ECONSTOR
Make Your Publications Visible.

A Service of
ZBW
Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Kunz, Johannes S.; Winkelmann, Rainer

Working Paper

An econometric model of health care demand with non-linear pricing

Working Paper, No. 204

Provided in Cooperation with:
Department of Economics, University of Zurich

Suggested Citation: Kunz, Johannes S.; Winkelmann, Rainer (2015) : An econometric model of health care demand with non-linear pricing, Working Paper, No. 204, University of Zurich, Department of Economics, Zurich,
https://doi.org/10.5167/uzh-112661

This Version is available at:
https://hdl.handle.net/10419/126596

Standard-Nutzungsbedingungen:

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

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

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

Terms of use:

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

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

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

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

University of Zurich

Department of Economics

Working Paper Series

ISSN 1664-7041 (print)
ISSN 1664-705X (online)

Working Paper No. 204

An econometric model of health care demand with non-linear pricing

Johannes S. Kunz and Rainer Winkelmann

September 2015

An econometric model of health care demand with non-linear pricing *

Johannes S. Kunz
University of Zurich

Rainer Winkelmann
University of Zurich and IZA

August 2015

Abstract

From 2004 to 2012, the German social health insurance levied a co-payment for the first doctor visit in a calendar quarter. We develop a new model for estimating the effect of such a co-payment on the individual number of visits per quarter. The model allows for a one time increase in the otherwise constant hazard rate determining the timing of doctor visits, and uses a difference-in-differences strategy to identify the reform effect. The model can be adapted to a situation where the reporting period and the calendar quarter differ. Using data from the German Socio-Economic Panel, we do not find an effect of the co-payment on demand for doctor visits.

JEL Classification: I10, C25
Keywords: Count data, Poisson process, co-payment, hurdle model

*Address for correspondence: University of Zurich, Department of Economics, Zurichbergstrasse 14, CH-8032 Zurich; email: rainer.winkelmann@econ.uzh.ch. We are grateful to two anonymous referees, Gregori Baetschmann, Helmut Farbmacher, Martin Grabka, Robert Jung, Joao Santos Silva, Kevin Staub, Joachim Winter, as well as participants of the 16th IZA/CEPR European Summer Symposium in Labour Economics for valuable comments, to Daniel Auer for very able research assistance, and to the German Institute for Economic Research (DIW Berlin) for providing access to the Socio-Economic Panel.

1 Introduction

Around 90% of the German population receive their health insurance coverage through the German statutory health insurance system (SHI). Before 2004, the SHI did not require any co-payment for doctor visits, although prescription drugs were subject to cost sharing for many years. Starting January 1st, 2004, the insured had to pay a 10 Euro fee for the first visit to a doctor in each calendar quarter ("Praxisgebühr"). Additional visits in the same quarter were free of charge. Thus the individual out-of-pocket expense became a non-linear function of utilization, dropping from 10 to 0 Euros after the first doctor visit in a quarter. Only individuals without any visit to a doctor could avoid paying the quarterly fee. Moreover, the fee did not apply to those with private health insurance (PHI). On January 1st, 2013, the co-payment was abolished.

Arguably, the introduction of a co-payment created an incentive to avoid doctor visits in a particular quarter of the year, at least at the extensive margin, for those close to indifference between consulting or not consulting a doctor. One therefore would expect that the probability of visiting a doctor at least once within a quarter should have fallen in the SHI population relative to the unaffected PHI population. By the same token, one might think that the number of doctor visits for those with at least one visit (the conditional-on-positives or intensive margin effect) should be unrelated to the co-payment (see Augurzky et al., 2006, for such a view) although our analysis below will provide reasons why this is not necessarily the case.

A number of researchers have conducted quantitative evaluations of the effect of that reform on overall demand and demand at the extensive margin (Augurzky et al., 2006, Schreyögg and Grabka, 2010, Farbmacher and Winter, 2013). So far, results have been mixed. Augurzky et al. (2006) report a negative and statistically significant difference-in-differences (DiD) coefficient in a logit model for "any visit". However, in their preferred specification they control for individual specific fixed effects, and the co-payment effect

switches sign and becomes insignificant. Schreyögg and Grabka (2010) estimate a hurdle-at-zero negative binomial model and find no effect in either part of the model. Farbmacher and Winter (2013) find a statistically significant 4 percentage point reduction of the probability of any visit.

Our paper makes two main contributions. First, rather than using ad-hoc reduced form count data or binary response models to estimate the reform effect, we develop a new approach based on a structural model of health care demand. In our model, individuals are exposed to random health shocks arriving according to a homogeneous Poisson process. Individuals are myopic and decide at each instance whether or not to visit a doctor. In the random utility decision model, the out-of-pocket costs for seeing a doctor drop non-linearly after the first visit, leading to an increased hazard rate for subsequent visits. The model introduces a dynamic aspect, absent in econometric models used in prior work, where an increased cost of a first visit has two effects: it becomes more likely that no visit occurs at all; moreover, a first visit, if it take place, will tend to do so later in the quarter, lowering the overall number of subsequent visits. This does not require forward looking behavior of agents (which in fact is ruled out by our model).

The second contribution is an internally consistent approach to deal with a discrepancy between calendar quarter (i.,e., pricing period) and reporting period. This is relevant if one uses survey data such as the Socio-Economic Panel (SOEP, Wagner et al., 2007) where respondents are asked to state the number of visits during the three-months period preceeding the day of interview. Usually, the reporting period overlaps with two calendar quarters, and "reporting-mismatch" arises if, unrecorded in the data and thus unobserved by the analyst, a visit has taken place between the start of the relevant calendar quarter and that of the reporting period. In this case, the first visit in the reporting period has an effective price of zero under treatment, not of 10 Euros, as would always be the case if the reporting period and the calendar quarter matched perfectly.

If one ignores this issue, it is easy to show that a misspecified likelihood function results. Farbmacher and

Winter (2013) recommend to base estimation on the subset of individuals who were interviewed close to the end of a calendar quarter. Even if one can treat the interview day as random, this approach takes a heavy toll in terms of sample size and thus precision. For example, in our data, the number of available observations drops from 32,888 to 6,236 if the sample is restricted to those interviewed within a $\pm$ 10 days-window around the end of a quarter. On the other hand, using our assumptions on the underlying stochastic process, it becomes possible to derive the correct probability model for mismatched observations, and thus to employ all available data for estimation. Extensions to allow for unobserved heterogeneity are available.

Regarding data and identification, this paper largely follows the lead of the prior literature. For example, Augurzky et al. (2006) employ two waves from the SOEP, 2003 as pre-reform period and 2005 as post-reform period. The control group consists of people with private health insurance and provides a baseline, counterfactual pre-and post reform trend in doctor visits. Any deviation from this baseline trend observed for the treated group (SHI) is then assumed to capture the effect of treatment. Since the co-payment was stopped in 2013, we can use its abolition as a second DiD experiment, by adding information on doctor visits from the 2012 and 2013 waves (for the subset of persons interviewed in the second quarter of 2013 or later). Assuming symmetric effects, a joint analysis of the two changes allows us to increase power.

While the main contribution of the paper is the development of a new model of demand for doctor visits with non-linear pricing, and thus methodological, our results add to the existing evidence regarding the absence of evidence for a robust effect of the 10-Euro copayment on health care utilization. Perhaps, household data from the SOEP provide simply too noisy an indicator of utilization, and researchers might benefit from access to richer information provided by insurance level data, as in Farbmacher et al. (2013). Or else, there really was no sizeable effect, because the amount was simply too small to be decision-relevant for most people, because it applied to the first visit only or because, in other cases, there simply was no choice but to pay the fee, e.g., for the chronically ill.

2 Modeling the number of doctor visits

In this section, we derive the distribution of the number of visits Y during a fixed time interval $(0, T)$ representing a calendar quarter (i.e., $T = 90$ if time is measured in days). Suppose that sickness events arrive according to a Poisson process with constant rate λ. The total number of sickness events N during a quarter is then Poisson distributed with mean λT. Let $X_j \in \{0, 1\}$ indicate the individual decision to visit a doctor ($X_j = 1$) or not ($X_j = 0$) at the j'th sickness event. The decision is made by comparing two utilities, utility u_1 with a visit and utility u_0 without. Let utility be a function of income y net of the cost of a visit c_j where the cost are allowes to be a function of j. A visit takes place, and $X_j = 1$, if

$$u_1(y - c_j) > u_0(y)$$

and

$$\Pr(X_j = 1) = \Pr[u_1(y - c_j) > u_0(y)] = p(y, c_j)$$

With constant cost $c_j = \bar{c}$, the probability of a visit is the same for all sickness events, and the total number of visits

$$Y = X_1 + \ldots + X_N \tag{1}$$

has a compound Poisson distribution with mean $\lambda \times p(y, \bar{c})$ (Feller, 1977). Under the aforementioned reform, the first visit during a quarter has a price of $\bar{c}$ and all subsequent visits are free. Thus, the cost of a visit is path dependent. For example, the cost for the second sickness event can be written as $c_2 = \bar{c} \times \mathbb{1}(X_1 = 1)$, and

$$\Pr(X_2 = 1) = p(y, \bar{c}) \times p(y, 0) + (1 - p(y, \bar{c})) \times p(y, \bar{c})$$

It follows that

$$\Pr(X_1 = 1) < \Pr(X_j = 1) \qquad j = 2, \ldots, N$$

as long as $p(y, \bar{c}) < p(y, 0)$.

The compound Poisson distribution (1) has an alternative representation in terms of an underlying stochastic process. It can be obtained from the number of "renewals" (i.e. completed time spells between visits) during a fixed time interval if we assume that the number of spells are i.i.d. exponentially distributed with constant rate $\lambda \times p(y, c)$. In our model, the constant rate assumption no longer applies. The non-linear pricing model introduces a one-time jump in the hazard rate: $\lambda_0 = \lambda \times p(y, c)$ is the hazard rate for the time to first visit, and $\lambda_1 = \lambda \times p(y, 0)$ that for the duration between subsequent visits. Under the assumptions of the model, $\lambda_0 < \lambda_1$. This "non-stationarity" violates the assumptions of a renewal process and leads to a new type of count data model.

2.1 The distribution of the number of visits

Given the above assumptions, the time of the first visit t has an exponential distribution with rate λ_0, whereas the number of further visits during the quarter between t and T is Poisson distributed with rate λ_1, $Y(t, T) \sim Poisson(\lambda_1(T - t))$. Therefore, for $k \geq 1$, the total number of visits during a quarter has probability function

$$\Pr[Y(0, T) = k] = \int_0^T \frac{\exp(-\lambda_1(T - t))[\lambda_1(T - t)]^{k-1}}{(k - 1)!} \lambda_0 \exp(-\lambda_0 t) dt \qquad . \tag{2}$$

where a first visit occurs between 0 and T, if at all, and we integrate over all these possible times. If we could observe t, we would directly estimate the parameters using the terms under the integral, $\Pr(Y = k - 1 | t, t < T; \lambda_1)$ and $f(t; \lambda_0)$. Our model applies to the case, where t is unobserved, as is

typically the case in general purpose household or health surveys where just the number, and not the times, of visits is recorded.

Note that the time of first visit is not a choice variable in our model. It results from the interplay between a stochastic sickness arrival process (which is unaffected by the co-payment) and a utility maximizing choice that trades off the instantaneous health benefit of a visit with its monetary cost. One can show (see e.g. Janardan, 1980, or Baetschmann and Winkelmann, 2014) that the integral (2) has a closed form solution, and the probability function is given by

$$f(y;\lambda_0,\lambda_1) = \frac{\lambda_0\lambda_1^{y-1}\exp(-\lambda_0 T)}{(\lambda_1-\lambda_0)^y}\left[1-\sum_{j=0}^{y-1}\frac{\exp(-(\lambda_1 T-\lambda_0 T))(\lambda_1 T-\lambda_0 T)^j}{j!}\right] \qquad y=1,2,\ldots \tag{3}$$

and $f(0;\lambda_0,\lambda_1) = \exp(-\lambda_0 T)$. The model will be referred to as "dynamic hurdle" model, and we write $f(y;\lambda_0,\lambda_1) = DHurdle(y;\lambda_0,\lambda_1)$. If $\lambda_0 = \lambda_1$, it can be shown that (3) simplifies to the probability function of the Poisson distribution.

2.2 Interpretation of parameters

The parameters of the model have a straightforward interpretation. λ_0 is the hazard rate for the first visit (or "stage 0" hazard), λ_1 the hazard rate for subsequent visits (or "stage 1" hazard). For instance, parameterizing $\lambda_0 = \exp(x'\beta_0)$, where x is a $(k \times 1)$ vector of covariates and β_0 a conformable vector of regression parameters, $\beta_{0j}\Delta x_j$ is the approximate relative change in λ_0 associated with a small change in x_j. In the context of two-part or zero-inflated models, β_0 and β_1 are often denoted as "extensive margin" and "intensive margin" effects, respectively.

The mean of the model has generic form

$$\mathrm{E}(Y(0,T)) = \Pr(y>0) + \mathrm{E}_t[\mathrm{E}Y(t,T)], \qquad 0 \le t \le T$$

Since $\mathrm{E}Y(t,T) = \lambda_1(T-t)$, where $T-t$ is the time from the first visit to the end of the calendar quarter, the intensive margin effect depends not only on λ_1 but on the expected duration of stage 1, and thus on λ_0, as well. In the model, a co-payment for the first visit means that it tends to happen later, leaving less time for accumulating further visits. It is therefore not the case that the number of visits after the first visit is unaffected by the co-payment as long as λ_1 is unchanged.

Using properties of the Poisson and exponential distributions, we obtain the following closed form expression for the mean:

$$\mathrm{E}(Y;\lambda_0,\lambda_1) = \lambda_1 T + (1-\lambda_1/\lambda_0)[1-\exp(-\lambda_0 T)] \tag{4}$$

As required, the expected value of the distribution reduces to the Poisson mean when $\lambda_0 = \lambda_1$. The expected value is greater than λ_1 when $\lambda_0 > \lambda_1$, and smaller otherwise. A relative small value of λ_0 is an indication of "zero-inflation", or "extra-zeros", relative to the Poisson model, a situation encountered in many count data applications (Mullahy, 1986).

2.3 Discussion

The implied model for the first visit is identical to that used in a class of hurdle count data models introduced by Mullahy (1986). The probability function of the fixed hurdle model is given by

$$\Pr(Y=y|x) = \begin{cases} p_0(\lambda_0) & \text{for } y=0 \\ (1-p_0(\lambda_0))\dfrac{f(y;\lambda_1)}{1-f(0;\lambda_1)} & \text{for } y \geq 1 \end{cases}$$

where $f(y;\lambda_1)$ denotes the probability function of a standard count data model, e.g., Poisson or negative binomial distribution, and $p_0(\lambda_0)$ is a complementary log-log model. Pohlmeier and Ulrich (1995) argue that such a hurdle model can be appropriate for modelling the demand for health care. In their interpretation,

the first contact decision for a general practitioner often triggers a number of re-appointments or referrals to specialists that are subject to a different mechanism and thus a different λ.

The standard hurdle model is not derived from an underlying stochastic process, however. It treats λ_0 and λ_1 as unrelated parameters that can be separately estimated. It ignores the random timing of the first visit, and thus the effect of λ_0 on the length of the period for which visits have a zero co-payment. In a fixed hurdle model, conditional-on-positives expressions such as $\Pr(Y = y|Y > 0, x)$ or $\mathrm{E}(Y|Y > 0, x)$ depend on λ_1 only, not on λ_0. It therefore rules out spill-over effects and cannot address path dependence generated by non-linear pricing.

By contrast, the dynamic hurdle model (3) naturally accounts for the timing of the first visit. The co-payment leads to a lower stage 0 rate and decreases the expected time available for subsequent visits. Although the timing of the first visit is unobserved, the corresponding count data model can be derived under the maintained assumptions. In the application below, we provide results for both fixed and dynamic hurdle models. While we argue that the dynamic model is *a-priori* preferable, it is of course possible that the underlying assumptions justifying the model are not satisfied in this particular application.

2.4 Identifying the effect of a co-payment on demand

In general, the two rates of the model can be expressed as functions of a number of exogenous factors x, such as prior health status, income, gender, employment status and the like. Suppose that $\lambda_{i0} = \exp(x_i'\beta_0)$ and $\lambda_{i1} = \exp(x_i'\beta_1)$. The above model suggests that with non-linear pricing, $\lambda_{i0} < \lambda_{i1}$, and thus

$$\exp(x_i'(\beta_0 - \beta_1)) < 1 \qquad \text{for all } i$$

However, attributing any such difference in rates to the existence of a co-payment for the first visit requires the absence of other explanations. But there are a number of factors that can rationalize a low initial

rate and a higher one thereafter, perhaps the leading one being explored in the aforementioned paper by Pohlmeier and Ulrich (1995), where visits occur in clusters and a first visit is followed by additional appointments for a given sickness spell. Thus a better identification strategy uses difference-in-differences. Specifically, the co-payment did apply between 2004 and 2012 for those covered by SHI. Privately insured people were not affected and can serve as control group. We consider the following hazard rate specification:

$$\lambda_{it,j} = \exp(\theta_{t,j} + \beta_{1,j}\, SHI_i + \beta_{2,j}\, COPAY_{it} + x'_{it}\gamma_j) \qquad j = 0, 1, \qquad t = 2003, 2005, 2011, 2013$$

where $\theta_{t,j}$ are year dummies (2003 is dropped) and $COPAY_{it}$ is a dummy variable equal to one if the person is covered by SHI and the year is either 2005 or 2011, and else equal to zero. Thus, $COPAY_{it} = 1$ indicates active treatment, and $\beta_{2,0}$ is the extensive margin treatment effect under the "parallel trends assumption". This assumption implies that the counterfactual hazard rate for a first visit for the SHI population in the absence of a co-payment is equal to the actual SHI rate when no co-payment was in place (e.g., in 2003) multiplied by the appropriate trend growth factor obtained from the PHI population (e.g., $\exp(\theta_{2005,0})$).

Similarly, one could formulate a DiD effect for the (second) hazard rate for further visits, $\lambda_{it,1}$. This offers a kind of placebo test, as, within the above model, the reform did not change the incentives conditional on a first visit, and no effect should therefore be observed (i.e., the null hypothesis $H_0 : \beta_{2,1} = 0$ should not be rejected). There are good reasons not to put too much weight on such a test, though. First, there was a concurrent increase in the co-payments for prescription drugs on January 1, 2004, and we know from previous research that such out-of-pocket expenses tend to reduce the number of doctor visits as well (Winkelmann, 2004). Second, a referral from the first doctor was needed in order to receive free consultations by further doctors (or specialists) introducing an additional element to the overall costs (here in terms of time and effort).

2.5 Dealing with mismatch

The empirical analysis is based on information from the GSOEP on the number of visits "during the previous three months". Since interviews typically do not take place at the end of a calendar quarter, the reporting period overlaps with two calendar quarters. As noted by Farbmacher and Winter (2013) the standard models are invalid in this case. By contrast, our model prescribes a method to deal with mismatch in a theory consistent way, owing to the derivation of the dynamic hurdle model from an underlying stochastic process, which the standard hurdle model lacks. Consider a reporting period $(0+r, T+r)$ that differs from the calendar quarter $(0, T)$ (e.g., $T = 90$ if time is measured in days). $r \in (0, T]$ is a known value in our data, since the day of the interview is recorded. Suppose a person gets interviewed on May 1st in a year. In this case, the relevant calendar quarter started on April 1st, $r = 30$, and the reporting period covers the final two months of the 2nd quarter and the first month of the 3rd quarter.

— — — Figure 1 about here — — —

This situation is illustrated in Figure 1, where visits are reported for periods B and C, whereas the calendar quarter includes A and B. In this case, the probability of no visit is the product of the probability of no visit in C times the probability of no visit in B, which in turn depends on whether or not a visit has taken place in A. In our model, the probability of a pre-reporting period visit is $\Pr(Y_A > 0) = 1 - \exp(-\lambda_0 r)$, and therefore,

$$\Pr(Y_B = 0) = (1 - \exp(-\lambda_0 r)) \exp(-\lambda_1 (T - r)) + \exp(-\lambda_0 r) \exp(-\lambda_0 (T - r)) \tag{5}$$

and

$$\Pr(Y_{B+C} = 0) = \Pr(Y_B = 0) \times \exp(-\lambda_0 r) \tag{6}$$

where $\Pr(Y_B = 0)$ is defined in (5). Ignoring mismatch would lead one to assume a different probability expression, in this case $\exp(-\lambda_0 T)$, and thus a misspecified model.

The expressions get more complex for $Y \geq 1$. The total number of events in the reporting period, Y, is then given by the sum $Y = Y_B + Y_C$, and

$$\Pr(Y = k) = \sum_{s=0}^{k} \Pr(Y_B = s) \Pr(Y_C = k - s) \tag{7}$$

This equation requires independence between two quarters which is guaranteed under the assumptions of the model. The assumption of full independence can be relaxed: independence conditional on observed or unobserved caracteristics (e.g. a common log-normally distributed unobserved heterogeneity term) would be sufficient.

Equation (7) depends on two probabilities. The second probability is easy to establish since $\Pr(Y_C = k - s) = DHurdle(k - s; \lambda_0, \lambda_1, r)$. The first probability is a mixture of two distributions:

1. There has been at least one visit in A, i.e., the arrival time of the first event, t, predates r. In this case, counts in period B follow a Poisson distribution with mean $\lambda_1(T - r)$.

2. The arrival time of the first event exceeds r. In this case, counts in period B follow a dynamic hurdle model with parameters λ_0, λ_1, and $T - r$.

Combining terms,

$$\Pr(Y_B = s) = (1 - \exp(-\lambda_0 r)) \times Poisson(s; \lambda_1(T - r)) + \exp(-\lambda_0 r) \times DHurdle(s; \lambda_0, \lambda_1, T - r)$$

Substituting these expressions into (7), it is clear that

$$\Pr(Y_{A+B} = k) = \Pr[Y(0, T) = k] \neq \Pr(Y_{B+C} = k) = \Pr[Y(r, T + r) = k]$$

and the model is far from stationary (unless $\lambda_0 = \lambda_1$, of course). The key point is that the standard model assumes observation period and calendar quarter to be identical. If the two diverge, it is not clear whether the hazard of the observation period starts at λ_0 or at λ_1, since we do not know, whether or not an event has already taken place in the previous quarter.

As a corollary, all standard models used in the previous literature on first-visit co-payment effects (probit, logit and hurdle count models) are misspecified, and thus inconsistent when applied to a sample from survey data with mismatched reporting period. For consistent parameter estimation, one can use a subset of observations for which reporting period and calendar quarter are roughly aligned. It would be better, though, to estimate a model, such as the one derived here, that explicitly accounts for mismatch and allows for fully efficient maximum likelihood estimation using the entire sample.

3 Data and Results

Data have been extracted from the Socio-Economic Panel (SOEP) that is made available by DIW Berlin. Four years are used, 2003, 2005, 2011 and 2013 and the analysis is restricted to individuals between the ages of 20 and 60. We don't impose a balanced panel, nor do we make the assumption of independence of observations across time. This affects the way standard errors should be computed, as we use clustered standard errors throughout. We consider two samples: The "restricted sample" includes all persons, who were interviewed within ± 10 days to the end of a calendar quarter. There are 6,236 such observations. In the restricted sample, the average distance to the nearest end-of-quarter is around 5 days. The second, the "full sample", includes all persons regardless of their time of interview. There are 32,888 such observations, and thus more than five times as many as in the restricted sample. The average distance to the nearest end-of-quarter in the full sample is about 24 days.

— — — Table 1 about here — — —

Table 1 reports means (and their standard errors) of variables employed in the estimation, separately for treatment group (SHI) and control group (PHI) and the two samples. Civil servants are excluded from the analysis due to their non-standard insurance arrangement. Selection into PHI is primarily based on income, as is evident in Table 1, where the mean log household income of PHI individuals is 0.5 above that of SHI individuals (in the full sample), corresponding to a 65 percent income difference. PHI individuals are also older on average (by about 3 years), more educated (by about 2 years) and more likely to be male. The regression analyses control for these factors, but the differences still raise the question of the comparability of the two groups, and hence of the validity of the parallel trends assumption underlying the DiD identification strategy. A formal test of this assumption was conducted by Farbmacher and Winter (2013), using the same kind of data from the SOEP, and they could not reject the null-hypothesis of parallel trends during the pre-treatment years.

The bottom panel of Table 1 splits the two samples further into two subsamples, the treatment years 2005 and 2011, and the non-treatment years 2003 and 2013 and shows for each group the average number of visits as well as the share with at least one visit. The non-treatment sample is somewhat smaller. The reason is that we had to drop all respondents who were surveyed during the first quarter of 2013, because their reporting period overlapped with the abolition of the co-payment by January 1st, 2013.

In terms of mean utilization, and using the full sample, we find that the SHI reported on average about 0.05 (or 2.5 percent) more visits in years without the co-payment in place than in years with co-payment. However, a similar trend is observed for the PHI individuals, who report about 0.1 fewer visits on average, so that the naive DiD effect, based on means only, is close to zero. A somewhat noisier picture emerges for the restricted sample, where the size of the control population (417 in treatment years and 310 in non-treatment years) leads to considerably more sampling uncertainty. If we look at the probability of

any visit instead – the extensive margin – not much is found either. The probability actually increases somewhat for SHI individuals in years where the co-payment is in place. Of course, it would be premature to dismiss the possibility of a treatment effect based on this descriptive evidence alone.

3.1 Baseline result

In a next step, we estimated the fixed and dynamic hurdle Poisson models with a full set of control variables. The upper half of Table 2 shows the results. In terms of overall fit, the simple Poisson model is clearly inferior to the two-part generalizations that introduce different parameters for the utilization (yes/no) decision and for the intensity of use. The dynamic hurdle model leads to a substantially higher loglikelihood value relative to the fixed hurdle model (-12,690.4 as compared to -13,252.4).

Recall that λ_0 is the hazard rate for the time to a first visit. The probability of no visit is then equal to the survivor rate $\exp(-\lambda_0)$. Harzard rate and survivor rate are inversely related, i.e., factors increasing the hazard rate lower the probability of no visit, and vice versa. With the exponential parameterization, the displayed coefficients indicate the predicted approximate relative change in the hazard rate associated with a unit change in the associated regressor. The exact relative change is obtained by applying the transformation $\exp(\hat{\beta}_j) - 1$. For instance, according to the dynamic hurdle model, the hazard rate for a first visit for men is $(\exp(-0.357) - 1) \times 100 = 30$ percent below that of women.

— — — Table 2 about here — — —

There are some interesting asymmetries between first-visit hazard (λ_0) and that of subsequent visits (λ_1). For instance, income has no effect on the former, but a statistically significant negative effect on the latter, where a 10 percent increase in income is predicted to reduce the hazard rate for each further doctor visit by 1.5 percent. A similar pattern holds for unemployment. Individuals with disabilities have higher hazard

rates in both states, and thus a higher predicted number of doctor visits, than individuals without.

The statistically insignificant point estimate of the treatment effect corresponds to a 1.6 percent reduction in the hazard rate for the first visit. The standard error is large, so sizeable positive or negative effects cannot be ruled out either, the data simply do not have sufficient power to tell. The effect on the stage 1 hazard (were a narrow interpretation of the model would predict none) is positive and relatively large, but again, the standard error is large and the coefficient is statistically insignificant as well.

3.2 Adding unobserved heterogeneity

All models can be extended to allow for unobserved heterogeneity, for instance by adding a multiplicative random effect to the exponential rates:

$$\tilde{\lambda}_i = \exp(x_i'\beta)u_i$$

Here, u_i captures individual level differences in the demand for doctor visits, for example due to differences in latent health. In the context of our dynamic hurdle model, it seems reasonable to assume that the two rates for the first and for subsequent visits are multiplied by the same factor. Moreover, we will assume that the heterogeneity term is the same for repeated observations on the same individual, and thus $\tilde{\lambda}_{it} = \exp(x_{it}'\beta)u_i$.

The models discussed in Section 2 now hold conditional on u_i. To take the models to the data, the u_i term has to be eliminated from the likelihood function by taking expectations. Specifically, suppose that $\ln u_i$ is normally distributed, independently of x_i, with mean $-0.5\sigma^2$ and variance σ^2. Then u_i is lognormal with $\mathrm{E}(u_i) = 1$ and $\mathrm{Var}(u_i) = \exp(\sigma^2) - 1$. The marginal likelihood has no closed-form solution, but it can be numerically approximated using Gauss-Hermite quadrature.

The lower part of Table 2 provides estimation results for the three models with log-normal unobserved heterogeneity. In the case of the fixed hurdle model, heterogeneity is introduced only for the conditional-on-positives part, and the λ_0 estimates are thus identical to those in Table 2. Clearly, the improvements over the models without unobserved heterogeneity are large and statistically significant across the board. Unobserved heterogeneity changes the log-likelihood ordering of the three models, the fixed hurdle model having the highest log-likelihood in this case.

Point estimates for the important predictors of health care utilization, i.e., unemployment, disability and gender, are rather stable, regardless of whether unobserved heterogeneity is allowed for or not. The estimates for the reform effect remain statistically insignificant. Note that our specification imposes that the introduction and the abolition of the co-payment have the same effect size (and opposite sign), an assumption, that may be too restrictive. By 2012, many insured, in particular younger ones, had switched to new types of SHI contracts that offered co-payment waivers in return for joining a primary care model. We therefore estimated all models using the introduction sample only (2003 and 2005, results are available on request) but all substantive conclusions remain unchanged.

3.3 Full sample results

We also estimated the adjusted dynamic hurdle model, as discussed in Section 2.5, for the entire sample, in this case based on 32,888 observations. Otherwise, the specification remained unchanged. Results shown here are for the specification witout unobserved heterogeneity. For the first visit, or stage 0, we obtained the log of the predicted hazard (with standard errors in parentheses):

$$\ln \hat{\lambda}_{it,0} = \underset{(.045)}{0.037}\,\mathrm{COPAY}_{it} + \underset{(.029)}{0.598}\,\mathrm{Disability}_{it} - \underset{(.017)}{0.353}\,\mathrm{Male}_i + \text{other terms}$$

The standard errors decrease roughly with the root of the sample size, compared the results presented in Table 2, and there is the expected efficiency gain from using the larger sample. The point estimates are in line with previous results. Again, no statistically significant reform effect is found.

For the stage 1 hazard, we get

$$\ln \hat{\lambda}_{it,1} = \underset{(.081)}{-0.041}\,\text{COPAY}_{it} + \underset{(.040)}{0.537}\,\text{Disability}_{it} - \underset{(.029)}{0.068}\,\text{Male}_i + \text{other terms}$$

The finding of a large and about evenly distributed disability effect remains robust, as does the pattern that men have a substantially lower stage 0 hazard than women, whereas their stage 1 hazard does not differ much.

4 Concluding remarks

This paper introduced a simple econometric model of health care demand under non-linear pricing, based on a Poisson process for the arrival of sickness events. In the model, a co-payment for the first visit during a calendar quarter potentially lowers the hazard rate for the first visit, leaving the subsequent hazard for further visits unchanged. The model was applied to an evaluation of a German health care reform of 2004 when a co-payment of 10 Euros was introduced for those covered by statutory health insurance. The co-payment was again abolished in 2013. In none of our various specifications, with or without unobserved heterogeneity and with or without reporting mismatch, did we find statistically significant effect of the co-payment on the number of doctor visits.

While the results thus are in line with those of two earlier studies by Augurzky et al. (2006) and Schreyögg and Grabka (2010) who also found no effect of the co-payment on utilization, the new methodological approach of this paper offers a number of additional insights that might prove useful for future research

in related contexts. For instance, the perspective of a stochastic process is useful to understand that any changes to the first hazard likely also affects the distribution of additional visits, simply because it changes the time left in the quarter to accumulate such visits.

Second, the approach also points towards a theory consistent way to derive the likelihood of mismatched observations, i.e., observations for which reporting period and calendar quarter do not coincide. Such an approach avoids a loss of information incurred by limiting the sample to people interviewed at the end of a calendar quarter, and thus increases power. Of course, such data problems could also be addressed by using data from insurance claims rather than survey data. However, such claim data have their own problems. First, they usually include only a very limited set of socio-economic control variabels, precluding certain types of analyses. Furthermore, they often do not allow for a DiD analysis, as it is unlikely that a useful control group can be established in a given claims dataset.

The models discussed in this paper should be useful in a number of other applications as well. For instance, in order to reduce absenteeism, firms have started to offer bonuses to workers with zero sick leave days per year. Depending on how the details of such schemes are designed, they may imply that the first absence is rather costly (the loss of the bonus) whereas subsequent absences have much lower cost. Similar non-linear pricing schemes can be observed for re-offenses in the context of fare dodging, were fines usually increase after the first offense. Moreover, this paper contributes to the literature on zero-inflated count data that are often encountered in health economic applications, adding a new, alternative model to the existing toolkit.

References

Augurzky, B., T.K. Bauer and S. Schaffner, S. (2006) Co-payments in the German health system - do they work? RWI Essen Discussion Paper No. 43.

Baetschmann, G., and R. Winkelmann (2014) A dynamic hurdle model for zero-inflated count data, unpublished manuscript.

Cameron, A.C. and P.K. Trivedi (1986) Econometric models based on count data: comparisons and applications of some estimators and tests, *Journal of Applied Econometrics* 1, 29-53.

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

Farbmacher, H., P. Ihle, I. Schubert, J. Winter and A.C. Wuppermann (2013) Heterogeneous effects of a nonlinear price schedule for outpatient care (No. 4499), CESifo Working Paper.

Feller, W. (1977), *An Introduction to Probability Theory and Its Applications*, second edition, New York: John Wiley & Sons.

Janardan, K.G. (1980), A stochastic model for the study of oviposition evolution of the pest callosobruchus maculatus on mung beans, phaseolus aureus, *Math. Biosciences*, 50, 231-238.

Mullahy, J. (1986), Specification and testing in some modified count data models, *Journal of Econometrics*, 33, 341-365.

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

Schreyögg, J., and M.M. Grabka (2010) Co-payments for ambulatory care in Germany: a natural experiment using a difference-in-difference approach, *European Journal of Health Economics*, 11(3),

331-341.

Wagner, G.G., J.R. Frick and J. Schupp (2007), The German Socio-Economic Panel Study (SOEP): scope, evolution and enhancements, *Schmollers Jahrbuch*, 127, 139-169.

Winkelmann, R. (2004), Health care reform and the number of doctor visitsan econometric analysis, *Journal of Applied Econometrics* 19, 455472.

Figure 1. *Mismatch between reporting period and calendar quarter.*

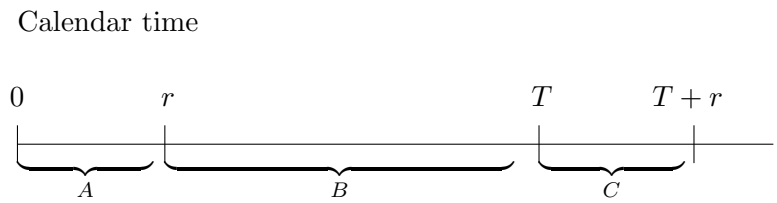

Table 1. *Descriptive statistics by insurance status for restricted and full sample (Means, cluster robust standard errors in parentheses)*

	Restricted Sample		Full Sample	
	PHI	SHI	PHI	SHI
Co-payment (yes/no)	0	0.580	0	0.577
		(0.006)		(0.002)
Distance to nearest end-of-quarter	4.642	5.254	23.721	23.811
	(0.115)	(0.043)	(0.224)	(0.079)
Age	44.281	40.899	43.890	41.004
	(0.418)	(0.162)	(0.226)	(0.090)
Male (yes/no)	0.597	0.452	0.614	0.447
	(0.021)	(0.008)	(0.012)	(0.004)
Years of schooling	14.096	12.032	14.148	11.969
	(0.124)	(0.038)	(0.071)	(0.022)
Disability (yes/no)	0.045	0.069	0.043	0.074
	(0.008)	(0.004)	(0.004)	(0.002)
Log net household income	10.954	10.490	10.962	10.466
	(0.025)	(0.009)	(0.014)	(0.005)
N	727	5509	3725	29163
Years with co-payment (2005,2011)				
Number of doctor visits	1.830	2.048	1.902	2.072
	(0.135)	(0.051)	(0.067)	(0.024)
Any visit (yes/no)	0.607	0.643	0.603	0.652
	(0.025)	(0.009)	(0.011)	(0.004)
N	417	3197	2096	16833
Years without co-payment (2003,2013)				
Number of doctor visits	2.216	2.039	1.988	2.123
	(0.197)	(0.061)	(0.076)	(0.028)
Any visit (yes/no)	0.619	0.638	0.622	0.649
	(0.027)	(0.010)	(0.012)	(0.004)
N	310	2312	1629	12330

Source: Socio-Economic Panel (SOEP), version 26, doi:10.5684/soep.v26

PHI: private health insurance; SHI: social health insurance

Table 2. *Poisson and hurdle models of health care utilization, restricted sample (N=6,236)*

	Poisson	Fixed Hurdle		Dynamic Hurdle	
	λ	λ_0	λ_1	λ_0	λ_1
without unobserved heterogeneity					
SHI	-0.196	0.007	-0.234	0.079	-0.447
	(0.094)	(0.082)	(0.093)	(0.076)	(0.169)
Co-payment (yes/no)	0.158	0.041	0.150	-0.016	0.297
	(0.119)	(0.104)	(0.123)	(0.098)	(0.209)
Unemployment (yes/no)	0.329	-0.014	0.369	-0.079	0.567
	(0.080)	(0.078)	(0.076)	(0.068)	(0.122)
Disability (yes/no)	0.823	0.817	0.535	0.654	0.563
	(0.055)	(0.075)	(0.054)	(0.060)	(0.086)
Male (yes/no)	-0.330	-0.380	-0.146	-0.357	-0.105
	(0.038)	(0.036)	(0.039)	(0.033)	(0.061)
Log net household income	-0.105	-0.053	-0.091	-0.025	-0.145
	(0.032)	(0.030)	(0.033)	(0.028)	(0.054)
Log likelihood	-14,781.7	-13,252.4		-12,690.4	
with unobserved heterogeneity					
SHI	-0.263	0.007	-0.241	0.104	-0.363
	(0.105)	(0.082)	(0.095)	(0.103)	(0.155)
Co-payment (yes/no)	0.275	0.041	0.175	-0.016	0.244
	(0.134)	(0.104)	(0.138)	(0.131)	(0.191)
Unemployment (yes/no)	0.288	-0.014	0.271	-0.119	0.542
	(0.093)	(0.078)	(0.095)	(0.093)	(0.117)
Disability (yes/no)	0.772	0.817	0.613	0.794	0.584
	(0.069)	(0.075)	(0.089)	(0.080)	(0.079)
Male (yes/no)	-0.335	-0.380	-0.185	-0.421	-0.150
	(0.038)	(0.036)	(0.045)	(0.044)	(0.056)
Log net household income	-0.090	-0.053	-0.106	-0.043	-0.111
	(0.032)	(0.030)	(0.032)	(0.037)	(0.049)
$\ln(\sigma)$	-0.002	-0.296		-0.517	
	(0.020)	(0.024)		(0.023)	
Log likelihood	-11,976.5	-11,781.3		-11,833.6	

Source: Socio-Economic Panel (SOEP), version 26, doi:10.5684/soep.v26
Dependent variable: Number of doctor visits
All models include three year dummies, a constant, a quadratic in age and level of schooling.
Standard errors in parentheses are clustered at the individual level.