University of Zurich

Socioeconomic Institute
Sozialökonomisches Institut

Working Paper No. 0311

Re-evaluating an Evaluation Study: The Case of the German Health Care Reform of 1997

Rainer Winkelmann

October 2003

Socioeconomic Institute
University of Zurich

Working Paper No. 0311

Re-evaluating an Evaluation Study: The Case of the German Health Care Reform of 1997

October 2003

Author's addresses Rainer Winkelmann
E-mail: winkelmann@sts.unizh.ch

Publisher Sozialökonomisches Institut
Bibliothek (Working Paper)
Rämistrasse 71
CH-8006 Zürich
Phone: +41-1-634 21 37
Fax: +41-1-634 49 82
URL: www.soi.unizh.ch
E-mail: soilib@soi.unizh.ch

Re-evaluating an Evaluation Study: The Case of the German Health Care Reform of 1997

Rainer Winkelmann*
University of Zurich, IZA Bonn and CEPR London

October 2003

Abstract

This paper reports on a re-evaluation of the German health care reform of 1997. A previous evaluation found a limited effect of a 4.4 percent reduction of the number of doctor visits in a sample of pharmacy customers. The re-evaluation based on a representative household survey, the German Socio-Economic Panel, yields a much larger effect. The paper uses this case study to discuss the methods and benefits of modern techniques of program evaluation.

JEL Classification: I11, I18, C25

Keywords: co-payment, counter-factual, differences-in-differences, count data, Poisson regression

Address for correspondence: Department of Economics, University of Zurich, Zürichbergstr. 14, CH-8032 Zürich. E-Mail: winkelmann@sts.unizh.ch. This paper is based on an invited lecture at the 23. Hochschulkurs aus Gesundheitsökonomik, Seefeld, 29. Sept.- 1. Oct. 2003. I am grateful to seminar participants for valuable comments.

1 Introduction

There is an increasing consensus that policy programs, such as active labor market policies or health policies, should be subject to independent scientific evaluation, either ex-ante based on pilot studies or ex-post. Of course, this leaves open the question how such policy evaluations should be conducted. This problem has been intensively discussed in the "evaluation community" which includes statisticians, biometricians, econometricians and many others, and some general rules of "good" evaluation have emerged. Not all evaluation studies, however, follow these principles, and it is the goal of this paper to use a particular reform, the 1997 health care reform in Germany, and an existing evaluation to illustrate the problems and pitfalls that can arise in any evaluation.

This paper, therefore, is in a certain sense a re-evaluation of an evaluation. I start by presenting the main features of the reform. Then I discuss the methodology and results of an existing evaluation. Next, I consider the evaluation problem from the point of view of modern evaluation research, by first repeating the basic concepts and quantities involved in such an evaluation, then adjusting the methods to the specific problem, and finally presenting the results. It turns out that the type of evaluation matters quite a bit for the results. Even so, the results of the two evaluations, the previous one and the one reported on in this paper, are eventually shown to be mutually compatible, only that the previous study does not measure the reform effect that one should really be interested in.

2 German Health Care Reforms

More than 90 percent of the German population is covered by statutory health insurance, which always includes coverage for prescription pharmaceuticals. The insurance system is mostly financed by mandatory payroll deductions. For employees, the premium is proportional to earnings (up to a contribution ceiling), and coverage automatically extends to non-working spouses and dependent children. Employees with earnings above the contribution ceiling have the option to remain insured, buy private insurance, or have no insurance at all. Special provisions exist for other groups, such as the unemployed or students. All in all, the system mixes insurance aspects with strong redistributive elements.

Because of cost overruns, the health care system has been subject to periodic reform. A reform schedule is attached in the Appendix. For example, between 1993 and 2003, there have been nine legislative changes made to the existing body of law concerning statutory health insurance. The question then arises, why I would pick the 1997 reform (the "2nd GKV Neuordnungsgesetz") for the purpose of this case study. There are really two reasons for this choice. First, the main element of the 1997 reform was an increase in patient's out-of-pocket expenses, as detailed below. Out-of-pocket expenses are an important instrument of any health care system, and knowing how they affect demand for health services therefore provides valuable information that can be used to assess expected effects of future such changes in Germany or elsewhere. Secondly, I picked the 1997 reform because of the methodological objective of the paper, namely to re-evaluate an existing evaluation. And such a prior evaluation had been conducted and made public for the 1997 reform.

The main change of the 1997 reform was an increase in prescription fees – referred to as co-payment in this paper – that must be made for all prescriptions. The amount varies by package size. Such co-payments were increased substantially on July 1, 1997, by a fixed amount of DM 6 relative to a year earlier. Since the absolute amount of the co-payment is a function of the package size, after the reform DM 9 for small, DM 11 for medium and DM 13 for large sizes, the relative effect of the 1997 reform was largest for small sizes, where it amounted to a 200 percent increase.

Not everyone was affected by the reform. First, there is the group of people with private insurance. A second group explicitly exempted from the increased co-payments are co-insured children under the age of 18. A third type of exemption status has been defined by the law through a combination of income and exposure, in an attempt to limit the financial burden related to the provision of health services for low-income families. In particular, low-income households with family gross income under DM 1700 (for singles) or DM 2350 (for couples) were exempt, as were all amounts beyond a cumulative annual co-payments of 2 percent of annual gross income.

The question is then how the increased co-payment for those subject to the 1997 reform affected their individual demand for health services. This is what we want to find out.

3 A Previous Evaluation

The 1997 reform was evaluated by Lauterbach, Gandjour and Schnell (2000, henceforth LGS). First, they asked the important question what outcome variable to consider. The main

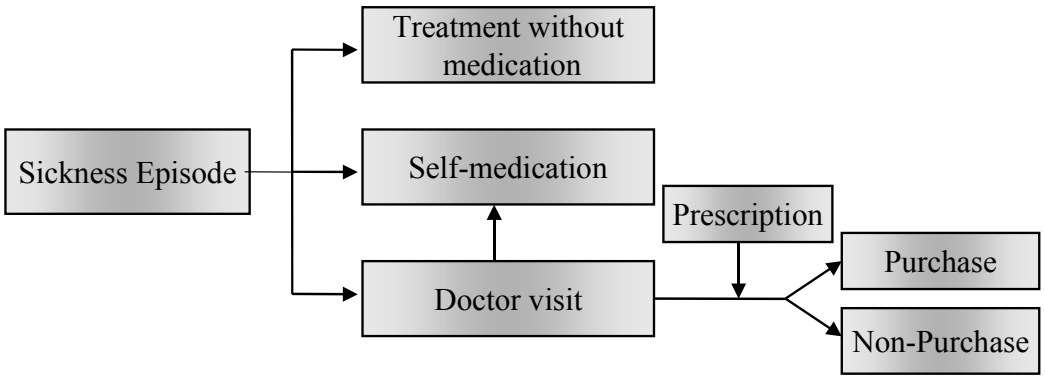
concern of the law makers was to lower expenditures while keeping the quality of the care provision the same. That this would be possible rested on the notion of previous inefficiencies based on moral hazard type incentive problems and a resulting "overutilization" of health services. Unfortunately, cost data – other than most aggregate ones – are very difficult if not impossible to come by. So it seems natural, as suggested by LGS, to concentrate instead on incidences of demand occasions, such as the number of prescriptions drug packages obtained, or the number of doctor visits. This focus also stresses the importance of individual patient behaviour that was the target of the reform.

The simple demand mechanism underlying the LGS evaluation is shown in Figure 1 (adapted from LGS, 2000). The reform changed the price for prescription drugs. However, because of technical complementarities respectively substitution possibilites, the quantity demanded for prescription drugs is not the only outcome variable of interest. Prescription drugs require a previous doctor visit. Hence, the demand for doctor visits should decrease. Also, the demand for alternative methods of treatment might increase.

LGS then proceeded with an own survey to obtain information on changes in demand. Between October and December of 1998, they distributed 10,000 questionnaires at pharmacies in the city of Cologne. Their target population were all visitors to a pharmacy that were covered by the statutory health insurance, not exempt from the co-payment, provided they visited the pharmacy due to an acute or chronic sickness. From the 10,000 questionnaires distributed, LGS obtained 695 responses.

Core elements of the questionnaire were questions such as: "Due to the recent health care

Figure 1: Prescription Drugs and the Demand for Doctor Visits



Source: Lauterbach et al. (2000)

reform, have you visited a physisician less frequently / obtained fewer prescription drugs?" Among those affirming these questions, there was then a follow-up question checking whether people renounced a single visit/drug purchase, or instead several visits/drug purchases.

With regards to doctor visits, the relative frequencies were as follows: 80.2% did say that their demand was unaffected; 8.6% did say that they renounced one visit they would have undertaken otherwise (without the reform); and 11.2% did say that they renounced repeatedly. The actual average reported number of visits during the previous 12 months was 9.2. If one assumes a figure of 3 fewer visits for those who stated that they renounced repeatedly, the hypothetical number of visits in the absence of the reform can be calculated as $9.2 + 1 \times 0.086 + 3 \times 0.112 = 9.62$. In this setting, the difference between 9.6 and 9.2 is the decrease in the number of doctor visits that can be attributed to the reform. It is the causal effect of the reform. In relative terms, this corresponds to a 4.4% reduction. LGS use

similar arguments to establish a reform effect in terms of prescription drugs, in this case a 5.3% reduction. They conclude that the effect of the reform on demand for health services by the statuorily insured was limited, and even more so for some selected sub-groups, such as the chronically ill.

So far the main results from the LGS evaluation. There are a number of more or less obvious concerns, though, that let one put a question mark behind the validity of these results. The most obvious problem is the low response rate and the small sample of 695 observations. This must not constitute a problem per se, though, as long the sample is not systematically biased. And indeed, LGS provide some evidence that the composition of their sample is similar to other studies targeting populations of pharmacy customers.

This raises the more fundamental problem that the focus on pharmacy customers was inappropriate in the first place, when the goal is to evaluate the consequences of the reforms on the demand of all members of the statutory health insurance. The reason is that the pharmacy sample is heavily biased towards sick people and frequent users of health services. This is an example of so-called endogenous sampling, where a higher value for the outcome variable, here the number of doctor visits, increases the probability of being included in the sample. As a consequence, the sample is in no way representative for the target population of the reform at large. Estimated reform effects will be biased inasmuch as the causal effect varies between frequent and less frequent users.

There are a few additional reasons why the results of this evaluation study may not be robust. First, it is one thing to *observe* how people changed their behaviour before and after

the reform, by taking two measurements, one before and one after the reform, and another to *ask* people what they perceive their change was, as done in the LGS study. Actually collecting data at two points in time seems clearly preferable, although it is more costly. Secondly, I will argue that any evidence from an evaluation study becomes much more compelling, when a non-affected comparison group, or "control-group" is included. This has not been done in the study, although it would have been possible. With a control group, one can account for example for other unobserved influences on demand behaviour that have changed over time.

4 Fundamental Concepts of Scientific Program Evaluation

This section offers a short introduction to the basic evaluation concepts, as introduced by Rubin (1974). For a more detailed discussion, see Wooldridge (2002, Chap. 18). At the beginning of each evaluation, one needs to specify the outcome variable of interest, Y. This can be the number of prescriptions or the number of visits to a doctor over a given period of time. If we define Y_1 as the value of the outcome variable with the reform, and Y_0 as the value of the outcome variable without reform, then the causal affect of the reform is given by $Y_1 - Y_0$. The fundamental problem of program evaluation is that we observe either Y_1 oder Y_0 but never both. Identification of the program effect requires additional assumptions.

Before we discuss these assumptions, we need some further definitions. We assume that measurements are made at two points in time, where $T = 1$ denotes the post-reform measurement and $T = 0$ denotes the pre-reform measurement. Similarly, we distinguish between

a treatment group subject to the reform ($D = 1$) and a control group $D = 0$. Finally, we are not interested in individual outcomes $Y_1 - Y_0$ but rather in population averages $\mathrm{E}(Y_1 - Y_0)$.

In evaluating a health care reform, the most sensible performance indicator is the so-called treatment effect on the treated:

$$\mathrm{E}(Y_1 - Y_0|D = 1, T = 1) = \mathrm{E}(Y_1|D = 1, T = 1) - \underbrace{\mathrm{E}(Y_0|D = 1, T = 1)}_{\text{unobservable}}$$

where the treated are those covered by the insurance. The main dilemma is that $\mathrm{E}(Y_0|D = 1, T = 1)$, i.e., the post-treatment outcome of the treatment group *had the reform not taken place* can never be observed. It is a so-called "counter-factual" outcome. The three identification strategies involve then replacing the unobserved outcome by a value obtained either using

1. a control group
2. a before/after comparison
3. or differences-in-differences (a combination of 1 and 2)

Control Group Comparison

In this case, we *assume* that

$$\mathrm{E}(Y_0|D = 1, T = 1) = \underbrace{\mathrm{E}(Y_0|D = 0, T = 1)}_{\text{observed}}$$

This assumption requires that treatment and control groups are similar. In particular, there may not be any systematic relationship between the fact that a person receives the treatment

(here: is insured by the statutory health insurance) and the outcome (here: e.g., the number of doctor visits). This requirement is for example fulfilled in controlled experiments, where the allocation to treatment and control is randomized. In the current context, this is clearly not the case. For example, privately insured persons (part of the control group) have higher incomes and therefore a better health status on average. The condition can be somewhat weakened by conditioning on covariates such that

$$\mathrm{E}(Y_0|D=1,T=1,X) = \mathrm{E}(Y_0|D=0,T=1,X)$$

Now, systematic differences between treatment and control group are fine, as long as they are associated with observable difference in X.

Before/After Comparison

In this case, we *assume* that

$$\mathrm{E}(Y_0|D=1,T=1) = \underbrace{\mathrm{E}(Y_0|D=1,T=0)}_{\text{observed}}$$

In words, this assumption states that the outcome variable for the treatment group would not have changed in the absence of the treatment. This requires the absence of any other influence factors that may have changed over time as well. Again, the condition can be weakend by including covariates:

$$\mathrm{E}(Y_0|D=1,T=1,X) = \mathrm{E}(Y_0|D=1,T=0,X)$$

Now, we require the absence of changes in unobserved influence factors over time.

Differences-in differences

In this case, the identifying assumption is

$$\mathrm{E}(Y_0|D=1,T=1) - \mathrm{E}(Y_0|D=1,T=0) = \mathrm{E}(Y_0|D=0,T=1) - \mathrm{E}(Y_0|D=0,T=0)$$

i.e., the *change* for the treatment group in the absence ot the treatment would have been the same as the actual *change* observed for the control group, in which case the counter-factual outcome can be expressed as

$$\mathrm{E}(Y_0|D=1,T=1) = \mathrm{E}(Y_0|D=1,T=0) + \mathrm{E}(Y_0|D=0,T=1) - \mathrm{E}(Y_0|D=0,T=0)$$

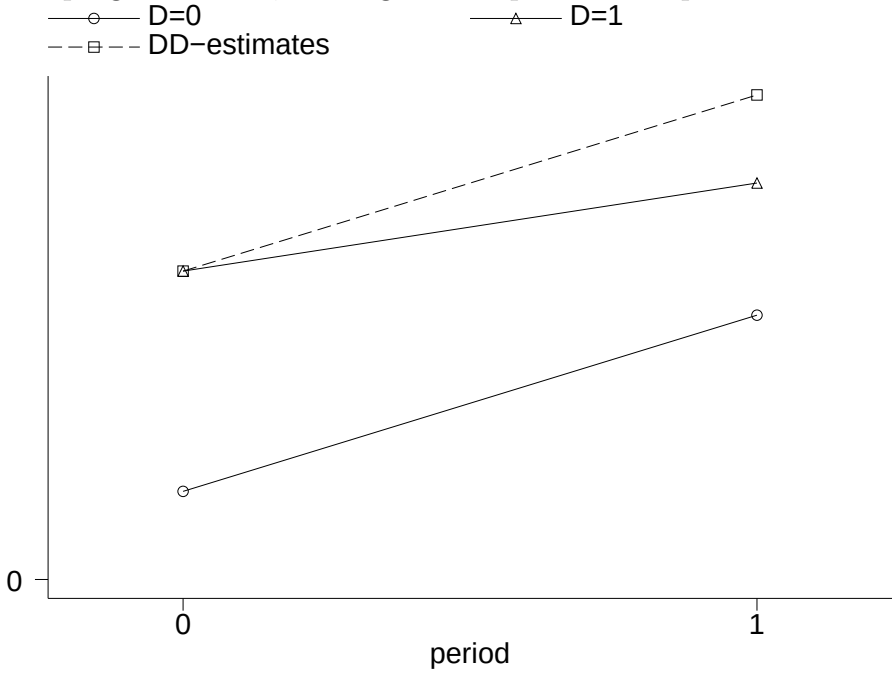
This assumption is less demanding than assumptions 1 and 2, since we

- do **not** require that treatment and control group are the same, and
- do **not** require that there were no changes in unobservables for the treatment group.

Identification based on difference-in-difference (dd) therefore is a preferred strategy, although the validity of the dd assumption cannot be tested.

The essence of the evaluation model is summarized in Figure 2. The two solid lines give the development of the outcome variable over time for the control (lower line) and treatment group (upper line), respectively. The dashed line shows, how the outcome of the treatment group would have developed based on the observed change for the control group. In this stylized example, the program effect would be negative (e.g., a reduction in demand), since the outcome in the treatment group fell *relative* to the outcome in the control group. This can occur despite the fact that absolute changes are positive for both groups. What matters in the dd framework are the relative changes. In this example, treatment effects

estimates basod on the two stronger assumptions would lead to opposite results, namely positive program effects, although the required assumptions are clearly invalid.



5 Implementation

There are many ways to implement the dd-estimator. Clearly the simplest one (but also the most restrictive one) is to embed it in a regression framework

$$\mathrm{E}(Y|D,T) = \beta_0 + \beta_1 D + \beta_2 T + \beta_3 D \times T$$

where $D \times T$ is an interaction between period and treatment status. In this case, we have

$$\mathrm{E}(Y_1|D=1,T=1) = \beta_0 + \beta_1 + \beta_2 + \beta_3$$

and

$$\mathrm{E}(Y_0|D=1,T=1) = (\beta_0 + \beta_1) + (\beta_0 + \beta_2) - (\beta_0)$$

Therefore, the treatment effect is simply

$$\mathrm{E}(Y_1|D=1,T=1) - \mathrm{E}(Y_0|D=1,T=1) = \beta_3$$

Under these simplifying assumptions, the treatment effect is assumed to be homogenous, i.e., there is no difference between the treatment effect on the treated, the average treatment effect and the marginal treatment effect. An advantage of the simplicity is that it can easily be implemented into a Poisson (or any other count data) regression model, thereby accounting for the nature of the dependent variable in the health utilization application that follows.

In particular, let

$$y_{it} = \exp(\beta_0 + \beta_1 D_i + \beta_2 T_t + \beta_3 D_i \times T_t + \gamma x_{it}) + u_{it} \tag{1}$$

where y_{it} is the number of doctor visits of person i at time t and x_{it} are additional control variables.

In equation (1), β_3 measures the reform effect. If it is negative, the demand for doctor visits in the treatment group fell relative to the demand in the control group after the imposition of the increased co-payments. Since the conditional expectation function is log-linear, the coefficient identifies the relative change in expected demand. The dependent variable being a count (the number of doctor visits in the previous quarter), efficient estimation requires that the conditional expectation function implied by (1) is embedded into a count data model. First, I use the Poisson probability function

$$f(y_{it}|\lambda_{it}) = \frac{\exp(-\lambda_{it})\lambda_{it}^{y_{it}}}{y_{it}!} \tag{2}$$

where $\lambda_{it} = \mathrm{E}(y_{it}|\lambda_{it})$ is defined as in (1). The model parameters are estimated by pseudo maximum likelihood. As long as the conditional expectation function is correctly specified this approach leads to consistent parameter estimates and valid inference (Gourieroux, Monfort and Trognon, 1984).

The alternative is to use a count data model with random effects. Let

$$\mathrm{E}(y_{it}|\lambda_{it}, \alpha_i) = \lambda_{it} \exp(\alpha_i) \tag{3}$$

where the conditional distribution of y_{it} given λ_{it} and α_i is a Poisson distribution, and α_i is independently gamma distributed. The resulting marginal model is a panel Poisson model of a negative binomial variety, with joint probability function for individual i

$$f(y_{i1}, \ldots, y_{iT}) = \frac{\Gamma(\sum_t y_{it} + \gamma)}{\Gamma(\gamma)} \left(\frac{\gamma}{\gamma + \sum_t \lambda_{it}} \right)^{\gamma} \frac{1}{(\gamma + \sum_t \lambda_{it})^{\sum_t y_{it}}} \prod_{t=1}^{T} \left(\frac{\lambda_{it}^{y_{it}}}{y_{it}!} \right) \tag{4}$$

where γ is an additional parameter. This model provides potentially more efficient parameter estimates than the Poisson model (see Winkelmann, 2003b).

6 A Re-Evaluation of the 1997 Reform

We see now that the LGS study does not fit into this standard evaluation framework. However, a full dd-estimator as described above can be implemented using publicly available data from the German Socio-Economic Panel (GSOEP). The GSOEP is a fully representative household survey that was initiated in 1984 (SOEP Group, 2001) and has been repeated annually ever since. For the purpose of this study, I selected a period of four years centered around the year of the reform, i.e., 1995, 1996, 1998 and 1999 (excluding 1997). The GSOEP

has a handful of variables relating to the usage of health service. One of them is the number of visits to a doctor during the previous 3 months. Direct information on prescription drug use is not available, and this aspect of the reform cannot be evaluated. I use observations on men and women from Sample A, i.e., persons associated with non-guestworker-households in the original sample for West Germany.

In the language of the aforementioned evaluation methodology, the outcome variable Y is the number of doctor consultations in the previous quarter. The treatment group, $D = 1$, includes all non-exempt persons insured in the statutory health insurance. The control group, $D = 0$, includes the privately insured and exempt groups (young or low income families). The two time periods are $T = 1$: 1998 ; and $T = 0$: 1996. Finally, the vector x_{it} stands for all other socio-economic characteristics controlled for in the regression. These include a second order polynomial in age, three indicators for the quarter of the interview, four indicators for employment status plus the variables *years of education*, *married*, *household size*, *active sport*, *good health*, *bad health*, *private insurance*, *welfare recipient* (the coding of these variable is explained in an appendix). Deleting observations with missing values on any of the dependent or independent variables, this sample comprises 37319 observations.

Two sets of full regression results, one for the Poisson model and one for the panel Poisson model with gamma heterogeneity, denoted here as a "panel negbin" model, are displayed in Table 1. No differences in differences are used here, but rather a simple pre-post reform comparison, which makes these results directly comparable to those in Winkelmann (2003a). In the Poisson model the standard errors are adjusted to account for heteroscedasticity of

unknown form and correlation between observations for the same person. A formal likelihood ratio test clearly rejects the Poisson model against the panel negbin model. A comparison of the estimated standard errors suggests that some efficiency is gained indeed by estimating the panel model.

The estimated parameters differ somewhat between the two specifications, although they tend to give the same qualitative results. In particular both models estimate a drop in the expected number of doctor visits in the post-reform period of 9-10 percent. Many of the other results are common in the literature: men have fewer doctor visits than women. The health indicators have the largest effect among all variables. A person in bad health is estimated to have four times as many doctor visits as a person in good health. Interestingly, engaging actively in sports increases the number of visits once we control for general health.

The full differences-in-differences estimates are shown in Table 2. The purpose of these comparisons is to see whether the reform effect can be causally attributed to the increased co-payments (rather than other aspects of the reforms, or third aggregate influences unrelated to the 1997 reform). The strategy is to compute the change in expected visits before and after the reform separately for a treatment group and a control group. Since the data are not from a real experiment, the assignment to treatment and control has to be based on evidence available in the survey. Three comparisons are considered. All of them are based on count data regressions with the full set of other control variables included. Specifically, I compare

- the pre-post change for privately insured persons (control) versus non-exempt statutory

insured (treatment)

- the pre-post change for youth aged 16-18 (control, since exempted from co-payment) versus those aged 19 and above who are not otherwise exempt.

- the pre-post change for poor individuals (control, since presumably exempted from co-payment) versus others (excluding the privately insured and the youth).

The division into treatment and control is admittedly imperfect. Most problematic is the grouping based on welfare and low income status. For example, income is likely subject to considerable reporting error. Moreover, a fraction of all persons identified as belonging to the treatment may be exempt as well, because they have reached their maximum cumulative contribution by past co-payments during the year. Arguably, the first two comparisons are less error prone. In addition, Table 2 also shows estimates from an additional model, where all three exempt groups are combined into a single control group. In this last comparison, I have 3547 observations as controls and 33772 observations for the treatment.

The reform effect is labeled here "differences in differences". It is negative in all cases, and also statistically significant except in the panel negbin estimates with poor persons as control group. The combined effect, based on the panel negbin model, is a 13 percent reduction in the expected number of visits of the treatment group relative to the control group. I also show in the table the two components of the differences-in-differences estimates, namely the pre-post change for the control group and the pre-post change of the treatment group. In the combined sample, and again based on the panel negbin model, the expected number

of visits for the control group increased by 3 percent, although this change is insignificant. The expected number of visits of the treatment group decreased by 10 percent, so that the overall reform effect is the aforementioned reduction by 13 percent. This reasoning assumes that in the absence of the reform, the change in the treatment group would have been equal to the observed change for the control group.

7 Discussion

So where do we stand now? The re-evaluation of the evaluation suggests a much larger reform effect. Using the dd-methodology and data from the German Socio-Economic Panel, the causal effect of the increased co-payments is estimated to be a reduction in the order of 10-13 percent. This is significantly larger than the 4% reduction in the number of doctor visits reported in LGS.

To understand this discrepancy, one has to remember the very selected population of the LGS study, where only pharmacy customers were included. The discrepancy could be due to the fact that the health care reform of 1997 had a larger effect on rare users relative to frequent users. And this seems to have been the case indeed. One indication is the result reported in LGS that chronically sick people were less affected than others. Another is the study by Winkelmann (1993a) that shows that the reform had a particularily large effect at the left tail of the distribution of doctor visits. The two studies therefore do not necessarily contradict each other. Nevertheless, this example illustrates that the evaluation methods matters quite a lot, and that only the method based on a representative survey

of all potential users and on dd-techniques provides a good estimator of the causal reform effect for a typical insured person, as it is relevant from a policy perspective.

References

Gourieroux, C., A. Monfort and A. Trognon (1984) Pseudo Maximum likelihood methods: Applications to Poisson models, *Econometrica* 52: 701-721.

Lauterbach KW, Gandjour A, Schnell G. 2000. Zuzahlungen bei Arzneimitteln, Powerpoint presentation at http://www.medizin.uni-koeln.de/kai/igmg/stellungnahme/zuzahlungen/index.htm, version 25.2.2000.

Rubin, D.B. (1974) Estimating Causal Effects of Treatments in Randomized and Nonrandomized Studies, *Journal of Education Psychology* 66, 688-701.

SOEP Group (2001) The German Socio-Economic Panel (GSOEP) after more than 15 years - Overview. In: Holst E, Lillard DR, DiPrete TA (eds.): Proceedings of the 2000 Fourth International Conference of German Socio-Economic Panel Study Users (GSOEP2000), *Vierteljahrshefte zur Wirtschaftsforschung* 70: 7-14.

Winkelmann, R. (2003a) Health Care Reform and the Number of Doctor Visits - An Econometric Analysis, forthcoming in: *Journal of Applied Econometrics*

Winkelmann, R. (2003b) *Econometric Analysis of Count Data* 4th ed., Springer: Heidelberg.

Wooldridge, J. M. (2002) *Econometric Analysis of Cross Section and Panel Data*, MIT Press: Cambridge, Mass.

Table 1. Maximum likelihood results.
Dependent variable: Number of doctor visits during the previous quarter.

	Poisson	Panel Negbin
Age/10	0.082**	0.066**
	(0.033)	(0.027)
Age squared/1000	-0.069**	-0.011
	(0.033)	(0.028)
Male	-0.160**	-0.250**
	(0.021)	(0.017)
Years of schooling/10	-0.000	-0.027
	(0.040)	(0.035)
Married	0.078**	0.060**
	(0.024)	(0.016)
Active sport	0.060**	0.040**
	(0.020)	(0.011)
Good health	-0.602**	-0.479**
	(0.019)	(0.011)
Bad health	0.809**	0.648**
	(0.021)	(0.010)
Welfare recipient	0.054	0.027
	(0.047)	(0.025)
Logarithmic income	0.051**	-0.018
	(0.023)	(0.013)
Private health insurance	0.059	0.028
	(0.043)	(0.022)
Post reform	-0.097**	-0.091**
	(0.014)	(0.006)
log-likelihood	-101630	-81770

Source: *German Socio-Economic Panel*, years 1995, 1996, 1998 and 1999. Model includes furthermore a constant, three indicator variable for the quarter of the interview (winter, spring, fall), four indicators of employment status (full-time, part-time, self-employed, unemployed) and household size. Robust standard errors in parentheses. Coefficients with ** are significant at the 5% level. Coefficients with * are significant at the 10% level. N=37319.

Table 2. Differences-in-Differences Estimates for the Reform Effect

Change pre-post	Poisson model		Panel Negbin	
	effect	standard error	effect	standard error
Private insurance				
Control group (yes)	0.031	(0.065)	0.044	(0.029)
Treatment group (no)	-0.098**	(0.013)	-0.090**	(0.006)
Difference in Difference	-0.130*	(0.067)	-0.134**	(0.030)
Youth				
Control group (age 16-18)	0.117	(0.102)	0.118	(0.073)
Treatment group (age 19-)	-0.099**	(0.013)	-0.088**	(0.006)
Difference in Difference	-0.216**	(0.103)	-0.207**	(0.074)
Poor				
Control group (yes)	0.125	(0.114)	-0.019	(0.053)
Treatment group (no)	-0.112**	(0.014)	-0.102**	(0.007)
Difference in Difference	-0.238**	(0.115)	-0.082	(0.053)
Combined†				
Control group (yes)	0.056	(0.052)	0.027	(0.025)
Treatment group (no)	-0.112**	(0.014)	-0.102**	(0.007)
Difference in Difference	-0.168**	(0.055)	-0.129**	(0.027)

These estimates are based on count data regression models with the same controls as in Table 1.

Coefficients with * are significant at the 10% level. Coefficients with ** are significant at the 5% level.

† Combines the three exempt sub-populations (privately insured, young, poor) into a single control group.

Maximum number of observations: 37319.

APPENDIX

Timetable of Health Care Reforms in Germany

1969: Einführung der Lohnfortzahlung im Krankheitsfall, Krankengeld
1972: Einführung des dualen Finanzierungssystems für Krankenhäuser
1973: Gesetz zur Verbesserung von Leistungen in der GKV
1974: Übernahme der Kosten für Rehabilitationsmassnahmen durch die GKV
1977 - 1984: Kostendämpfungsgesetze
1989: Gesundheitsreformgesetz (GRG)
1993: Gesundheitsstrukturgesetz (GSG)
1997: Beitragsentlastungsgesetz (BeitrEntlG)
1997: 1. und 2. GKV-Neuordnungsgesetz (1./2. GKV-NOG)
1999: Gesetz zur Stärkung der Solidarität in der GKV (GKV-SolG)
2000: Gesundheitsreform 2000
2002: Gesetz zur Einführung von Chronikerprogrammen
2002: Gesetz zur Einführung von Fallpauschalen für Krankenhausleistungen
2003: Beitragssicherungsgesetz
2004: Modernisierung der GKV (bill passed in German Parliament on 26.9.2003)

Working Papers of the Socioeconomic Institute at the University of Zurich

The Working Papers of the Socioeconomic Institute can be downloaded from http://www.soi.unizh.ch/research/wp/index2.html

0311	Re-evaluating an Evaluation Study: The Case of the German Health Care Reform of 1997 Rainer Winkelmann, October 2003, 23 p.
0310	Downstream Investment in Oligopoly Stefan Buehler and Armin Schmutzler, September 2003, 33 p.
0309	Earning Differentials between German and French Speakers in Switzerland Alejandra Cattaneo and Rainer Winkelmann, September 2003, 27 p.
0308	Training Intensity and First Labor Market Outcomes of Apprenticeship Graduates Rob Euwals and Rainer Winkelmann, September 2003, 25 p.
0307	Co-payments for prescription drugs and the demand for doctor visits – Evidence from a natural experiment Rainer Winkelmann, September 2003, 22 p.
0306	Who Integrates? Stefan Buehler and Armin Schmutzler, August 2003, 32 p.
0305	Strategic Outsourcing Revisited Stefan Buehler and Justus Haucap, July 2003, 22 p.
0304	What does it take to sell Environmental Policy? An Empirical Analysis for Switzerland Daniel Halbheer, Sarah Niggli and Armin Schmutzler, 2003, 30 p.
0303	Mobile Number Portability Stefan Buehler and Justus Haucap, 2003, 12 p.
0302	Multiple Losses, Ex-Ante Moral Hazard, and the Non-Optimality of the Standard Insurance Contract Michael Breuer, 2003, 18 p.
0301	Lobbying against Environmental Regulation vs. Lobbying for Loopholes Andreas Polk and Armin Schmutzler, 2003, 37 p.
0214	A Product Market Theory of Worker Training Hans Gersbach and Armin Schmutzler, 2002, 34 p.
0213	Weddings with Uncertain Prospects – Mergers under Asymmetric Information Thomas Borek, Stefan Buehler and Armin Schmutzler, 2002, 37 p.
0212	Estimating Vertical Foreclosure in U.S. Gasoline Supply Zava Aydemir and Stefan Buehler, 2002, 42 p.
0211	How much Internalization of Nuclear RiskThrough Liability Insurance? Yves Schneider and Peter Zweifel, 2002, 18 p.
0210	Health Care Reform and the Number of Doctor Visits ? An Econometric Analysis Rainer Winkelmann, 2002, 32p.
0209	Infrastructure Quality in Deregulated Industries: Is there an Underinvestment Problem? Stefan Buehler, Armin Schmutzler and Men-Andri Benz, 2002, 24 p.
0208	Acquisitions versus Entry: The Evolution of Concentration Zava Aydemir and Armin Schmutzler, 2002, 35 p.
0207	Subjektive Daten in der empirischen Wirtschaftsforschung: Probleme und Perspektiven. Rainer Winkelmann, 2002, 25 p.
0206	How Special Interests Shape Policy - A Survey Andreas Polk, 2002, 63 p.

0205 Lobbying Activities of Multinational Firms
Andreas Polk, 2002, 32 p.

0204 Subjective Well-being and the Family
Rainer Winkelmann, 2002, 18 p.

0203 Work and health in Switzerland: Immigrants and Natives
Rainer Winkelmann, 2002, 27 p.

0202 Why do firms recruit internationally? Results from the IZA International Employer Survey 2000
Rainer Winkelmann, 2002, 25 p.

0201 Multilateral Agreement On Investments (MAI) - A Critical Assessment From An Industrial Economics Point Of View
Andreas Polk, 2002, 25 p.