restrict the analysis to homogeneous effects. Rückert *et al.* (2008) stratify their results by observable characteristics and find that younger and healthier adults and individuals with low income are more likely to report that they have delayed doctor visits. Winkelmann (2006) studies the effects of a different reform of the German SHI using a quantile count data model; he documents heterogeneous reactions of the number of doctor visits to the reform, a result that is in line with the findings we report below.

affect these health measures which are thus endogenous to the reform. But health itself could be an important source of heterogeneity in the reform effects, for instance as current health status might explain individuals' expectations of their future health care use. Our finite mixture approach allows us to omit the plausibly endogenous health measures from our main analysis, while still allowing for health-related heterogeneity through the model's latent classes. We use the health measures in a post-estimation analysis later on.

With this paper, we add to the literature on the price sensitivity of health care demand. Many studies in this literature have focused on the impact of the current out-of-pocket price on health care demand without taking the possible impact of nonlinearities in the price schedule into account (see Aron-Dine *et al.* (2012) for a discussion and Einav *et al.* (2013) for a recent study that explicitly investigates incentives introduced by nonlinearities in health insurance plans). A challenge in accounting for the impact of nonlinear price schedules on the price sensitivity of demand is that individual expectations and thus the relevant prices are typically unobservable to the researcher. Some studies have solved this problem by constructing individuals' relevant or expected prices based on observable sociodemographic characteristics, prior use of health care services, or experiences of individuals in the same health plan (Ellis, 1986; Contoyannis *et al.*, 2005; Meyerhoefer and Zuvekas, 2010; Aron-Dine *et al.*, 2012). However, modeling individual expectations is difficult and in practice requires strong and implausible assumptions, including homogeneity of the expectations process (Manski, 2004). Our finite mixture approach has the advantage that in the spirit of a reduced-form model, we can account for the nonlinearity in the price schedule, without having to predict relevant prices and specify expectations. Furthermore, we do not need to make assumptions on whether and to what extent individuals are forward looking. This is important as evidence by Aron-Dine *et al.* (2012) suggests that individuals are neither fully myopic nor fully forward looking.

The finite mixture model we develop also fits into a recent literature on modeling heterogeneity in health care decisions and in the effects of reforms on individual behavior more generally. Examples include Deb and Trivedi (1997), Deb and Trivedi (2002), Atella *et al.* (2004), Bago d'Uva (2006), Deb *et al.* (2011) and Ayyagari *et al.* (2013) who use finite mixture models to analyze heterogeneity in health-related decisions and outcomes, Winkelmann (2006) who develops a quantile model for count data to analyze heterogeneity in the effect of a different copayment in the SHI, and Djebbari and Smith (2008) who extend and apply quantile treatment effects and random coefficient models to test for heterogeneity of the effects of a Mexican welfare program.

Our results indicate that the average probability of no doctor visit significantly increased after the reform, by about 3.5 percentage points. This result suggests that the reform has had an effect on the demand for doctor visits on average. However, the average effect only delivers the correct price response under the assumption that individuals are completely myopic and ignore the impact that their demand has on future prices. Results of our finite mixture model show that there are heterogeneous reactions to the introduction of the new copayment. We estimate models with two, three and four latent classes. According to the NEC, the model with two latent classes is the most appropriate one. This model shows that for 36% of individuals the probability of no doctor visit increased by around 7 percentage points, while there was no significant change for the remaining individuals. Post-estimation analyses show that the individuals who react to the reform are the relatively healthy. To the extent that healthy individuals have lower expected future health care needs, these results are in line with the theoretical predictions of Keeler *et al.* (1977). However, we cannot fully exclude that the results reflect other sources of unobserved heterogeneity, such as heterogeneity in price sensitivity with respect to health or income or heterogeneity in the degree of myopia. We return to this issue in the concluding section.

The paper is structured as follows: The next section describes the health care reform in more detail. Section 3 introduces the dataset and Section 4 explains our estimation strategy. In Section 5 the results are presented and Section 6 concludes.

2   Incentive Effects of the Reform

The health care reform that we analyze became effective in January 2004. The reform implemented various financial incentives in the SHI with the aim of increasing patients' cost consciousness. Figure 1 shows the changes in Germany's consumer price index for medical care. The reform appears to have permanently increased the prices for medical care. According to this index, it has been the largest health care reform in Germany for more than a decade. The central and most radical element of the reform was the introduction of a per-quarter fee for doctor visits. While patients did not have to copay for doctor visits before the reform, since 2004 they pay € 10 for the first visit to a physician in each quarter of the year. Further visits to the same doctor within a quarter are free of charge. Visits in the same quarter to other doctors are also exempt from the fee if patients get a referral by the doctor whom they visited first. Alternatively, patients can visit other doctors without referral and pay the fee again. As paying the fee can only be avoided by not visiting any physician within a quarter, we expect that the introduction of

4

the per-quarter fee has increased the probability of no doctor visit in a quarter.

Figure 1: Germany's Consumer Price Index for Medical Care
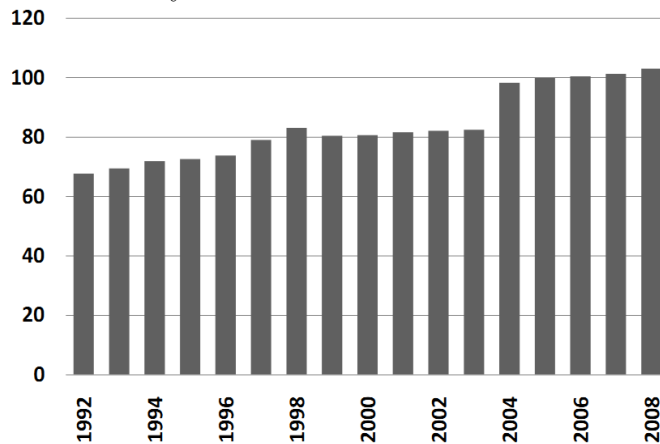
The theoretical models of health care demand by Keeler *et al.* (1977) and Ellis (1986) suggest that this effect might be heterogeneous across individuals as the per-quarter fee has introduced a nonlinearity in the price schedule. The reason for this is its implicit deductible. The fee has to be paid only at the first visit in a calendar quarter. Hence, given a referral the patient's price for doctor visits drops to zero after the first visit. This nonlinearity generates varying incentives depending on individuals' health status. Keeler *et al.* (1977), for instance, show that under uncertainty a rational individual facing a deductible will not base decisions on nominal prices. The authors instead argue that "the greater the chance that future expenditures will exceed the deductible, the cheaper today's visit to the doctor". A rational and forward-looking individual will thus anticipate that the price drops to zero at a certain consumption level. In the German case, the intertemporal effect on prices is relatively easy to assess. It only depends on individuals' expectations on their probability of visiting a physician in the next three months. In the extreme case, their behavior is unaffected by the per-quarter fee once they know that they have to visit a doctor in a certain calendar quarter (for example to get a new prescription for a drug for a chronic disease). The effective price for doctor visits is thus lower for individuals with chronic conditions. Hence, if demand for outpatient care depends on effective prices, we expect a weaker reaction of demand among individuals with high risk of doctor visits than among low-risk individuals.[2]

---

[2]In this argument we assume that individuals continue to visit doctors in case of major conditions and take their drugs despite the increase in copayments. This assumption can be justified by the low copayment

An additional effect of the per-quarter fee on individual behavior could be that individuals try to fit as many foreseeable doctor visits as possible into a quarter once they have paid the fee. This would allow them to avoid having to pay the fee in a later quarter. Since we are mainly interested in whether the per-quarter fee affected the demand for doctor visits at all, and whether the effects are in line with theoretical predictions, we focus on the effect on the probability of no doctor visit in this study. Consequently, our results do not pertain to the reform's effects on the overall number of doctor visits and on health care costs.

3   Data

The analysis is based on insurance claims data from the largest German health insurance plan in the years 2002 to 2005. The data contains information on a 18.75% random subsample of all individuals in the German state of Hesse who are insured with this health plan. At the beginning of each year, a refreshment sample is drawn in order to keep the sample representative of the insured population.[3]

The data contains information on all doctor visits. Specifically, for all visits we observe the type of doctor visited, the diagnoses made (classified at the ICD-10 level), and the drugs prescribed. As we are interested in the reactions to the introduction of the per-quarter fee, we aggregate the data from the level of the doctor visit to the quarterly level (within individual). Furthermore, we group information on the different doctor visits into those to general practitioners (GPs) and those to specialists. The dependent variables are thus the number of GP visits per quarter and the number of specialist visits per quarter, at the level of the insured individual.

The main advantage of using claims data is that doctor visits are reliably observed (whereas in survey data they are based on respondent self-reports which are typically rather noisy). A drawback of claims data, however, is that the only information on an individual's health is that contained in the diagnoses codes and the list of prescribed drugs. This information is only available for individuals who have seen a doctor. Other than that, only information on the age and sex of the individual is available. The claims data thus comprise only very few variables that do not depend on whether an individual

---

level. Generally, individuals in Germany will not get into severe financial troubles due to out-of-pocket expenditures.

[3]See *http://www.pmvforschungsgruppe.de/content/02_forschung/02_b_sekundaerd_1.htm* for a short description of the data in German.

has visited a doctor.

Another disadvantage of this claims dataset is that it consists only of publicly insured individuals. It does not contain any adults who were not affected by the reform (primarily the privately insured) who could serve as a control group. Within the SHI, individuals younger than 18 are generally exempt from paying the per-quarter fee. These individuals, however, are not suitable as a control group for the entire adult population.[4] We thus exclude all individuals younger than 19 from our analysis and revert to before-after comparisons to identify the effects of the reform in the adult population. Our results therefore rely on the assumption that in the absence of the reform, no considerable changes in health care use would have occurred. We shed light on the plausibility of this assumption by comparing changes across the two years prior to the reform.

Our sample is further restricted to those individual-quarter observations for which the entire quarter is covered. Individual-quarter observations are excluded, for example, if the individual switches from or to a different insurer within the quarter. This sample restriction ensures that the length of the period at risk is the same for each observation.

Table 1 shows descriptive statistics for the third quarter of each year.[5] The average age is almost unchanged over time, reflecting the inclusion of refreshment samples. The average number of doctor visits in our sample is around 4.5 per quarter. On average, individuals visit a GP a little less than once a month and a specialist 1.7 times per quarter. This level of physician service use is high in international comparison, but it is in line with statistics on doctor visits in Germany available from other sources (see Grobe *et al.*, 2010).

While the average number of doctor visits per quarter does not show a clear change after the reform, two possible effects of the increased copayments become evident in table 1. Between 2003 and 2004, the fraction of individuals with at least one GP visit and the fraction with at least one specialist visit in the third quarter both decline, from 63% to 60% for GPs and from 46% to 42% for specialists. Individuals thus seem to reduce contacts with either type of doctor after the reform.

As the copayment can only be avoided by seeing neither type of physician, we are particularly interested in how the probability of no doctor visit within a quarter changed after the reform. Descriptive statistics on these quantities are reported at the bottom of table 1. While in the third quarter of 2002 and 2003, roughly 27% of the sample visit neither a GP nor a specialist, this is the case for 33% of individuals in the years after the

---

[4]This feature of the reform suggests a natural division into treatment and control group among teenagers. We will conduct a difference-in-difference analysis for teenagers in a follow-up study.

[5]The descriptive statistics are very similar for the other quarters of the year.

Table 1: Descriptive Statistics

|	3Q 2002		3Q 2003		3Q 2004		3Q 2005	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Age	51.87	18.71	52.16	18.75	52.31	18.81	52.17	18.88
19-39	0.31	0.46	0.30	0.46	0.29	0.46	0.29	0.46
40-59	0.31	0.46	0.31	0.46	0.32	0.47	0.33	0.47
60-79	0.31	0.46	0.31	0.46	0.31	0.46	0.30	0.46
≥ 80	0.07	0.26	0.08	0.27	0.08	0.27	0.08	0.27
Female	0.52	0.50	0.52	0.50	0.52	0.50	0.52	0.50
# GP visits	2.77	4.12	2.75	4.11	2.77	4.16	2.85	4.33
GP>0	0.63	0.48	0.63	0.48	0.60	0.49	0.59	0.49
# GP visits truncated at 0	4.43	4.45	4.40	4.45	4.65	4.51	4.83	4.72
# Specialist visits	1.75	3.65	1.70	3.49	1.63	3.57	1.71	3.73
Specialist>0	0.46	0.50	0.46	0.50	0.42	0.49	0.42	0.49
# Specialist visits truncated at 0	3.83	4.60	3.70	4.37	3.89	4.64	4.05	4.85
GP= 0 & Specialist = 0	0.27	0.44	0.27	0.44	0.33	0.47	0.33	0.47
N	256,071		249,851		246,379		248,328	

reform. The average probability of no doctor visit per quarter thus increases by about 6 percentage points after the reform.

The change in the probability of no doctor visit in 2004 compared to 2003 is also depicted in figure 2. This figure shows the changes in the third quarter of 2004 compared to the third quarter of 2003 separately for men and women in different age groups and with different health status. The health status is captured by the Charlson Index (Charlson *et al.*, 1987). This index is based on 17 diseases identified from the diagnoses codes available in our dataset. Each disease is assigned a weight between 1 and 6, depending on its severity.[6] The Charlson Index is the sum of these weights, truncated at 2. A value of 0 thus indicates that an individual had no diagnosis of any of the Charlson conditions, and a value of 1 or 2 indicates the presence of one or more severe co-morbidities.

As the Charlson Index is based on diagnoses codes which are only available if an individual has seen a doctor, the Charlson Index might itself have been affected by the introduction of the copayment. In order to mitigate this problem, we construct the Charlson Index for individuals in the third quarter of each year based on their diagnoses in the two prior quarters. For example, the "Charlson 0" group in 2003 contains all individuals who had no diagnosis of any Charlson condition in the first two quarters of 2003. Since individuals with conditions as severe as those counted in the Charlson Index will likely

---

[6]A list of the diseases and corresponding weights is displayed in Appendix A.

Figure 2: Change in the Probability of No Doctor Visit in 2004 compared to 2003 by Age and Charlson Index

(a) Men

(b) Women
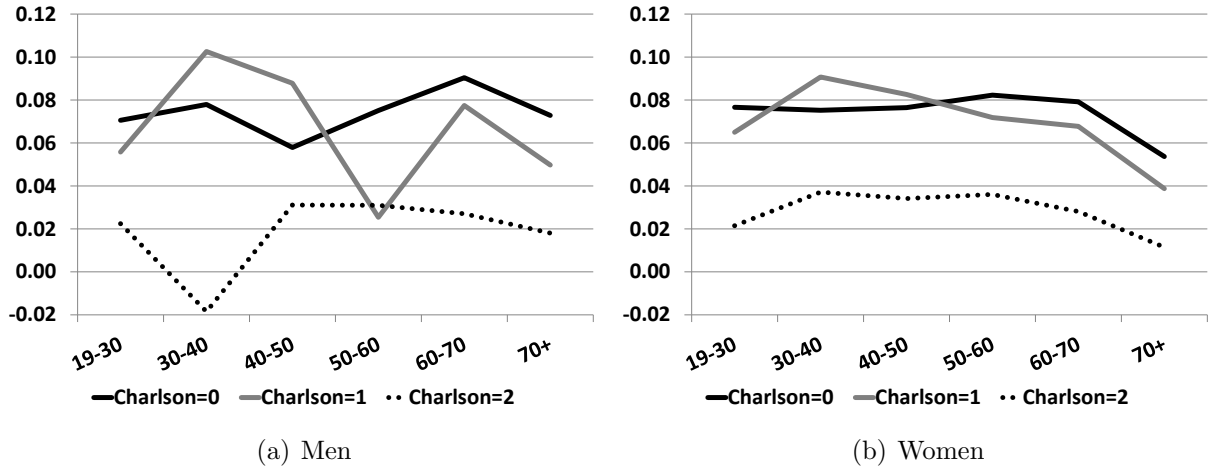
have to visit a doctor at some point during six months, we hope that the index based on diagnoses in 6 months should be affected less by the reform than a measure based on one quarter alone. Furthermore, by focusing on the quarters prior to the quarter for which we describe the change in the probability to visit a doctor, we eliminate the direct link between doctor visits in the quarter and our health measure.

Figure 2 presents first descriptive evidence of heterogeneous effects across the different groups. The probability of no doctor visit generally increases after the reform with similar magnitude for both genders. These increases are much smaller for individuals with a Charlson Index of 2 than for the other groups. Sicker individuals thus seem to react less to the reform than healthier ones. This is true for all age groups and both genders.

Overall, figure 2 indicates that the change in the probability of no doctor visit is stronger for healthy individuals. These descriptive results, however, rely on an endogenous measure of health. In order to test our hypothesis of heterogeneous reactions without having to rely on the health information in the data, we use a finite mixture model. This model allows us to estimate different effects for separate groups in the population without having to explicitly stratify the data by observable characteristics *a priori*.

4   Econometric Framework

In this section we develop a finite mixture model that accommodates a bivariate binary outcome and correlated error terms in a panel data setting. Our data consists of a panel of

individuals observed over time and with outcomes related to two different medical special-ties. Individuals can seek care from GPs (y_1) and/or specialists (y_2). Suppose there are J types (or classes) of individuals ($j = 1, 2, ..., J$) who may react differently to the changing financial incentives. The number of different types J is determined by the data according to different selection criteria that are further explained at the end of this section. The object of interest is the change over time in the probability of visiting a doctor,

$$Pr(y_k|t = a, group = j) - Pr(y_k|t = b, group = j) \equiv \zeta_{kj} \qquad k = 1, 2 \qquad (1)$$

where j denotes the type, k refers to GP and specialists, respectively, and t is an indicator for years before (b) and after (a) the reform. In our application heterogeneous reform effects can, for instance, be caused by varying price sensitivities and/or differences in the perception of the price change. The latter is particularly important as some individuals may not react to the financial incentive at all since the deductible element of the reform makes them consider the change as an increase in insurance premium. So even if their price sensitivity is negative, we might not see changes in the demand for health care in this group.

The overall change in the probability of visiting a doctor for the population can be estimated from the group-specific reform effects as a weighted average across all groups

$$Pr(y_k|t = a) - Pr(y_k|t = b) = \sum_{j=1}^{J} Pr(group = j)\zeta_{kj}. \qquad (2)$$

Reporting the overall effect is informative if the entire population is affected by the finan-cial incentive. However, considering a case with a (possibly) large group that does not react to the reform at all, reporting the overall effect may hide even strong reactions in smaller groups. Equation (2) suggests the use of a finite mixture model to explore these heterogeneities. In the following we give a formal description of the model and generalize the basic idea to the bivariate case incorporating correlated error terms and also additional control variables.

The dataset is an unbalanced panel, and each individual i is observed in T_i quarters. Over time and across the two physician specialities, the dataset thus includes $2 \cdot T_i$ obser-vations for individual i. Suppose that individual i belongs to a latent class j for the entire observational period. The probability of belonging to class j is π_j. Within a latent class, we use bivariate probits to jointly model the decision to visit a GP and/or a specialist. While we are mostly interested in whether individuals visit any doctor within a quarter, indepen-dent of the type of doctor visited, using information on two decisions (GP and specialist

visits) in the estimation of the model should improve the classification of individuals into latent classes. The joint probability of the dependent variables over the observed period is the product of T_i independent probabilities, given fixed class membership, i.e.,

$$Pr(y_{1i}, y_{2i}|x_i, \theta_j) = \prod_{t=1}^{T_i} \Phi_2[(2y_{1it} - 1)x_{it}\beta_j, (2y_{2it} - 1)x_{it}\gamma_j, (2y_{1it} - 1)(2y_{2it} - 1)\rho_j] \quad (3)$$

where x_i denotes the vector of covariates that includes age, sex, seasonal fixed effects and year fixed effects. $\Phi_2(\cdot)$ denotes the bivariate normal cdf and θ_j contains the vector of parameters for GP visits (β_j), the vector of parameters for specialist visits (γ_j), and the parameter ρ_j which captures covariation of the errors in the underlying structural model.

The log-likelihood function that corresponds to this model is given by

$$ln(L) = \sum_{i=1}^{I} ln\left(\sum_{j=1}^{J} \pi_j Pr(y_{1i}, y_{2i}|x_i, \theta_j)\right) \quad (4)$$

where I is the number of individuals in the dataset and J is the number of latent classes.[7] In order to obtain a dataset of manageable size, we use a 3% random subsample in this part of our analysis. This gives us 120,521 observations from a little more that 7,500 individuals per quarter.[8]

In order to facilitate the interpretation of our results, we calculate the average marginal effects of the variables in x_i on $Pr(y_{1i} = 1|x_i, \theta_j)$ and $Pr(y_{2i} = 1|x_i, \theta_j)$. Standard errors of the marginal effects are calculated using the delta method. As the reform effect is captured by changes in the probability of no doctor visit, we also calculate marginal effects on the probability of no doctor visit, i.e. on $Pr(y_{1i} = 0, y_{2i} = 0|x_i, \theta_j)$.[9]

Furthermore, using Bayes' theorem we calculate posterior probabilities of membership in the different latent classes for each individual i as

$$Pr(y_{1i}, y_{2i} \in h|y_{1i}, y_{2i}, x_i, \widehat{\theta}) = \frac{\widehat{\pi}_h Pr(y_{1i}, y_{2i}|x_i, \widehat{\theta}_h)}{\sum_{j=1}^{J} \widehat{\pi}_j Pr(y_{1i}, y_{2i}|x_i, \widehat{\theta}_j)} \quad (5)$$

where $Pr(y_{1i}, y_{2i}|x_i, \widehat{\theta}_h)$ is defined as in equation (3). The posterior probabilities on the one hand help to judge how well the different latent classes are separated by the estimation, which is a basic objective of mixture models. On the other hand, one can assign individuals

---

[7]The likelihood function is maximized directly using the Newton-Raphson algorithm. First and second derivatives are calculated numerically in Stata's optimization package.

[8]A comparison of this subsample with the entire sample is available upon request.

[9]See appendix B for details.

to a specific latent class based on their posterior probabilities and then characterize each latent class using observable characteristics. In addition to the variables age and sex that are included in x_i, we use the health information contained in the data in this post-estimation analysis. As this information is not included in x_i, it is external to the estimation and can therefore be used to verify the classification.

So far, we have not specified J, i.e. the number of latent classes in the finite mixture model. For model selection, information criteria such as the well-known Akaike information criterion (AIC) and the Bayesian information criterion (BIC) are often used. Both criteria emphasize the goodness of model fit, but differ in their way how they take the parsimony of the model into account. However, these criteria do not reflect the model's ability to generate well-separated clusters – an important objective of mixture models. We therefore use a third criterion for model selection, the normalized entropy criterion (NEC). This criterion was proposed by Celeux and Soromenho (1996) to assess the optimal number of clusters in a mixture model. It determines the overlap of the mixture components; the smaller the overlap between different components, the better the separation between the latent classes. The NEC is defined as

$$NEC(J) = \frac{\sum_{j=1}^{J} \sum_{i=1}^{I} t_{ij} ln(t_{ij})}{L(J) - L(1)}$$

where t_{ij} represents the posterior probability of class membership shown in equation (5), and $L(\cdot)$ is the likelihood function displayed in equation (4), which depends on the number of latent classes J. The entropy term in the numerator measures the overlap of the mixture components. A low value of the NEC therefore points to a model with well-separated clusters.

5   Results

Table 2 reports the NEC, AIC and BIC criteria for the finite mixture bivariate probit model described in the last section with different number of latent classes (LC). In addition to finite mixture models (FMM) with 2, 3, and 4 latent classes, we estimate a standard one-component bivariate probit model.[10] The AIC and BIC displayed in table 2 indicate that the model with 4 latent classes fits the data best. The NEC criterion, on the other hand, prefers the model with just 2 latent classes. As the criteria come up with conflicting results, we need additional guidance to choose between the models. While NEC indicates that the

---

[10]Estimations with more than 4 latent classes failed to converge and are likely overparameterized.

Figure 3: Posterior probabilities for different model specifications

(a) FMM with 2 latent classes   (b) FMM with 4 latent classes
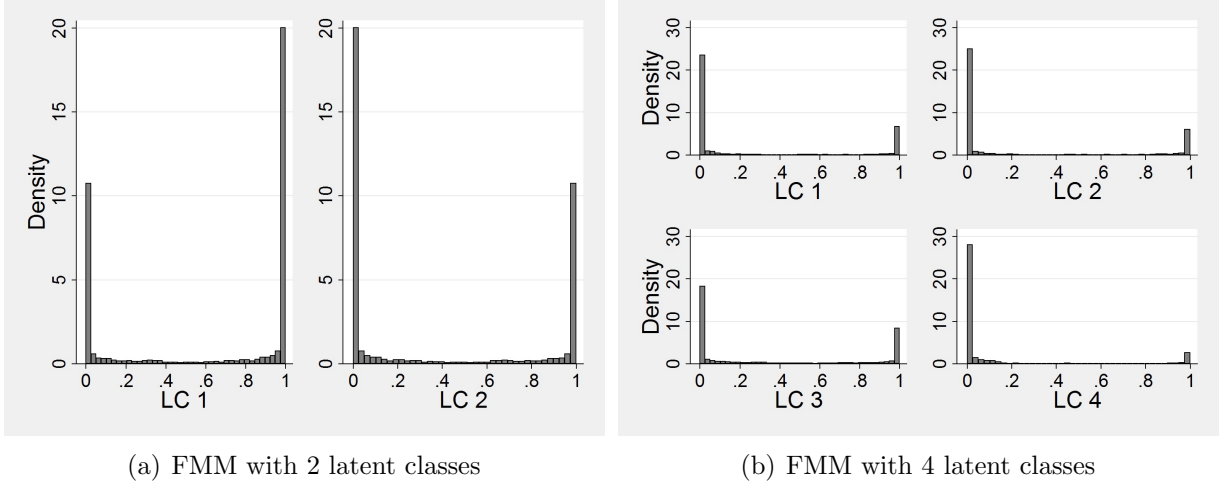
model with 2 latent classes results in the optimal number of clusters, figure 3 indicates that both models seem to do reasonably well in separating the sample into classes based on posterior probabilities. This figure plots the densities of the posterior probabilities of class membership for the models with 2 and 4 latent classes. For both models, each probability has large spikes at 0 and 1, indicating that individuals can be easily assigned to latent classes based on these probabilities.

Another aspect that is typically considered in the choice of the number of latent classes in finite mixture models is interpretability of the different groups. Cameron and Trivedi (2005), p. 624, argue that "[t]he most desirable situation is one in which the components have an interpretation." Finite mixture models in health economics, therefore, often contain only 2 latent classes, which are interpreted as the group of *low* or *infrequent* users and the group of *high* users of health care (see, for instance Deb and Trivedi, 2002; Bago d'Uva, 2006). For the sake of interpretation, we therefore focus on the results of the model with 2 latent classes, which is also favored by NEC. The results from the model with 4 classes do not differ qualitatively from the ones of the model with 2 latent classes and are presented in appendix C.

Marginal effects and their standard errors based on the results of our finite mixture bivariate probit model with 2 latent classes are reported in table 3. In addition to the marginal effects in the different latent classes, table 3 displays *overall effects*. The latter are derived as weighted averages of the effects in the 2 latent classes; the estimated prior probabilities of class membership, $\widehat{\pi}$ and $1 - \widehat{\pi}$, serve as weights. For comparison, the

13

Table 2: Model Selection

Model	N	LogL	df	AIC	BIC	NEC
Bivariate Probit	120,521	-148,116.6	25	296,283	296,526	–
2 LC FMM Bivariate Probit	120,521	-130,342.6	51	260,787	261,282	**0.034**
3 LC FMM Bivariate Probit	120,521	-122,594.1	77	245,342	246,089	0.050
4 LC FMM Bivariate Probit	120,521	-118,899.9	103	**238,006**	**239,005**	0.058

last column of table 3 reports marginal effects and standard errors based on the standard bivariate probit model.

The overall marginal effects and the marginal effects of the standard bivariate probit model have the same signs and are similar in magnitude. Women have a higher probability of visiting a GP and of visiting a specialist at least once in a quarter than men. The probability of visiting a GP or a specialist at least once increases with age, particularly so for individuals aged 40 to 60. The probabilities of visiting either type of doctor at least once are lower in the summer months (quarter 2 and 3) compared to the first quarter of a year. Overall, the marginal effects thus show expected signs.

Of particular interest are the changes in the probabilities of doctor visits over the years. The effects of the year dummies capture these changes, compared to the reference year 2002. Overall, the results of the finite mixture bivariate probit model and the standard bivariate probit concordantly show significant reductions in the probabilities in the years 2004 and 2005 compared to 2002. In 2003, however, there is no significant difference to 2002. As 2003 is a pre-reform year, the absence of an effect in 2003 supports the assumption that there would have been no changes in the outcome variable in the absence of the reform.

The first two columns of table 3 present the reform effects for the two latent classes. The prior probability of belonging to class 1 is estimated as 62.5%. In this relatively large class, the changes in the probability of either doctor visit in the post-reform years are much smaller than in the other class. The results in table 3 thus suggest that the reform had an effect on the demand for doctor visits and that this effect might be heterogeneous across individuals. However, recall that the per-quarter fee has to be paid at the first visit to a doctor in a quarter, independent of the type of doctor visited, while all additional visits to other doctors are free of charge. We thus present an analysis of the change in the probability of not visiting any doctor. The marginal effects on $Pr(y_{1i} = 0, y_{2i} = 0|x_i, \theta_j)$ are displayed in figure 4.

The overall effect displayed in figure 4 is not significant in the pre-reform year 2003

Table 3: FMM Bivariate Probit – Marginal Effects

	LC 1	LC 2	Overall	BiProbit
GP				
Female	0.011	0.013	0.012	0.067***
	(0.009)	(0.014)	(0.010)	(0.007)
Age Splines				
Age 19-40	0.005***	-0.001	0.003***	0.002***
	(0.001)	(0.002)	(0.001)	(0.001)
Age 40-60	0.003*	0.005	0.004**	0.007***
	(0.001)	(0.003)	(0.001)	(0.001)
Age 60-80	-0.002	-0.002	-0.002	0.000
	(0.001)	(0.003)	(0.002)	(0.001)
Age > 80	-0.001	0.005	0.002	-0.008*
	(0.003)	(0.006)	(0.004)	(0.003)
Quarter Dummies				
Quarter 2	-0.006	-0.018***	-0.011***	-0.011***
	(0.004)	(0.005)	(0.003)	(0.003)
Quarter 3	-0.013***	-0.032***	-0.020***	-0.020***
	(0.004)	(0.006)	(0.003)	(0.003)
Quarter 4	0.018***	0.003	0.012***	0.010***
	(0.004)	(0.006)	(0.003)	(0.003)
Year Dummies				
2003	0.011	-0.005	0.005	0.001
	(0.004)	(0.006)	(0.003)	(0.003)
2004	-0.005**	-0.052***	-0.023***	-0.027***
	(0.005)	(0.008)	(0.004)	(0.004)
2005	-0.015**	-0.027**	-0.020***	-0.027***
	(0.005)	(0.009)	(0.004)	(0.004)
Specialist				
Female	0.131***	0.112***	0.124***	0.164***
	(0.011)	(0.023)	(0.010)	(0.007)
Age Splines				
Age 19-40	-0.002	0.003*	0.000	-0.001*
	(0.001)	(0.001)	(0.001)	(0.001)
Age 40-60	0.007***	0.000	0.004**	0.008***
	(0.002)	(0.002)	(0.001)	(0.001)
Age 60-80	-0.008***	0.000	-0.005***	-0.006***
	(0.002)	(0.003)	(0.001)	(0.001)
Age > 80	-0.009*	-0.014*	-0.011***	-0.014***
	(0.003)	(0.006)	(0.003)	(0.003)
Quarter Dummies				
Quarter 2	-0.004	-0.002	-0.003	-0.004
	(0.004)	(0.005)	(0.003)	(0.003)
Quarter 3	-0.009*	-0.012*	-0.010**	-0.010**
	(0.004)	(0.005)	(0.003)	(0.003)
Quarter 4	-0.002	-0.003	-0.003	-0.003
	(0.004)	(0.005)	(0.003)	(0.003)
Year Dummies				
2003	0.016***	-0.016**	0.004	0.002
	(0.005)	(0.006)	(0.004)	(0.004)
2004	-0.017**	-0.063***	-0.034***	-0.037***
	(0.006)	(0.007)	(0.004)	(0.004)
2005	-0.010	-0.060***	-0.029***	-0.033***
	(0.006)	(0.007)	(0.004)	(0.004)
ρ	0.278***	0.364***		0.489***
	(0.013)	(0.015)		(0.008)
π_j	0.625			

Notes: π_j is the probability of class membership in latent class (LC) j. Standard errors for the marginal effects in parentheses. They are calculated using the delta method. Standard errors of the underlying coefficients are clustered at the individual level. *p<0.05; **p<0.01; ***p<0.001

Figure 4: Changes in the Probability of No Doctor Visit
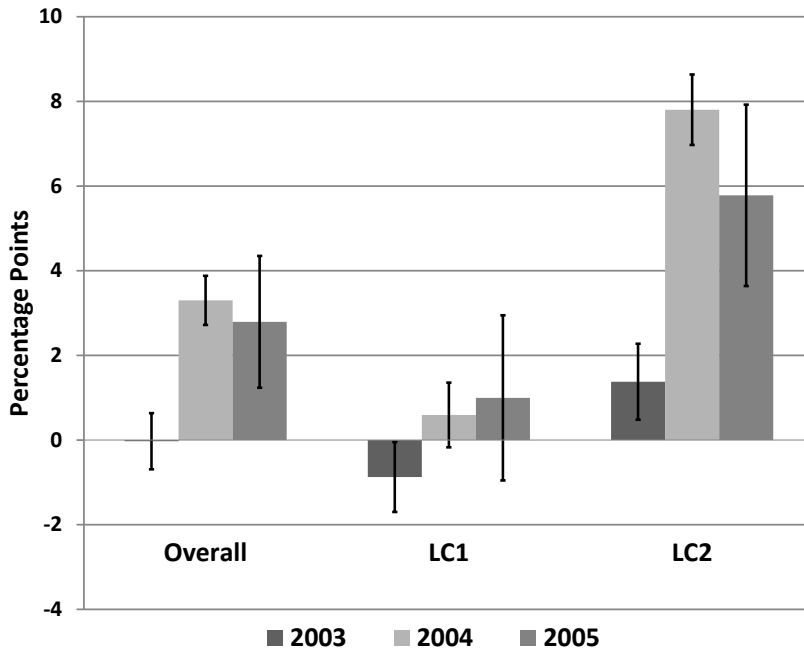
compared to 2002, but large and highly significant in the post-reform years 2004 and 2005. This indicates that the probability of no doctor visit increased by about 3.5 percentage points after the reform. Furthermore, figure 4 presents evidence of heterogeneous reform effects: While individuals in latent class 1 did not react to the reform, there are strong reactions in latent class 2. In the latter class, the probability of no doctor visit already increased by around 1 percentage point in 2003. The increases in the post-reform years 2004 and 2005, however, are a lot larger. They lie between 6 and 8 percentage points. The results thus suggest that there is a group of individuals who changed their behavior after the reform, while the behavior in the other group did not change.

Naturally, the question arises who the individuals are that belong to the different latent classes. Given the parameter estimates, we derive the posterior probabilities for each individual i and assign each individual to the latent class with the highest posterior probability. Table 4 reports averages of observed characteristics for the different latent classes. Besides age and sex, table 4 includes information on current health, as summarized by the Charlson Index, and on prescription drugs that individuals got from GPs and specialists. The latter are measured in defined daily doses (DDDs).[11] These give a rough measure of

---

[11]DDDs are defined by the World Health Organization, see

Table 4: Comparison of Group Means

	LC 1	LC 2
Age	54.88	47.82
Female	0.58	0.41
Charlson Index	0.94	0.20
DDD	182.46	22.41
Exempt from co-payments	0.15	0.02
% of individuals	63.70	36.30

Notes: Fractions of Charlson Index do not add
up to 1 because of missing observations.
DDD stands for defined daily dose.

drug consumption within the different classes, adjusted for the fact that different drugs are of different potency. The DDDs do not take into account, however, which amount of the drugs were actually prescribed.

The last row of table 4 shows that about 64% of individuals were assigned to latent class 1 and 36% to latent class 2. Comparing the different latent classes, it becomes evident that latent class 1 is on average older than latent class 2 and has a larger share of women. In addition, there are differences concerning the Charlson Index and the DDDs. Individuals in latent class 1 have a higher average Charlson Index and higher average DDD. Individuals in latent class 1 thus have more severe diagnoses and take more potent drugs on average, indicating that these individuals are sicker than individuals in the other latent classes. This supports the hypotheses that among the individuals who do not react to the reform are the relatively sick.

Additionally, table 4 shows the fraction of observations in each latent class that is exempt from copayments. Before and after the reform in 2004, individuals could apply for an exemption from copayments if the amount of their copayments exceeded 2% of their gross yearly income. Welfare recipients and chronically sick individuals could be entirely exempt from copayments before the reform. Since 2004, however, the 2% rule also applies to welfare recipients. The chronically sick can only be exempt from further payments within a year once they have already copaid 1% of their gross yearly income. Individuals can apply for these exemptions already at the beginning of each year. In order to do so, they have to pay either 1% or 2% of their gross annual income up front. The information on whether individuals are exempt from copayments is only available after the reform. The results in table 4 thus only include observations in the post-reform period.[12]

http://www.whocc.no/ddd/definition_and_general_considera/ for a definition.

[12]For all variables that are available before the reform, the results look almost identical in the pre-reform

Given the exemption rules described above, one could argue that individuals who know *ex ante* that they will be exempt from all copayments do not react to the introduction of the per-quarter fee. In line with this argument, table 4 shows that latent class 1 includes a larger fraction of observations that are exempt from copayments. However, 85% of observations in latent class 1 are not exempt from copayments and still do not react to the introduction of the per-quarter fee. The nonlinearity of the per-quarter fee thus results in heterogeneous reactions not only through the general exemption rules.

Overall, the results from the finite mixture model confirm that the introduction of the per-quarter fee in the SHI had heterogeneous effects: Around 64% of individuals are assigned *ex post* to the class that does not react to the reform, while there are a strong reactions among the rest. Consistent with the theory of Keeler *et al.* (1977), the results indicate that among the individuals who do not react to the introduction of the per-quarter fee are the ones who are sick and thus likely expect that they cannot avoid paying the fee.

6   Conclusion

In this paper, we present an empirical analysis of the heterogeneous effects of a nonlinear price schedule that was introduced in the SHI. The nonlinearity takes the form of a copayment for doctor visits that only has to be paid for the first visit in each quarter of the year. The price for doctor visits individuals face in the remainder of a quarter drops to zero once the fee has been paid. Theoretical models of health care demand in the presence of nonlinear price schedules suggest that the per-quarter fee changes the probability of no doctor visit differently across individuals. In particular, individuals who expect that they will need to visit a physician at some point within the quarter have a lower incentive to change their demand for doctor visits than individuals who expect that no visit is necessary. As the expectations with respect to the necessity of visits within a given quarter likely depend on individuals' health status, we expect that the reform effects vary between individuals with good and bad health.

In a descriptive analysis, we find that individuals in worse health react less to the reform than healthier individuals, independent of their age. Our measure of health, however, depends on the outcome variable – namely, it is observed only if individuals visit a doctor. We therefore omit this information from our econometric model and instead allow for unobserved heterogeneity by using a finite mixture model. The results of this model show

---

period.

that some individuals react to the reform while others do not. Characterizing the different groups indicates that those individuals who do not react are in worse health. Our results are thus in line with the theoretical predictions of how nonlinear price schedules affect the demand for medical care.

The finite mixture model does not allow to determine the exact source of heterogeneity. In addition to heterogeneity in the relevance of price changes that are due to forward-looking behavior, differences in health care needs could be another source of heterogeneity, as some individuals might not be able to avoid visiting a doctor due to their health status. Furthermore, heterogeneity in the price sensitivity might also arise due to unobserved variation in income or due to preference heterogeneity, such as differences in the degree of myopia with some individuals taking future considerations into account in their decision making and others not. As the intertemporal consequences of the German copayment are salient and relatively easy to assess, however, it is likely that at least some of the heterogeneity that we document reflects forward-looking behavior.

Our findings allow two conclusions. First, the per-quarter fee affects the demand for doctor visits on average, indicating that this demand is sensitive to prices. Second, and more importantly, the effect on demand is heterogeneous across individuals and varies with individuals' health status. Analyses of the average effect will miss these important heterogeneities, making it impossible to asses the consequences of reforms for specific, vulnerable groups.

References

Aron-Dine A, Einav L, Finkelstein A, Cullen M. 2012. Moral hazard in health insurance: How important is forward looking behavior? NBER Working Paper 17802.

Atella V, Brindisi F, Deb P, Rosati F. 2004. Determinants of access to physician services in Italy: a latent class seemingly unrelated probit approach. *Health Economics* **13**(7): 657–668.

Augurzky B, Bauer TK, Schaffner S. 2006. Copayments in the German health system - Do they work? RWI : Discussion Papers.

Ayyagari P, Deb P, Fletcher J, Gallo W, Sindelar JL. 2013. Understanding heterogeneity in price elasticities in the demand for alcohol for older individuals. *Health Economics* **22**(1): 89–105.

Bago d'Uva T. 2006. Latent class models for utilisation of health care. *Health Economics* **15**(4): 329 – 343.

Bitler M, Gelbach J, Hoynes H. 2006. What mean impacts miss: distributional effects of welfare reform experiments. *American Economic Review* **96**(4): 988–1012.

Cameron A, Trivedi P. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press.

Celeux G, Soromenho G. 1996. An entropy criterion for assessing the number of clusters in a mixture model. *Journal of Classification* **13**(2): 195–212.

Charlson M, Pompei P, Ales K, MacKenzie C. 1987. A new method of classifying prognostic comorbidity in longitudinal studies: Development and validation. *Journal of Chronic Diseases* **40**(5): 373–383.

Contoyannis P, Hurley J, Grootendorst P, Jeon S, Tamblyn R. 2005. Estimating the price elasticity of expenditure for prescription drugs in the presence of non-linear price schedules: an illustration from Quebec, Canada. *Health Economics* **14**(9): 909–923.

Deb P, Trivedi P. 1997. Demand for medical care by the elderly: a finite mixture approach. *Journal of Applied Econometrics* **12**(3): 313–336.

Deb P, Trivedi PK. 2002. The structure of demand for health care: latent class versus two-part models. *Journal of Health Economics* **21**(4): 601–625.

Deb P, Gallo W, Ayyagari P, Fletcher J, Sindelar J. 2011. The effect of job loss on overweight and drinking. *Journal of Health Economics* **30**(2): 317–327.

Djebbari H, Smith J. 2008. Heterogeneous impacts in PROGRESA. *Journal of Econometrics* **145**(1-2): 64–80.

Einav L, Finkelstein A, Schrimpf P. 2013. The Response of Drug Expenditure to Non-Linear Contract Design: Evidence from Medicare Part D. NBER Working Paper 19393.

Ellis R. 1986. Rational behavior in the presence of coverage ceilings and deductibles. *The RAND Journal of Economics* **17**(2): 158–175.

Farbmacher H, Winter J. 2013. Per-period co-payments and the demand for health care: Evidence from survey and claims data. *Health Economics* **22**(9): 1111–1123.

Grobe T, Dörning H, Schartz F. 2010. BARMER GEK Arztreport. Schriftenreihe zur Gesundheitsanalyse.

Keeler E, Newhouse J, Phelps C. 1977. Deductibles and the demand for medical care services: The theory of a consumer facing a variable price schedule under uncertainty. *Econometrica* **45**(3): 641–655.

Manski C. 2004. Measuring expectations. *Econometrica* **72**(5): 1329–1376.

Meyerhoefer C, Zuvekas S. 2010. New estimates of the demand for physical and mental health treatment. *Health Economics* **19**(3): 297–315.

Rückert I, Böcken J, Mielck A. 2008. Are German patients burdened by the practice charge for physician visits('Praxisgebuehr')? A cross sectional analysis of socio-economic and health related factors. *BMC Health Services Research* **8**(232).

Schreyögg J, Grabka MM. 2010. Copayments for ambulatory care in Germany: a natural experiment using a difference-in-difference approach. *European Journal of Health Economics* **11**(3): 331–341.

Winkelmann R. 2006. Reforming health care: Evidence from quantile regressions for counts. *Journal of Health Economics* **25**(1): 131–145.

Appendices

A   Charlson Index

<div align="center">Table 5: Charlson Score</div>

	Charlson Comorbidity	Assigned Weights
1	Myocardial infarction	1
2	Congestive heart failure	1
3	Peripheral vascular disease	1
4	Cerebrovascular disease	1
5	Dementia	1
6	Chronic pulmonary disease	1
7	Rheumatic disease	1
8	Peptic ulcer disease	1
9	Mild liver disease	1
10	Diabetes without complications	1
11	Diabetes with chronic complications	2
12	Hemiplegia or paraplegia	2
13	Renal disease	2
14	Cancer	2
15	Moderate or severe liver disease	3
16	Metastatic carcinoma	6
17	AIDS/HIV	6

B Marginal Effects

This appendix describes how the marginal effects on the probabilities of each type of doctor visit and on the joint probability that neither type of visit occurs ($Pr(y_{1i} = 0, y_{2i} = 0 | x_i, \theta_j)$) are calculated. Standard errors of the marginal effects are derived using the delta method. Each marginal effect is calculated for each individual i and for each of the J latent classes. We report the average marginal effects of each latent class j as the average over all individual marginal effects in this class. Furthermore, the weighted average of the effects in the different latent classes gives an overall effect. For continuous explanatory variables, the marginal effects are calculated using the calculus method. Marginal effects of binary variables are calculated with the finite difference method.

For a continuous variable x, the marginal effects on $Pr(y_{ki} = 1 | x_i, \widehat{\theta}_j)$ with $k \in \{1, 2\}$ in latent class j for individual i are calculated as

$$ME_{xi}Pr(y_{1i} = 1 | x_i, \widehat{\theta}_j)_j = \widehat{\beta}_{x,j} \phi(x_i' \widehat{\beta}_j) \tag{6}$$

$$ME_{xi}Pr(y_{2i} = 1 | x_i, \widehat{\theta}_j)_j = \widehat{\gamma}_{x,j} \phi(x_i' \widehat{\gamma}_j) \tag{7}$$

where $\phi()$ stands for the standard normal p.d.f. We report $\frac{1}{I} \sum_{i=1}^{I} ME_{xi}Pr(y_{ki} = 1 | x_i, \widehat{\theta}_j)_j$.

The marginal effects of a continuous variable x on the joint probability for y_{1i} and y_{2i} in latent class j are calculated for individual i

$$
\begin{aligned}
ME_{xi}Pr(y_{1i} = 0, y_{2i} = 0 | x_i, \widehat{\theta}_j)_j &= \frac{\partial \Phi_2(q_{1i} x_i' \widehat{\beta}_j, q_{2i} x_{it}' \widehat{\gamma}_j, q_{1i} q_{2i} \widehat{\rho}_j)}{\partial x_i} \\
&= q_{1i} \widehat{\beta}_{x,j} \phi(q_{1i} x_{it}' \widehat{\beta}_j) \Phi \left(\frac{q_{2i} x_{it}' \widehat{\gamma}_j - q_{1i}^2 q_{2i} \widehat{\rho}_j x_{it}' \widehat{\beta}_j}{\sqrt{1 - \widehat{\rho}_j^2}} \right) \\
&+ q_{2i} \widehat{\gamma}_{x,j} \phi(q_{2i} x_{it}' \widehat{\gamma}_j) \Phi \left(\frac{q_{1i} x_i' \widehat{\beta}_j - q_{1i} q_{2i}^2 \widehat{\rho}_j x_i' \widehat{\gamma}_j}{\sqrt{1 - \widehat{\rho}_j^2}} \right)
\end{aligned} \tag{8}
$$

where $\Phi_2()$ stands for the cumulative bivariate normal distribution function, $\Phi()$ indicates the standard normal c.d.f., and $q_{ki} = 2y_{ki} - 1$, with $k \in \{1, 2\}$.

The calculation of marginal effects of discrete variables is illustrated for the year dummies. The marginal effects of the years 2003, 2004 and 2005 with 2002 as reference in latent class j for individual i are calculated as

$$ME_{year,i}Pr(y_{1i} = 1 | x_i, \widehat{\theta}_j)_j = \Phi(x_i' \widehat{\beta}_j + \widehat{\beta}_{year,j}) - \Phi(x_i' \widehat{\beta}_j) \tag{9}$$

$$ME_{year,i}Pr(y_{2i} = 1 | x_i, \widehat{\theta}_j)_j = \Phi(x_i' \widehat{\gamma}_j + \widehat{\gamma}_{year,j}) - \Phi(x_i' \widehat{\gamma}_j) \tag{10}$$

for the marginal probabilities, and as

$$ME_{year,i}Pr(y_{1i}=0,y_{2i}=0|x_i,\widehat{\theta}_j)_j = \Phi_2\left(q_{1i}(x_i'\widehat{\beta}_j+\widehat{\beta}_{year,j}),q_{2i}(x_i'\widehat{\gamma}_j+\widehat{\gamma}_{year,j}),q_{1i}q_{2it}\widehat{\rho}_j\right)$$

(11)

$$- \Phi_2(q_{1i}x_i'\widehat{\beta}_j,q_{2i}x_i'\widehat{\gamma}_j,q_{1i}q_{2i}\widehat{\rho}_j)$$

for the joint probability of y_{1i} and y_{2i}. $\widehat{\beta}_j$ and $\widehat{\gamma}_j$ now stand for the vectors of parameter estimates for all variables excluding the year indicators. Marginal effects for the variable female and the quarter dummies are calculated analogously.

The overall marginal effect, i.e. the marginal effect averaged over the latent classes, for any continuous or discrete variable x is then derived as weighted average of the marginal effects across the different latent classes,

$$ME_{xi} = \sum_{j=1}^{J} \pi_j ME_{xi,j}\,.$$

(12)

Again, averages over all individuals are reported.

Standard errors for the average marginal effects are derived using the delta method that delivers the variance for each average marginal effect as

$$Var(ME) = \nabla_g' Var(\theta)\nabla_g\,,$$

(13)

where θ is the vector of all parameters that are estimated (β_j,γ_j,ρ_j, and p_j, where $\pi_j = \frac{exp(p_j)}{1+\sum_{c=1}^{C}exp(p_c)}$ with $C=J-1$), and ∇_g stands for the gradient of the marginal effect, $ME = g(\theta)$, with respect to θ. In order to calculate the variance of the average marginal effect, we set each element in the gradient to its sample average.

Table 6: Assignment to Latent Classes

2 LCs \ 4 LCs	LC 1	LC 2	LC 3	LC 4
LC 1	42.43	0.00	48.59	8.98
LC 2	0.00	65.98	14.1	19.93

C   Results for 4 Latent Classes

Table 6 displays the overlap of class assignment based on posterior probabilities after estimation of the models with 2 and 4 latent classes. Latent class 1 in the model with 2 latent classes is mostly split between latent classes 1 and 3 in the model with 4 latent classes. Latent class 2 is split between latent classes 2, 3 and 4 in the model with 4 classes. 66% of individuals assigned to latent class 2 in the 2-class model remain in latent class 2 in the 4-class model. 14% are assigned to class 3 and 20 % to class 4. While there is thus some overlap in assignment to classes between the different models, there is no reshuffling of individuals between latent classes 1 and 2. Already the 2-class model identifies the 2 extreme groups of individuals; less extreme individuals however are assigned to either of the 2 groups in the 2-class model and then separated into classes 3 and 4 in the 4-class model.

The main results of the model with 4 latent classes are displayed in figure 5. As in the model with 2 latent classes, latent class 1 contains individuals for whom the probability of no doctor visits does not change after the reform, while for individuals in latent class 2 the probability of no doctor visit increases significantly after the reform. Qualitatively, the results are thus similar to the ones in the 2-class model. The effects in the two additional latent classes, however, are similar to the effects in latent class 2, even though many individuals who are assigned *ex post* to latent class 3 are assigned to latent class 1 in the 2-class model. While the results of the 4-class model thus confirm the conclusion drawn from the 2-class model that some individuals react to the reform while others don't, there are fewer individuals in the 4-class model who do not react. Table 7 indicates that again the individuals who do not react to the reform are less healthy than the other individuals, which supports the importance of forward-looking behavior in the reaction to nonlinearities in health care price schedules.

Figure 5: 4 Latent Classes - Changes in the Probability of No Doctor Visit
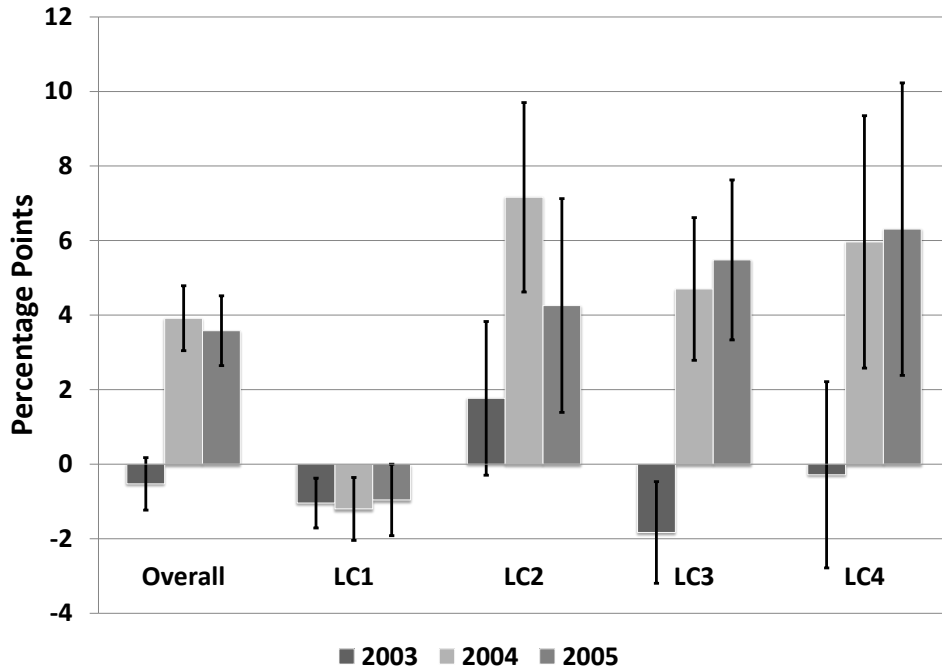
Note: Error bars indicate 99%-confidence intervals.

Table 7: 4 Latent Classes - Comparison of Group Means

	LC 1	LC 2	LC 3	LC 4
Age	57.43	44.03	56.72	44.64
Female	0.60	0.40	0.51	0.62
Charlson Index	0.98	0.17	0.85	0.47
DDD	201.26	13.07	156.22	79.73
Exempt from copayments	0.18	0.02	0.12	0.08
% of individuals	36.24	29.91	26.67	7.18