Assessing the Effectiveness of

Health Care Cost Containment Measures[‡]

Nicolas R. Ziebarth
SOEP at DIW Berlin and TU Berlin[*]

December 23, 2010

[‡]I would like to thank Christian Boehler, Martin Karlsson, Jochen Kluve, Michael Kvasnicka, Therese Nilsson, Nigel Rice, Hendrik Schmitz, Matt Sutton, Tom Siedler, Gert G. Wagner, and participants in seminars at the 2008 Latin American Meeting of the Econometric Society (LAMES), the 2009 Health Economists' Study Group (HESG) Meeting in Manchester, the 2009 Annual Conference of the Royal Economic Society (RES), the 2009 Meeting of the European Economic Association (EEA), the 2009 conference of the European Association of Labour Economists (EALE), the SOEP Brown Bag, and the Berlin Network of Labour Market Researchers (BeNA) for their helpful comments and discussions. The paper will also be presented at the 2011 meeting of the American Economic Association (AEA) in Denver. I take responsibility for all remaining errors in and shortcomings of the article.
[*]German Institute for Economic Research (DIW Berlin), Socio-Economic Panel Study (SOEP), Mohrenstraße 58, 10117 Berlin, Germany, and Berlin University of Technology (TUB) `e-mail: nziebarth@diw.de`

Abstract

Using SOEP panel data and difference-in-differences methods, this study is the first to empirically evaluate the effectiveness of four different health care cost containment measures within an integrated framework. The four measures investigated were introduced in Germany in 1997 to reduce moral hazard and public health expenditures in the market for convalescent care. Doubling the daily copayments was clearly the most effective cost containment measure, resulting in a reduction in demand of about 20 percent. Indirect measures such as allowing employers to cut statutory sick pay or paid vacation during health spa stays did not significantly reduce demand.

Keywords: health expenditures, cost containment measures, copayment, convalescent care, SOEP

JEL classification: H51; I11; I18; J22

1 Introduction

For decades health expenditures have increased exponentially in most industrialized countries. In the US, health spending increased a staggering 787 percent between 1980 and 2007. In reunified Germany, health expenditures increased from 1992 to 2008 by 60 percent, consuming more than 10 percent of GDP in 2010 (German Federal Statistical Office, 2010). In light of these figures, it is no surprise that rising health care costs are one of the most contentious issues and a matter of great concern for policy makers worldwide.

Researchers have identified various key factors behind rising health expenditures, including demographic change, increasing national incomes, and technological change. Newhouse (1992) and Weisbrod (1991) are among the first to identify technological change as the dominant driving force, a conjecture that is difficult to empirically prove (Weisbrod, 1991; Newhouse, 1992; Okunade, 2004; Di Matteo, 2005; Civan and Koksal, 2010).

While the main causes of rising health expenditures seem clear, the question of how to deal with them remains unresolved. There is an extremely wide variety of organizing health care systems in different countries, but none of them have emerged with an optimal model. This comes as no surprise if one thinks about the very different objectives that the different health care systems are designed to achieve: reducing the burden on the social security system and taxpayers, achieving equal access to care, providing universal coverage, avoiding state rationing, allowing freedom to choose medical providers and insurance plans, or promoting medical progress, to name just a few.

The literature analyzes the optimal organization of health care theoretically as well as empirically, although the majority of work is theoretical in nature. Some attention is given to the supply side, particularly to the question of how to optimally organize and finance a hospital system with the aim of balancing quality of care against costs (Ellis and McGuire, 1996; Sloan et al., 2001; Propper et al., 2004; Bazzoli et al., 2008). Analogously, the same question can be raised for the outpatient sector and physicians (Mariñoso and Jelovac, 2003; Dusheiko et al., 2006; Karlsson, 2007). Especially in the US—a market still dominated by private health care providers—there is considerable debate surrounding the question of whether Health Maintenance Organizations (HMOs) can help reduce health expenditures while maintaining quality (Goldman et al., 1995; Hill and Wolfe, 1997; Keeler et al., 1998; Deb and Trivedi, 2009). In Europe, on the other hand, key concerns revolve around issues of direct rationing (by public authorities) and indirect rationing (through waiting times) (Propper et al., 2002; Schut and de Ven, 2005;

Felder, 2008; Siciliani et al., 2009).

In the demand-side research, cost-sharing is identified as the main tool used to reduce moral hazard and overconsumption of medical services (Pauly and Blavin, 2008; van Kleef et al., 2009). In this strand of the literature, the RAND Health Insurance Experiment (HIE) is still the largest and most influential health policy study to this day. In this 1970s era study, families at six different sites in the US were randomly assigned to 14 different health insurance plans with a varying degree of cost-sharing and observed for periods up to five years (Manning et al., 1987). Since then, a great amount of publications on the impact of cost-sharing on the demand for medical care emerged from the HIE, most published in the 1980s (see Zweifel and Manning (2000) for an overview). But apart from the HIE, there is only scant empirical evidence of causal effects of cost-containment measures on the demand for health care. A handful of studies empirically investigate how increased copayments affect the demand for doctor visits (Chiappori et al., 1998; Voorde et al., 2001; Cockx and Brasseur, 2003; Winkelmann, 2004; Gerfin and Schellhorn, 2006). Schreyögg and Grabka (2010) analyze the effects of the copayments for doctor visits introduced in Germany. Using a difference-in-differences setup, similar to the one in this study, as well as the same dataset, they do not find any significant behavioral reactions in the aftermath of the reform.

To the best of my knowledge, this is the first paper evaluating the effectiveness of four different cost containment measures within an integrated framework. In Germany, beginning in 1997, various health reforms were implemented to reduce the demand for convalescent care. Before the reforms went into effect, experts claimed that around a quarter of all convalescent care therapies were unnecessary (Schmitz, 1996; Sauga, 1996). In 1995, 1.9 million patients in Germany underwent convalescent care therapy and more than €7 billion (0.4 percent of GDP) was spent on these programs (German Federal Statistical Office, 2010). Ziebarth (2010) shows that the price elasticity of demand for convalescent care treatments is inelastic and about -0.4—an estimate that is very much in line with the consensus price elasticity estimates in the literature on health care (Wedig, 1988; Keeler et al., 1988; Zweifel and Manning, 2000). Hence, it is plausible to assume that convalescent care is a good proxy for health care in general.

The first reform evaluated here doubled the daily copayments for convalescent care. The second increased waiting times between two treatments and reduced the legally codified standard length of the therapy. The third reform gave employers the right to deduct two days of paid vacation for every five days that employees were unable to work while in convalescent care. The fourth reform cut statutory sick pay from 100 to 80 percent of foregone gross wages during convalescent care.

The first two reforms only affected people insured under the German Mandatory Health Insurance (MHI), while people insured under the second tier of the German health insurance system—the Private Health Insurance (PHI)—were not affected. The other two reforms, which concerned the cut in paid leave, only affected private-sector employees. Thus, I can define various subgroups that were affected differently by the reforms. By means of conventional difference-in-differences models and SOEP panel data, I then disentangle the causal effects of these cost containment measures on the demand for convalescent care. One main objectives of this paper is to evaluate the effectiveness of direct cost containment measures such as copayment increases, which apply to the entire population, as compared to indirect measures such as decreasing legal minimum requirements, which only increase employers' options to regulate work conditions at the firm level.

My empirical results show that doubling the copayments was, by far, the most effective cost containment instrument. It led to a significant decrease in the demand for convalescent care programs of about 20 percent. Moreover, evidence from administrative data suggests that the reduction in the legally defined standard length of the therapies was effective in reducing the average duration of treatments. However, I do not find evidence that the cuts in paid leave significantly reduced the demand for convalescent care programs.

Based on administrative data, back-of-the-envelope calculations suggest that all reforms jointly reduced annual public spending for convalescent care by €800 million or 13 percent. Although the length of treatments decreased, the doubling of daily copayments raised additional revenues of about €400 million per year.

In the next section, I describe some features of the German health care system and give more details about the reform. In Section 3, the dataset and the variables used are explained, and in the subsequent section, I specify my estimation and identification strategy. Estimation results are presented in Section 5 and I conclude with Section 6.

2 The German Health Care System and the Policy Reforms

The German health care system is actually comprised of two independent health care systems that exist side by side. The more important of the two is the Mandatory Health Insurance (MHI), which covers about 90 percent of the German population. Employees whose gross income from salary is below a defined income threshold (in 2010, €49,950 per year) are compulsorily insured under the MHI. High-income earners who exceed that threshold as well as self-employed people have the right to choose between the MHI and private health insurance. Non-working spouses and dependent children are covered at no cost by the MHI family insurance. Special regulations apply to particular groups such as students and the unemployed, but most of these are MHI-insured. Everyone insured under the MHI is subject to a generous universal benefit package, which is determined at the federal level and codified in the Social Code Book V (SGB V). Coinsurance rates[1] are prohibited in the MHI and thus, apart from copayments, health services are fully covered. The MHI is one pillar of the German social security system (German Ministry of Health, 2010).

The MHI is primarily financed by mandatory payroll deductions that are not risk-related. For people with gainful employment, these contributions are split equally between employer and employee up to a contribution ceiling (2010: €45,000 per year). Despite several health care reforms that tried to remedy the problem of rising health care expenditures, contribution rates have risen from 12.6 percent in 1990 to 15.5 percent in 2011, mainly due to demographic changes, medical progress, and system inefficiencies (German Federal Statistical Office, 2010).

The second track of the German health care system is Private Health Insurance (PHI). The main groups of private insurance holders are private-sector employees above the aforementioned income threshold, public-sector employees[2], and the self-employed. Privately insured people pay risk-related insurance premiums determined by an initial health checkup. The premiums exceed the expected expenditures in younger age brackets, since health insurance providers build up reserves for each insured person for rising expenditures with increased age. Coverage is provided under a range of different health plans, and insurance contracts are subject to private law. Consequently, in Germany, public health care reforms apply only to the MHI, not to the PHI.

[1] Coinsurance rates are important for private health insurance providers. They differ from copayments. While a copayment is typically a fixed amount that the insured person has to pay per day of treatment or for specific medical devices or medications, a coinsurance rate defines a percentage of the costs that an insured person has to pay when using the system. For example, private health insurance providers may offer 80/20 health plans in which the insured person pays 20 percent of all costs incurred while the health insurance provider pays the remaining 80 percent. Often, health insurance providers limit the total amount that an individual has to spend out-of-pocket with a so-called coinsurance cap, which might be €2,000 per year.

[2] We need to distinguish between two types of employees in the German public sector: first, civil servants with tenure (*Beamte*), henceforth called "civil servants," most of whom purchase PHI to cover the 50 percent of health expenditures that the state does not reimburse (*Beihilfe*), and second, employees in the public sector without legal tenure (*Angestellte im öffentlichen Dienst*), henceforth called "public servants," who receive some additional benefits but are mainly insured under the MHI (under the same conditions as everyone else).

It is important to keep in mind that compulsorily insured persons have no right to choose the health insurance system or benefit package. They are compulsorily insured under the standard MHI insurance scheme. Once an optionally insured person (a high-income earner, self-employed person, or civil servant) opts out of the MHI system, it is practically impossible to switch back. Employees above the income threshold are legally prohibited from doing so, while those who fall below the income threshold in subsequent years may do so under certain conditions, but any reserves that they have built up under PHI policies are not transferable (neither between PHI and MHI, nor between different private health insurance providers).[3] In reality, switching to a private health insurance provider may be regarded as a lifetime decision, and switching between the MHI system and PHI—as well as between PHI providers—is therefore very rare.

2.1 The German Market for Convalescent Care

In Europe, especially in Germany, there is a long tradition of health spa treatments to improve poor health. Since the time of the Roman Empire, doctors have sent patients to "take the waters" to recover from various disorders. In Germany, convalescent care treatments are usually combined with various types of physical therapy, often including electrotherapy, massage, underwater exercise, ultrasonic therapy, health and diet education, stress reduction therapy, and cold and hot baths as well as mud packs. Convalescent care therapies require the patients to follow a strict daily schedule.

The German MHI is one of the few health insurance systems worldwide that, apart from small copayments, fully covers convalescent care therapies at health spas. It may therefore come as no surprise that the German market for convalescent care is said to be the largest worldwide. In 1995, a total of €7.646 billion was spent on convalescent care, accounting for more than 4 percent of all health expenditures in Germany. Around 1,400 medical facilities with 100,000 full-time (equivalent) staff members treated 1.9 million patients, who stayed an average of 31 days each (German Federal Statistical Office, 2010).

Convalescent care therapy—referred to in Germany as a *Kur* or cure—requires a physician's prescription, and the individual must submit an application for treatment to his or her MHI sickness fund. The role of the patient in the application process is central. On the one hand, well informed patients may push their doctors to recommend them for convalescent care, and doctors may comply simply out of the fear of losing patients given the competition on the

[3] Until 2009, accrued reserves for rising health expenditures with increased age were not transferable at all. But since January 1, 2009, a strictly defined level of transferability between PHI providers is compulsory.

market and free choice of doctors for those insured under the MHI. On the other hand, patients may not accept their doctor's recommendation for convalescent care. After the application, the MHI fund determines whether the preconditions for treatment are fulfilled and authorizes the therapy. The wording of the preconditions can be found in the German social legislation, Social Code Book V (SGB V, article 23 para. 1, article 40, para. 1). After authorization by the MHI sickness fund, the prescribed treatment is provided in an approved medical facility under contract with the MHI fund. These medical facilities are usually located in scenic rural villages licensed by the state as *Kurorte*, or spa towns. For a village to be granted such a license, it needs to fulfill several conditions established in state legislation: pure air and location near the seaside or mineral springs. The idea of providing patients a healthy change of environment is integral to the treatment program.

2.2 The Cost Containment Policy Reforms

At the end of 1996, the German government implemented four health care reforms. The first three were designed to dampen the demand for convalescent care programs directly, based on the suspicion of a high degree of moral hazard in the market for convalescent care. Prior to the reform, experts estimated that around a quarter of all treatments prescribed were unnecessary (Schmitz, 1996; Sauga, 1996). The fourth reform was designed to tackle moral hazard in the decision to take sick leave and may have indirectly affected the demand for convalescent care as well.

The first reform doubled daily copayments. In West Germany, as of January 1, 1997, copayments for convalescent care therapies increased from DM 12 (€6.14) per day to DM 25 (€12.78) per day. In East Germany, the copayments increased from DM 8 (€4.09) to DM 20 (€10.23) per day. This reflects an increase of 108 (150) percent.[4] To illustrate how drastic this copayment increase really was, I multiply the daily copayment rates by the average length of stay according to the Federal Statistical Office (German Federal Statistical Office, 2010). The absolute increase per treatment amounted to around €150 in East and West Germany. Before the reform and in relation to the monthly net wages of those who received convalescent care in my sample, the total copayment per treatment was 12 percent of the monthly net wage in East Germany and 13 percent in West Germany. After the copayment increase, the total copayment sum per treatment approximately doubled to 25 (East) and 24 (West) percent of the average

[4] Passed on November 1, 1996, this law is the *Gesetz zur Entlastung der Beiträge in der gesetzlichen Krankenversicherung (Beitragsentlastungsgesetz - BeitrEntlG), BGBl. I 1996 p. 1631-1633.*

monthly net wage.

The second reform reduced the standard length of convalescent care therapies from four to three weeks. Only the medical personnel of the facility—after consultation with the sickness fund—have the authority to approve deviations from the standard, legally codified, length of therapy. Together with this reduction in therapy duration, waiting times were increased from three to four years between treatments. Both reform elements—the reduced standard length of therapy and the extended waiting period—are only effective conditional on the non-existence of urgent medical reasons for treatment.

The third reform allowed employers to deduct two days of paid vacation for every five days that an employee was unable to work due to convalescent care therapy. The fourth reform decreased statutory short-term sick pay from 100 to 80 percent of foregone gross wages. German social legislation provides employees with paid leave for convalescent care treatments *in addition* to paid vacation. Hence, one would expect that the latter two reforms, which allowed employers more leeway in reducing paid leave, to have an effect on the demand for convalescent care.[5]

[Insert Table 1 about here]

Table 1 displays the various subgroups of insured people who were affected differently by the four cost containment measures. Subgroup (1) comprises the vast majority of Germans: private-sector employees who are insured under the MHI. They were affected by all reforms discussed above. I define them as *Treatment Group 1*.

In contrast, subgroups (2) to (5) were not affected by either the cut in statutory sick pay or the cut in paid vacation. Non-working and self-employed people are not eligible for paid leave. Public-sector employees and apprentices were exempted from the cuts in paid leave. However, since they were insured under the MHI, they were affected by the first two reforms. I call these subgroups jointly *Treatment Group 2*.

Subgroups (6) to (9) were completely unaffected by all legislative changes; they also consist of the non-working, the self-employed, apprentices, and public sector employees, none of whom were affected by the cut in paid leave. However, in contrast to *Treatment Group 2*, subgroups (6) to (9) were insured under the PHI and, thus, reforms one and two did not apply to them either. I define subgroups (6) to (9) jointly as *Control Group*.[6]

[5] Passed on September 15, 1996, this law is the *Arbeitsrechtliches Gesetz zur Förderung von Wachstum und Beschäftigung (Arbeitsrechtliches Beschäftigungsförderungsgesetz), BGBl. I 1996 p. 1476-1479*. The law went into effect on October 1, 1996.

[6] Private-sector employees insured under the PHI are not included in my working sample. They were *only*

In total, I obtain three mutually exclusive subsamples that were affected differently by the reforms. Thus, in the empirical assessment, I use three distinct main models in which I compare these subsamples to evaluate the effectiveness of the reforms. To this end, I generate three treatment indicators that I will explain in more detail in Section 3.3 below.

3 Dataset and Variable Definitions

3.1 Dataset

The empirical analysis relies on micro data from the German Socio-Economic Panel Study (SOEP). The SOEP is an annual representative household survey that started in 1984 and meanwhile includes more than 20,000 respondents. Wagner et al. (2007) provide further details. Information on convalescent care treatments is only available for two post-reform years. Hence, for the core analyses, I use data from the 1995 to 1999 waves, which include time-invariant information, current information, and retrospective information about the previous year. Since the dependent variable contains information about the calendar year prior to the interview, I employ data on the years 1994 to 1998.[7]

I exclude respondents under the age of 18, who are exempted from copayments, and focus on the subgroups defined in Table 1.

3.2 Dependent Variable and Covariates

The SOEP contains various questions about health insurance and the use of health care services. The dependent variable *convalescent care* measures whether a respondent received convalescent care at a health spa in the calendar year prior to the interview; it takes the value one if that was the case, and zero if not. In other words, *convalescent care* measures the overall incidence of convalescent care programs. The variable has been generated from the following question, which was asked in every wave from 1995 to 1999: "Did you go to a health spa for convalescent care in 199X?" In German, this question is even clearer because of the well-known umbrella term *Kur* and the inpatient treatment this entails, at a location other than the recipient's place

affected by reforms three and four but not by the increase in copayments or in waiting times. However, in my sample, they consist of only 150 respondents per year and, thus, I cannot use them to obtain precise estimates.

[7] If the respondent was interviewed in two subsequent waves, e.g., in 1994 and 1995, I match time-variant data from questions posed in the first year dealing with the first year with retrospective data obtained from questions posed in the second year dealing with the first year. For example, in 1995, respondents were asked about their current health status and about their insurance status during the previous year. Hence, I use the 1994 data on health status together with the 1995 data on insurance status if the respondent was interviewed in both years.

of residence, a *Kurort* or spa town, which minimizes measurement errors. The fact that we do not know the exact period of the therapy does not severely hamper the analysis, especially since such treatments are usually not carried out over Christmas or New Year's. Hence, there should be no doubt as to whether the therapy was in 1996 or in 1997.

While *convalescent care* can be considered a fairly good measure of the incidence of convalescent care treatments, the SOEP does not include a measure of their duration. However, as explained above, the length of treatment is regulated by social law and deviations from it are solely determined by the medical personnel and the MHI sickness fund, not by the patient. Therefore, the empirical analysis focuses mainly on the effects on the incidence, which is the key behavioral parameter in this setting and mainly influenced by the patient. I use aggregated administrative data on the average duration of treatments as an additional outcome measure in descriptive assessments later on.

In my main empirical models, I make use of various control variables. These control variables capture personal and family-related characteristics such as *age*, *female*, *immigrant*, *partner*, and *children*. Moreover, I control for educational characteristics by using data on the highest educational degree obtained. An important determinant of the demand for convalescent care programs is the health status of the respondents, which I observe and control for (in form of self-assessed health). I also include covariates that measure whether the person was employed full-time, part-time, marginally, or not at all.[8] I additionally control for gross monthly income. To capture time-invariant regional characteristics, I make use of 15 state dummies. Regional labor market dynamics are controlled for by the inclusion of the annual state unemployment rate. Time trends are captured by year dummies. A list of the covariates, as well as its means and standard deviations are found in Appendix A.

3.3 Treatment Indicators

In Section 2.2, I define three mutually exclusive subsamples that were affected by different reform elements, as shown in Table 1. In the next section, I make use of three distinct models to assess the effectiveness of the various reforms. This requires three distinct treatment indicators for the three models to compare the different subsamples.

T1 has a one for employees in *Treatment Group 1* and a zero for respondents in the *Control*

[8] Non-employment in particular may change quickly. Hence, the assignment of respondents to the treatment and control group might be imprecise. Since the MHI/PHI status is very stable over time, the imprecision lies between the different subgroups that were insured under the MHI as well as between the different subgroups that were insured under the PHI.

Group. By using this treatment indicator in *Model 1*, I compare those who were affected by all reforms with those who were completely unaffected to assess the net effect of all reforms jointly on the demand for convalescent care programs.

T2 has a one for employees in *Treatment Group 2* and a zero for respondents in the *Control Group*. Thus, in *Model 2*, I contrast those who were affected by the first two reforms with the non-treated. In this model, my main intention is to evaluate the effectiveness of the copayment doubling, i.e., the first reform. In extended robustness checks, I will also assess the effect of the second reform by means of *Model 2*.

T3 is used in *Model 3*, which assesses the effectiveness of the cuts in paid leave. For this purpose I compare *Treatment Group 1* with *Treatment Group 2*. Thereby I extract the effect of the first two reforms from the net reform effect to obtain the effect of the cuts in paid leave.

4 Estimation Strategy

4.1 Difference-in-Differences

I would like to measure how each reform affected the incidence of convalescent care programs. Thinking of the policy intervention as a treatment, I fit probit models of the form:

$$P[y_{it}=1|\boldsymbol{X_{it}}] = \Phi(\alpha + \beta post97_t + \gamma T_{it} + \delta \underbrace{(post97_t \times T_{it})}_{DiD_{it}} + \boldsymbol{s'_{it}}\boldsymbol{\psi} + \rho_t + \phi_s + \epsilon_{it}) \tag{1}$$

where y_{it} stands for the incidence of convalescent care programs, *convalescent care*. $post97_t$ is a dummy that takes on the value one for post-reform years and zero for pre-reform years. Depending on the model, T_{it} stands for one the three treatment indicators (see Section 3.3 above). The interaction term between the two dummies gives us the difference-in-differences (*DiD*) estimator. To evaluate how the reform affected the outcome variable y_{it}, henceforth, I always compute and display the marginal effect of the interaction term $\frac{\Delta\Phi(.)}{\Delta(post97\times T)}$.[9] $\Phi(.)$ is the cumulative distribution function for the standard normal distribution. By including additional time dummies, ρ_t, I control for common time shocks. State dummies, ϕ_s, account for permanent differences across the 16 German states along with the annual state unemployment rate that

[9] Puhani (2008) shows that the advice of Ai and Norton (2004) to compute the discrete double difference $\frac{\Delta^2\Phi(.)}{\Delta post97\Delta T}$ is not relevant in nonlinear models when the interest lies in the estimation of a treatment effect in a difference-in-differences model. Using treatment indicators, the average treatment effect on the treated is given by $\frac{\Delta\Phi(.)}{\Delta(post97\times T)} = \Phi(\alpha + \beta \text{post97} + \gamma \text{T} + \delta \text{DiD} + \boldsymbol{s'}\boldsymbol{\psi} + \rho + \phi + \epsilon) - \Phi(\alpha + \beta \text{post97} + \gamma \text{T} + \boldsymbol{s'}\boldsymbol{\psi} + \rho + \phi + \epsilon)$ which is exactly what is calculated and presented throughout the paper.

controls for changes in the tightness of the regional labor market and that is included in the $K \times 1$ column vector $\boldsymbol{s_{it}'}$. The other $K-1$ regressors are made up of personal controls including health status, educational controls, and job-related controls as explained in Section 3.1.

4.2 Identification

The identification strategy relies on DiD estimation and hence on the assumption of a common time trend of the outcome variable for treatment and control group in the absence of the policy intervention. This assumption should hold conditional on all available covariates. In almost all natural experiments and non-randomized settings, controlling for a rich set of covariates is important since the control and treatment groups differ with respect to most of the observed characteristics. This is also true in the present case, as Table 2 shows. For example, in comparison to the *Control Group*, *Treatment Group 1* includes more females and immigrants, and the employees are less educated. As compared to the *Control Group*, the people in *Treatment Group 2* are younger and more likely to be full-time employed.

[Insert Table 2 about here]

As can be seen in Table 3, the most important driver of the demand for convalescent care programs is health status. Not surprisingly, age also plays a role, as well as income. Immigrants are less likely to receive convalescent care, probably because of information asymmetries.

[Insert Table 3 about here]

Again, I would like to stress that the econometric specifications adjust the sample composition to the various personal, educational, and job-related characteristics of the respondents. Recall that the health status of the respondents is observed and controlled for. Likewise, adjustments are made for time effects, persistent differences between states, and the annual state unemployment rate.

The key identifying assumption, the common time trend assumption, is likely to hold. It assumes the absence of unobservables that generate different outcome *dynamics* for the treatment and control group. It is worth mentioning that a selection on observables story is very plausible in the present setting. In the first place, it is the MHI/PHI insurance status that determines treatment (see Table 1). Almost all factors that determine whether respondents are insured under the MHI or PHI—such as occupational status and income—are observed.

A method to check the absence of distorting unobservable effects is to estimate placebo regressions for years without a reform. I make use of this method in the next section.

Figure 1 shows the evolution of the outcome variable for *Treatment Group 1, 2* and the *Control Group* over time.[10] Even without the correction for observables, we observe parallel evolution in the three groups during the pre-reform years. After the reform, the incidence of convalescent care programs in the control group remained fairly stable, whereas we observe a clear, distinct, and parallel decrease for the treatment groups.

[Insert Figure 1 about here]

Compositional changes within the treatment and control groups might have an impact on the outcome variable. For example, in *Treatment Group 2*, the share of self-employed or public-sector employees may change over time, which might affect or even produce the trend in the outcome variable. However, the share of self-employed people within *Treatment Group 2* only fluctuated between 5.31 percent and 5.86 percent between 1994 and 1998. The other subgroups showed similar fluctuations, also remaining very stable over time.

The drawbacks and limitations of DiD estimation are extensively debated. A particular concern is the underestimation of OLS standard errors due to serial correlation in the case of long time horizons as well as unobserved (treatment and control) group effects (Bertrand et al., 2004; Donald and Lang, 2007; Angrist and Pischke, 2009). To address the serial correlation issue, we focus on short time horizons. In addition, to provide evidence on whether unobserved common group errors might be a serious threat to our estimates, in robustness checks, we cluster on the state$\times$year ($16\times5 = 80$ clusters) level (Angrist and Pischke, 2009).

A crucial issue in most studies trying to evaluate policy reforms is, besides the absence of a control group, selection into or out of the policy intervention. Selection issues are addressed since I am in the unusual position of having a framework in which two almost totally independent health care systems exist side by side, as explained in Section 2. On the one hand, this provides a well-defined control group. On the other hand, I do not need to fear that reform-induced selection has distorted the results, as there is virtually no switching between the MHI and the PHI, and since all MHI-insured persons are covered by universal health plans. Due to strict German regulations, a switch to the PHI was only legally allowed for a small fraction of optionally MHI-insured individuals, and I am able to identify and exclude these cases when

[10] As shown later, there is evidence that distorting effects play a role due to the announcement of the reform at the end of 1995. Hence, the two uncontaminated pre-reform years, 1994 and 1995, are contrasted to the two post-reform years, 1997 and 1998.

running robustness checks. In my dataset, only 1.6 percent of those who were insured under the MHI for at least one year switched to the PHI between 1994 and 1998. The rate did not increase after the reform. Only 1.3 percent of those who were insured under the MHI in 1995 switched to the PHI in 1997 or 1998.

We need to consider the possibility of pull-forward effects. Convalescent care programs are usually planned several months or even years in advance. Since the first policy reform plans were made public at the end of 1995 (Handelsblatt, 1995), it may be that a significant portion of the MHI-insured received their convalescent care therapy in 1996 instead of 1997. In the empirical application, I check for anticipation effects.

Admittedly, it may have been that, due to rising awareness and increased political pressure, the MHI funds were more restrictive in their authorization of therapy programs during the period when the reforms were under political discussion, i.e., in 1996. As for anticipation effects that might have been triggered by the insured, one can test for such effects by either excluding the year 1996 from the analysis or by adding an interaction term between 1996 and the treatment indicator to the analysis.

To be able to fully attribute changes in the incidence to changes in the demand for convalescent care programs, supply-side effects should not play a role. I found no indications of supply-side constraints. In contrast, there are reports about the deepest crisis in the market for convalescent care since the end of the Second World War (Handelsblatt, 1998). Dozens of medical facilities and health spas closed and, hence, there is strong evidence of excessive supply. This is also supported by official statistics stating that the utilized bed capacity of all facilities strongly decreased, from 83.2 percent in 1996 to 62.3 percent in 1997 (German Federal Statistical Office, 2010).

Individuals insured under the MHI who were for some reason exempted from copayments are not identifiable. For example, people whose annual copayments for pharmaceuticals, health care services, or medical devices exceeded a certain percentage of their disposable household income could have applied for a case of hardship.[11] However, at that time, the *German Spa Association* claimed that the public was widely unaware of the exemption clauses. Therefore this should not downwardly bias the results severely.

As mentioned in Section 2.2, the third reform allowed employers to deduct two days of paid vacation for every five days that an employee was absent from work due to convalescent care

[11] The usual threshold is 2 percent of disposable household income; for people with chronic diseases it is 1 percent.

therapy. The fourth reform cut statutory (short-term) sick pay. In contrast to the other reforms, these two reforms are rather indirect cost containment measures since they decreased the statutory minimum standards. Since employers are always free to provide fringe benefits on top of statutory requirements, reforms three and four simply increased employers' capacity to act. I cannot observe which employers enforced these reforms strictly and passed on the decrease in social law minimum standards one-to-one to their employees. Anecdotal evidence and polls suggest that this might have been the case for about 50 percent of all potentially treated, i.e., private-sector employees (Ridinger, 1997; Jahn, 1998). Using all private-sector employees jointly as treatment group, Ziebarth and Karlsson (2010) show that the cut in statutory short-term sick pay significantly reduced absenteeism. Since I apply the same approach in this setting, I should be able to identify potential reform effects. Indeed, one of the main objectives of this paper is precisely to evaluate the effectiveness of direct cost containment measures such as co-payment increases, which apply to the entire population, as compared to indirect measures such as decreasing legal minimum requirements, which only increase employers' options to regulate work conditions at the firm level.

As a last point, it should be kept in mind that the identification strategy for the difference-in-differences regression models is based on various specifications. In total, I estimate three distinct models, each of which compares different mutually exclusive and differently affected subsamples. In addition, I run various robustness checks, which enables me to automatically cross-check the consistency and plausibility of the reform effects identified.

5 Results

Assessing the reforms' effectiveness

Table 4 shows the results for *Model 1, 2*, and *3*. For each model, I display the "raw" difference-in-differences (DiD) estimate as well as the estimates obtained from a Probit and an OLS specification with the full set of covariates. The raw estimate represents what we see in Figure 1, which displays the unconditional trends for the various subsamples over time. All models in Table 4 use an unbalanced panel, and each column represents one DiD model. *DiD* always stands for the DiD estimate.

Model 1 makes use of the treatment indicator *T1* and compares the pre-post-reform outcome difference for *Treatment Group 1* to the pre-post-reform outcome difference for the *Control*

Group. Since *Treatment Group 1* was affected by all four cost containment measures and the *Control Group* by none, *Model 1* estimates the net effect of all reforms on the incidence of convalescent care programs. Column (1) gives the raw estimate, column (2) the Probit estimate, and column (3) the OLS estimate under the inclusion of all covariates.

All three estimates for *Model 1* yield significantly negative reform effects on the incidence of convalescent care programs. Moreover, all three estimates are fairly robust and lie within the same confidence intervals. The Probit and the OLS estimates in columns (2) and (3) are especially close to one another, which suggest that functional form assumptions do not seem to matter here. The pre-reform incidence of convalescent care programs for *Treatment Group 1* is 0.0355, i.e., 3.55 percent. Relating the percentage point estimate (-0.0081) from my preferred specification in column (2) to this pre-reform incidence rate suggests that all reforms jointly decreased the demand for convalescent care programs by 22.8 percent.

[**Insert Table 4 about here**]

Model 2 disentangles the effects of reforms one and two from the effects of reforms three and four. Reform 1 doubled the daily copayments for convalescent care treatments. Reform 2 reduced the legally codified standard length of the therapy and increased the waiting times between two therapies. Reform 3 cut statutory sick leave, while Reform 4 cut paid vacation in case of work absences due to convalescent care treatments. *Model 2* contrasts those who were affected by reforms one and two (*Treatment Group 2*) with those who were completely unaffected by all health reforms (*Control Group*). It employs the treatment indicator *T2*.

Again, all three estimates are similar in magnitude: all are negative and significantly different from zero, they are insensitive to the inclusion of covariates, and the results from the OLS and Probit models barely differ. All *DiD* point estimates fall within the same confidence intervals than the ones in *Model 1*. The average pre-reform convalescent care incidence for *Treatment Group 2* was 0.0502, and hence the -0.0136 percentage point estimate of the Probit model in column (5) translates into a reform-induced decrease of about -27.1 percent. This suggests that reforms one and two are responsible for the decrease in demand for convalescent care programs. In the robustness checks below, I provide evidence that the copayment doubling is probably responsible for the bulk of this decrease. My findings below suggest that the increase in waiting times did not contribute much to the decrease and that the legally codified reduction in the standard length of treatments merely reduced the average duration of treatments.

Model 3 compares those affected by all four reforms (*Treatment Group 1*) to those affected by reforms one and two (*Treatment Group 2*). I thereby assess the effects of reforms three and four jointly, i.e., the cuts in paid leave. The results of *Model 3* strongly confirm the findings of *Model 1* and *Model 2*: columns (7) to (9) of Table 4 all yield point estimates that are very close to zero and not statistically different from zero. The point estimates are even positive and the standard errors are fairly tight. All in all, I do not find any evidence that the cuts in paid leave induced any significant reduction in the demand for convalescent care programs. I have two explanations for this finding. First, the cut in vacation days may not have been a binding constraint, since many employees use all or part of their paid vacation for convalescent care. Although entitled to take paid leave *in addition* to their paid vacation, many employees fear negative job consequences, especially when unemployment rates are high. Second, the cut in sick pay did not necessarily impose any limitation on the insured since their decision may have been between either going to a convalescent care facility or simply staying home to recover. In any case, the patient would have been on sick leave. If necessary, physicians usually recommend treatments in spa towns, but if patients prefer to stay home on sick leave, their wishes are usually respected.

The entire setup and the fact that all results are based on a comparison of three mutually exclusive subsamples gives rise to another means of calculating the effects for *Model 2* and the first two reforms: one can simply subtract the estimates from *Model 3* from those from *Model 1*, i.e., subtract the effects of reforms three and four from the net effect of all reforms. It is easy to see that this exercise yields very consistent alternative estimates for *Model 2* that are almost identical to the direct estimates in columns (5) and (6) (0.0097 for the Probit and 0.0141 for the OLS model).

Robustness checks

Table 5 displays various robustness checks. In all cases, I focus on *Model 2* and the Probit specification with all covariates included.[12]

The first column of Panel A is the same estimate as the one in column (5) of Table 4 (-0.0136). This is the "standard" estimate. Column (2) excludes 1996 from the specification. Since the copayment doubling was first announced in December 1995, is it likely that pre-reform 1996 is

[12] Here I focus on *Model 2* since it includes more observations than *Model 1* and therefore yields a more precise estimation. Moreover, as such, I am able to run checks on the effectiveness of Reform 2. The results for *Model 1* are very similar and available upon request from the author.

contaminated by either pull-forward effects triggered by the insured or by supply-side effects triggered by MHI sickness funds or the SPI. The MHI and SPI might have been more restrictive in the authorization of treatments due to rising public awareness and political pressure. Indeed I evidence of this. Omitting 1996, the DiD estimate shrinks slightly and translates into a decrease of about 21 percent in demand. Column (3) also supports this result, since the short-run reform effect obtained by comparing 1996 to 1997 is larger than the standard estimate in column (1) and -0.018.[13]

Reform 2 increased the waiting period between treatments for MHI-insured from three to four years. The last column in Panel A tests whether the increase in waiting times reduced the incidence of convalescent care programs in the short run. The extension of the waiting period did not apply to individuals needing urgent medical treatment. As detailed in Section 2, people insured under the MHI have free choice of doctors, and there are almost no waiting times for doctor appointments in Germany. Thus, it is unlikely that the increase in waiting times had a substantial effect, since finding a doctor to write a prescription for treatment is not usually difficult. The increase in the waiting period forced patients who received treatment in 1994 (1995) to wait until 1998 (1999) instead of 1997 (1998) in the absence of urgent medical reasons. Thus, if the increased waiting period had a substantial impact, I would measure a stronger reform effect for 1997 than for 1998. Column (5) of Table 5 shows that the reform effect in 1998 was not stronger than in 1997. I take this as evidence that the increased waiting period had no significant (short-term) effect on the demand for convalescent care.

The second element of Reform 2 was the reduction of the standard length of convalescent care from four to three weeks. The standard length is codified in the Social Code Book and applies universally to everyone who is insured under MHI. Exceeding the standard length is only possible in case of urgent medical reasons. The decision to deviate from the legally codified standard length can only be made by the attending physician after consulting the sickness fund to authorize the prolongation. Since the SOEP does not include information on the length of therapy, I cannot estimate the effect of the reduction in the standard length using a regression model. However, official data is available on the average treatment length and the total number of days spent in inpatient medical facilities for convalescent care treatments. These official data represent average values for the whole of Germany. According to these data, the average treatment length for all insured individuals decreased by almost 4 days from 31.0 (30.2) days in 1995 (1996) to 27.3 (26.4) days in 1997 (1998) (German Federal Statistical Office, 2010). The

[13] Please note that this is not true in a strict statistical sense since the confidence intervals overlap.

figures provide evidence that reducing the legally codified standard length was an effective tool to reduce the real length of treatments. On the other hand, it is unlikely that reducing the legal standard length of therapies had a substantial impact on the *incidence* of convalescent care therapy, i.e., on the decision to go to a health spa. However, there is no way to empirically prove this assumption.

[Insert Table 5 about here]

Panel B of Table 5 presents additional robustness checks. The first three columns prove that treatment selection or panel attrition pose no threat to the results. In the first column, I balance the sample. In column (2), I weight the standard regression with the inverse probability that a respondent did not drop out of my sample in the post-reform period. In the third column, I exclude the only population group from my sample that could have selected themselves out of the treatment. Only respondents who were optionally insured under the MHI system had the possibility by opting out of the MHI. However, as noted before, opting out is essentially a lifetime decision and therefore very rare. The DiD estimates from all three robustness checks are close in size to the standard estimate in column (1) of Panel A and confidence intervals largely overlap. Each estimate is significantly different from zero.

I exclude health variables in column (4) since the health status might be endogenous if measured after a convalescent care therapy.[14] The resulting estimate is very robust.

The last column in Panel B clusters standard errors on a higher aggregated level to test whether the common group error structure might be a serious issue in this setting (Angrist and Pischke, 2009). As can be seen, there is no evidence of this.

[Insert Table 6 about here]

In Table 6, I display placebo regressions for *Model 1, 2,* and *3* and Probit as well as OLS specifications. Placebo regressions are a common means to test the common time trend assumption. Finding significant reform effects for years without a reform would cast serious doubts on the plausibility of the common time trend assumption. I use 1994 and 1995 as pseudo-reform years

[14] Keep in mind that health status refers to the time of the interview, whereas the information about convalescent care programs is sampled retrospectively for the previous calendar year. As explained at the beginning of Section 3, if a respondent was interviewed in two subsequent years, I match the current health status information in year t_0 with the convalescent care information from year t_1 which refers to year t_0. Since two-thirds of all interviews were carried out between January and March, the health status is likely to have been measured before the medical treatment.

and, apart from that, the same setup as above. All twelve placebo regression estimates are close to, and not significantly different from, zero.

Reduction in health expenditures

Since reducing health expenditures was the main intention behind the policy reforms, I perform a rough calculation of the decrease in public health expenditures using official data. Official data is available on the total sum that was spent on convalescent care by the public social insurance. Taking the simple difference in expenditures in 1997/1998 vs. 1994/1995 yields a total savings estimate of €835 million per year. This represents a decrease in spending of 12.5 percent (German Federal Statistical Office, 2010). It should be kept in mind, however, that time trends are included in this savings estimate.

As copayments were doubled, this reform raised additional revenues. However, official data show that the total number of convalescent care days consumed decreased by 22 percent from 57 million in 1994/1995 to 44.5 million in 1997/1998 (German Federal Statistical Office, 2010). Multiplying each sum by the pre- and post-reform copayments and taking the difference suggests that increasing copayments not only effectively dampened the demand for convalescent care but it also raised additional revenues of about €435 million per year.[15]

6 Discussion and Conclusion

In this article, I empirically assess the effectiveness of different cost containment measures within a unifying framework. In Germany, from 1997 on, four different health reforms were implemented in order to dampen the demand for convalescent care therapies, to fight moral hazard, and to decrease public health expenditures. At that time, experts claimed that around a quarter of all convalescent care therapies were unnecessary (Schmitz, 1996; Sauga, 1996). In 1995, the German public social insurance system spent €7.6 billion for 1.9 million convalescent care treatments. Given the price elasticity of demand, convalescent care treatments can be considered good proxies for health care demand in general (Wedig, 1988; Ziebarth, 2010).

Two of the health care cost containment measures evaluated applied solely, but universally, to those insured with public health insurance. In Germany, public health insurance coexists with private health insurance providers and strict legal regulations prevent switching between

[15] Under the assumption that 18.8 percent of all therapies were undertaken by East Germans (German Federal Statistical Office, 2010) who were charged lower copayments (see Section 2.2 for details.)

the two independent systems. Privately insured people were not affected by the two reforms and concerns about treatment selection are addressed. Moreover, since the other two of the four cost containment measures only applied to employees in the private sector, I am able to define various subsamples that were affected differently by the reforms. Hence, my empirical findings are based on various difference-in-differences models that compare different mutually exclusive subsamples.

The consistency of the findings across the models, combined with the robustness checks, allows me to conclude the following: first, all reforms, combined, decreased the demand for convalescent care therapies by about 20 percent. Second, doubling the daily copayments for convalescent care treatments was by far the most effective cost containment measure. This measure was responsible for the major part of the total decline in demand.

Third, descriptive evidence from official data suggests that a legally codified reduction in the standard length of the therapies was effective in reducing the true length of the therapies. Fourth, I find no evidence that increasing the waiting times between two treatments had any significant effect on the *decision* to go for convalescent care.

Fifth, while all these policy measures applied universally to every publicly insured person, two other measures evaluated here applied in a rather indirect way. They reduced statutory minimum standards and increased the employers' options to set firm-specific employment conditions. The first of these indirect measures allowed employers to deduct two days of paid vacation for every five days that an employee was unable to work due to a convalescent care therapy. The second cut statutory sick pay for which employees are eligible during convalescent care treatments. I do not find any evidence that these soft cost containment measures were effective in reducing the demand for convalescent care programs. These findings let me conclude that, sixth, indirect measures that reduce statutory minimum conditions in the labor market are far less effective in achieving a specific predetermined policy goal; direct measures that apply universally are much more effective.

As a last exercise, using administrative data, I roughly calculate the reduction in public health expenditures that was induced by all reforms. My back-of-the-envelope calculations suggest that public health expenditures decreased by about €800 million (-12.5 percent) per year due to the decline in the demand for convalescent care. Moreover, doubling the daily copayments raised additional revenues of about €400 million per year.

The question to what degree such policy reforms succeed in reducing moral hazard or whether they actually lead to adverse health outcomes is difficult to quantify and is beyond the scope of this paper. The overall decrease in demand fits well with the *a priori* claims by health experts that a quarter of all pre-reform therapies were unnecessary. Although it is unlikely that moral hazard was completely eliminated by the reforms, it is probable that the majority of the decrease is due to a reduction in moral hazard and led to greater efficiency in the convalescent care market. On the other hand, if medically necessary therapies were not provided, this may have led to adverse health outcomes and, in the long run, to even higher overall health expenditures.

Especially in the case of convalescent care, it is difficult to balance the prevailing degree of moral hazard against potential long-term health improvements that may reduce health expenditures and exert positive external effects. Some studies find positive health effects of health spa stays: patients with chronic diseases experienced reductions in pain and blood pressure, and for a sample of employees, beneficial effects on physical and particularly mental health, such as improved sleep quality, were found (Sekine et al., 2006; Cimbiz et al., 2005; Constant et al., 1998). While two of these studies are purely correlation-based, Constant et al. (1998) estimate the short-term effects of a randomized trial on 224 patients with chronic lower back pain. However, I am unaware of studies evaluating the long-term health effects of convalescent care therapies. Assessing the long-term effects of health care on health outcomes as well as on health expenditures is a promising field for future research.

References

Ai, C. and E. C. Norton (2004). Interaction terms in logit and probit models. *Economics Letters 80*, 123–129.

Angrist, J. D. and J.-S. Pischke (2009). *Mostly Harmless Econometrics: An Empiricist's Companion* (1 ed.). Princeton University Press.

Bazzoli, G. J., H.-F. Chen, M. Zhao, and R. C. Lindrooth (2008). Hospital financial condition and the quality of patient care. *Health Economics 17*(8), 977–995.

Bertrand, M., E. Duflo, and M. Sendhil (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Chiappori, P.-A., F. Durand, and P.-Y. Geoffard (1998). Moral hazard and the demand for physician services: first lessons from a natural experiment. *European Economic Review 42*, 499–511.

Cimbiz, A., V. Bayazit, H. Hallaceli, and U. Cavlak (2005). The effect of combined therapy (spa and physical therapy) on pain in various chronic diseases. *Complementary Therapies in Medicine 13*, 244–250.

Civan, A. and B. Koksal (2010). The effect of newer drugs on health spending: do they really increase the costs? *Health Economics 19*(5), 581–595.

Cockx, B. and C. Brasseur (2003). The demand for physician services: evidence from a natural experiment. *Journal of Health Economics 22*, 881–913.

Constant, F., F. Guillemin, J. F. Collin, and M. Boulangé (1998). Use of spa therapy to improve the quality of life of chronic low back pain patients. *Medical Care 36*(9), 1309–1314.

Deb, P. and P. K. Trivedi (2009). Provider networks and primary-care signups: do they restrict the use of medical services? *Health Economics 18*(12), 1361–1380.

Di Matteo, L. (2005). The macro determinants of health expenditure in the United States and Canada: assessing the impact of income, age distribution and time. *Health Policy 71*(1), 23–42.

Donald, S. G. and K. Lang (2007). Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics 82*(2), 221–233.

Dusheiko, M., H. Gravelle, R. Jacobs, and P. Smith (2006). The effect of financial incentives on gatekeeping doctors: evidence from a natural experiment. *Journal of Health Economics 25*(3), 449–478.

Ellis, R. P. and T. G. McGuire (1996). Hospital response to prospective payment: moral hazard, selection, and practice-style effects. *Journal of Health Economics 15*(3), 257–277.

Felder, S. (2008). To wait or to pay for medical treatment? Restraining ex-post moral hazard in health insurance. *Journal of Health Economics 27*(6), 1418–1422.

Gerfin, M. and M. Schellhorn (2006). Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits—Swiss evidence. *Health Economics 15*(9), 1011–1020.

German Federal Statistical Office (2010). *Federal Health Monitoring.* www.gbe-bund.de, last accessed at August 5, 2010.

German Ministry of Health (2010). www.bmg.bund.de, last accessed at August 5, 2010.

Goldman, D. P., S. D. Hosek, L. S. Dixon, and E. M. Sloss (1995). The effects of benefit design and managed care on health care costs. *Journal of Health Economics 14*(4), 401–418.

Handelsblatt (1995). Reformpaket der Koalition. Kur soll künftig Urlaub kosten. *Handelsblatt 19.12.1995*, 3.

Handelsblatt (1998). Gesundheitsreform sorgt für Aderlaß. *Handelsblatt 10.03.1998*, 21.

Hill, S. C. and B. L. Wolfe (1997). Testing the HMO competitive strategy: an analysis of its impact on medical care resources. *Journal of Health Economics 16*(3), 261–286.

Jahn, J. (1998). Lohnfortzahlung: Gerichte stehen vor Herkulesaufgabe. *Handelsblatt 124: 02.07.1998*, 4.

Karlsson, M. (2007). Quality incentives for GPs in a regulated market. *Journal of Health Economics 26*(4), 699 – 720.

Keeler, E. B., G. Carter, and J. P. Newhouse (1998). A model of the impact of reimbursement schemes on health plan choice. *Journal of Health Economics 17*(3), 297–320.

Keeler, E. B., W. G. Manning, and K. B. Wells (1988). The demand for episodes of mental health services. *Journal of Health Economics 7*, 369–392.

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

Mariñoso, B. G. and I. Jelovac (2003). GPs' payment contracts and their referral practice. *Journal of Health Economics 22*(4), 617–635.

Newhouse, J. P. (1992). Medical care costs: how much welfare loss? *Journal of Economic Perspectives 6*(3), 3–21.

Okunade, A. A. (2004). Concepts, measures, and models of technology and technical progress in medical care and health economics. *The Quarterly Review of Economics and Finance 44*(3), 363–368.

Pauly, M. V. and F. E. Blavin (2008). Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance. *Journal of Health Economics 27*(6), 1407–1417.

Propper, C., S. Burgess, and K. Green (2004). Does competition between hospitals improve the quality of care? Hospital death rates and the NHS internal market. *Journal of Public Economics 88*(7-8), 1247–1272.

Propper, C., B. Croxson, and A. Shearer (2002). Waiting times for hospital admissions: the impact of GP fundholding. *Journal of Health Economics 21*(2), 227–252.

Puhani, P. A. (2008). The treatment effect, the cross difference, and the interaction term in nonlinear "difference-in-differences" models. IZA Discussion Paper Series 3478, IZA. http://www.iza.org, last accessed at August 14, 2009.

Ridinger, R. (1997). Einfluss arbeitsrechtlicher Regelungen auf die Beschäftigungsentwicklung im Handwerk–Ergebnisse von Befragungen von Handwerksbetrieben im 3. Quartal 1997. Dokumentation, Zentralverband des Deutschen Handwerks. http://www.zdh.de, last accessed at August 14, 2009.

Sauga, M. (1996). Gesundheit im eigenen Bett: Bonn will den Anstieg der Kur-Ausgaben begrenzen. Die Bäderlobby macht mobil. *Handelsblatt 11.01.1996*, 21.

Schmitz, H. (1996). Politiker, Wissenschaftler und Praktiker suchen Wege, die Verschwendung bei Rehabilitation und Heilmaßnahmen zu stoppen. Kuren sollten für Arbeitnehmer kein preiswerter Urlaub sein. *Handelsblatt 01.04.1996*, 2.

Schreyögg, J. and M. M. Grabka (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.

Schut, F. T. and W. P. M. M. V. de Ven (2005). Rationing and competition in the Dutch health-care system. *Health Economics 14*(S1), S59–S74.

Sekine, M., A. Nasermoaddeli, H. Wang, H. Kanayama, and S. Kagamimori (2006). Spa resort use and health-related quality of life, sleep, sickness absence and hospital admission: the Japanese civil servants study. *Complementary Therapies in Medicine 14*, 133–143.

Siciliani, L., A. Stanciole, and R. Jacobs (2009). Do waiting times reduce hospital costs? *Journal of Health Economics 28*(4), 771–780.

Sloan, F. A., G. A. Picone, D. H. TaylorJr., and S.-Y. Chou (2001). Hospital ownership and cost and quality of care: is there a dime's worth of difference? *Journal of Health Economics 20*(1), 1–21.

van Kleef, R., W. van de Ven, and R. van Vliet (2009). Shifted deductibles for high risks: more effective in reducing moral hazard than traditional deductibles. *Journal of Health Economics 28*(1), 198–209.

Voorde, C. V. D., E. V. Doorslaer, and E. Schokkaert (2001). Effects of cost sharing on physician utilization under favourable conditions for supplier-induced demand. *Health Economics 10*(5), 457–471.

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

Wedig, G. J. (1988). Health status and the demand for health: results on price elasticities. *Journal of Health Economics 7*, 151–163.

Weisbrod, B. A. (1991). The health care quadrilemma: An essay on technological change, insurance, quality of care, and cost containment. *Journal of Economic Literature 29*(2), 523–52.

Winkelmann, R. (2004). Health care reform and the number of doctor visits – an econometric analysis. *Journal of Applied Econometrics 19*(4), 455–472.

Ziebarth, N. and M. Karlsson (2010). A natural experiment on sick pay cuts, sickness absence, and labor costs. *Journal of Public Economics 94*(11-12), 1108–1122.

Ziebarth, N. R. (2010). Estimating price elasticities of convalescent care programmes. *The Economic Journal 120*(545), 816–844.

Zweifel, P. and W. G. Manning (2000). Moral hazard and consumer incentives in health care. In A. J. Culyer and J. P. Newhouse (Eds.), *Handbook of Health Economics* (1 ed.), Volume 1, Chapter 8, pp. 409–459. Elsevier.

Figure 1: Incidence of Convalescent Care Programs

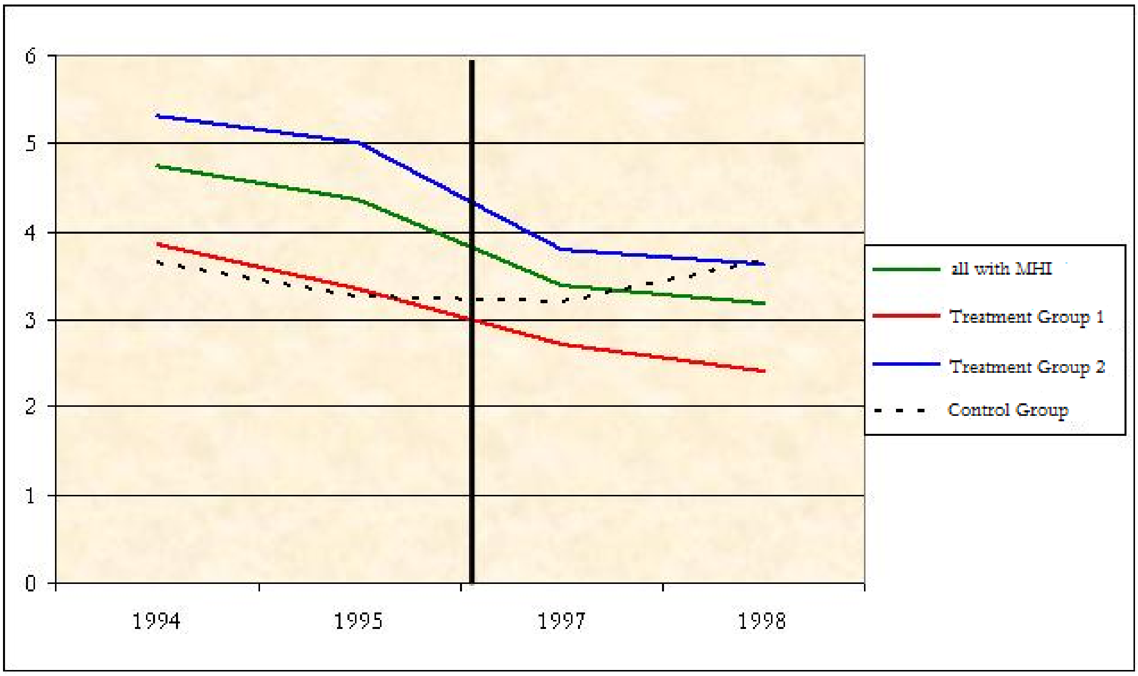

Table 1: Identification and Definition of Subgroups and Subsamples

	Reform 1: Copayment doubling	**Reform 2: Waiting time increase**	**Reform 3: Paid vacation reduction**	**Reform 4: Sick pay decrease**
Private sector with MHI (1) (*Treatment Group 1*)	yes	yes	yes	yes
Self-employed with MHI (2)	yes	yes	no	no
Non-working with MHI (3)	yes	yes	no	no
Public sector with MHI (4)	yes	yes	no	no
Apprentices with MHI (5) (*Treatment Group 2*)	yes	yes	no	no
Self-employed with PHI (6)	no	no	no	no
Non-working with PHI (7)	no	no	no	no
Public sector with PHI (8)	no	no	no	no
Apprentices with PHI (9) (*Control Group*)	no	no	no	no

Table 2: Variable Means by Treatment and Control Group

Variable	**Treatment Group 1**	**Treatment Group 2**	**Control Group**
Convalescent Care	0.032	0.045	0.032
Personal characteristics			
Female	0.397	0.614	0.364
Age	37	47	45
Age squared	1,693	2,576	2,231
Immigrant	0.220	0.170	0.064
East Germany	0.246	0.309	0.134
Partner	0.798	0.671	0.745
Children	0.470	0.364	0.404
Good health	0.598	0.463	0.609
Bad health	0.106	0.201	0.110
Educational characteristics			
Dropout	0.051	0.073	0.027
8 years of completed schooling	0.368	0.425	0.206
10 years of completed schooling	0.319	0.267	0.299
12 years of completed schooling	0.034	0.025	0.049
13 years of completed schooling	0.112	0.113	0.387
Certificate degree	0.116	0.077	0.027
Job characteristics			
Full-time employed	0.831	0.197	0.671
Part-time employed	0.130	0.046	0.053
Marginally employed	0.040	0.010	0.008
Civil servant	0.000	0.010	0.427
Public servant	0.000	0.679	0.645
Self employed	0.000	0.056	0.258
Apprentice	0.000	0.057	0.012
Gross wage per month	1,860	618	2,126
Regional unemployment rate	11.706	12.317	11.031
N	23,530	4,261	37,758

In contrast to Appendix A, this table gives mean values separately for the treatment and control groups. As detailed in Section 3, *convalescent care* is the overall incidence of convalescent care programs.

Table 3: Determinants of Convalescent Care

Variable	Coefficient	Standard Error
Personal characteristics		
Female (d)	-0.0006	0.002
Age	0.0023***	0.000
Age squared /1,000	-0.0169***	0.003
Immigrant	-0.0056**	0.002
East Germany	0.0017	0.007
Partner	-0.0025	0.002
Children	-0.0014	0.002
Good health	-0.0230***	0.002
Bad health	0.0398***	0.003
Educational characteristics		
8 years of completed schooling	0.0044	0.004
10 years of completed schooling	0.0100**	0.004
12 years of completed schooling	0.0106	0.007
13 years of completed schooling	0.0046	0.004
Other certificate	0.0045	0.004
Job characteristics		
Full-time employed	0.0017	0.002
Part-time employed	-0.0035	0.003
Marginally employed	-0.0033	0.005
Gross wage per month/1,000	-0.0019**	0.001
Regional unemployment rate	-0.0028***	0.001
R-squared	0.0947	
χ^2	1,542	
N	65,549	

* p<0.10, ** p<0.05, *** p<0.01; marginal effects, which are calculated at the means of the covariates, are displayed. Dependent variable is *convalescent care* and measures the incidence of all convalescent care programs. Standard errors in parentheses are adjusted for clustering on person identifiers. Regression includes state dummies. Left out reference categories are dropout and non-employed.

Table 4: Assessing the Reforms' Effectiveness: Net Effect, Copayment Effect, and Effect of Cut in Paid Leave

	Model 1: Net effect			**Model 2: Copayment effect**			**Model 3: Cut in paid leave**		
Variable	**Raw**	**Probit**	**OLS**	**Raw**	**Probit**	**OLS**	**Raw**	**Probit**	**OLS**
DiD	**-0.0129****	**-0.0081****	**-0.0096***	**-0.0165*****	**-0.0136****	**-0.0163*****	**0.0035**	**0.0016**	**0.0045**
	(0.0057)	**(0.0041)**	**(0.0057)**	**(0.0057)**	**(0.0056)**	**(0.0057)**	**(0.0031)**	**(0.0024)**	**(0.0031)**
Treatment indicator	0.0046	0.0055***	0.0106**	0.0193***	0.0051*	0.0077*	-0.0147***	0.0010	-0.0058**
(*T1, T2, or T3*)	(0.0042)	(0.0020)	(0.0044)	(0.0042)	(0.0028)	(0.0044)	(0.0023)	(0.0020)	(0.0026)
Post-reform dummy	0.0035	-0.0045	-0.0089	0.0035	-0.0007	-0.0018	-0.0130***	-0.0149***	-0.0214***
(*post97*)	(0.0053)	(0.0041)	(0.0066)	(0.0053)	(0.0050)	(0.0063)	(0.0021)	(0.0029)	(0.0037)
Year 1997 (d)		0.0002	0.0000		0.0012	0.0018		0.0015	0.0015
		(0.0020)	(0.0030)		(0.0021)	(0.0027)		(0.0018)	(0.0021)
Year 1996 (d)		-0.0047**	-0.0083**		-0.0049**	-0.0078**		-0.0051***	-0.0069
		(0.0019)	(0.0038)		(0.0020)	(0.0034)		(0.0017)	(0.0028)
Year 1995 (d)		-0.0021	-0.0036		-0.0022	-0.0035		-0.0020	-0.0028
		(0.0018)	(0.0033)		(0.0019)	(0.0031)		(0.0016)	(0.0025)
Educational covariates	no	yes	yes	no	yes	yes	no	yes	yes
Job covariates	no	yes	yes	no	yes	yes	no	yes	yes
Personal covariates	no	yes	yes	no	yes	yes	no	yes	yes
Regional unempl. rate	no	yes	yes	no	yes	yes	no	yes	yes
State dummies	no	yes	yes	no	yes	yes	no	yes	yes
Year dummies	no	yes	yes	no	yes	yes	no	yes	yes
R-squared	0.0006	0.1217	0.0428	0.0012	0.0901	0.0339	0.0019	0.0956	0.0054
χ^2/F-stat	6	793	12	16	992	19	36	1454.9678	8
N	27,791	27,791	27,791	42,019	42,019	42,019	61,288	61,288	61,288

* p<0.1, ** p<0.05, *** p<0.01; in columns (2), (4), and (6), marginal effects are displayed; they are calculated at the means of the covariates except for *T1 (T2, T3)*(=1) and *DiD*(=1). Dependent variable is *convalescent care* and measures the incidence of convalescent care programs (see Section 3). Every column represents one regression model; all columns except for (2), (4), and (6) estimate OLS models. Columns (1) to (3) use *T1*, columns (4) to (6) use *T2*, and columns (7) to (9) use *T3* (see Section 3 for further details). Standard errors in parentheses are adjusted for clustering on person identifiers.

Table 5: Robustness Checks

Panel A					
	Standard	**w/o 1996**	**'96 vs. '97**	**flexible**	
DiD	-0.0136**	-0.0109*	-0.0180*	DiD98	-0.0115***
	(0.0056)	(0.0062)	(0.0098)		(0.0042)
				DiD97	-0.0089**
					(0.0043)
				DiD96	0.0084
					(0.0080)
Panel B					
	balanced sample	**weighted**	**w/o optionally insured**	**no health covariates**	**cluster state×year**
DiD	-0.0153**	-0.0155**	-0.0125**	-0.0144**	-0.0144**
	(0.0067)	(0.0061)	(0.0057)	(0.0061)	(0.0057)

* p<0.1, ** p<0.05, *** p<0.01; marginal effects are displayed; they are calculated at the means of the covariates except for *T2*(=1) and *DiD*(=1). Dependent variable is *convalescent care* and measures the incidence of convalescent care programs (see Section 3). Every cell represents one probit DiD model. All models are similar to the one in column (5) of Table 4, i.e., they estimate the copayment effect using Model 2 and comparing *Treatment Group 2* to the *Control Group* (see Section 3). Column (1) in Panel A is the standard DiD estimate, i.e., the estimate in column (5) of Table 4. Column (2) in Panel A excludes the year 1996 from the regression and is the estimate excluding anticipation effects. Column (3) in Panel A contrasts the year 1996 to the year 1997 and thus estimates the reforms' short-run effect. Column (5) in Panel A shows the most flexibel of all specifications. Instead of interacting the post-reform dummy *post97* with the treatment indicator *T2*, it includes three alternative interaction terms: *Year1996*×*T2* (*DiD96*), *Year1997*×*T2* (*DiD97*), and *Year1998*×*T2* (*DiD98*). Column (1) in Panel B uses a balanced sample and thus excludes panel attrition effects. Column (2) in Panel B weights the regression with the inverse probability that a person does not drop out of the sample in post-reform years. Column (3) in Panel B excludes the only respondents that could have selected themselves out of the treatment, i.e., optionally MHI insured people. Column (4) in Panel B excludes all health measures from the list of covariates. Column (5) in Panel B clusters the standard errors at a higher aggregated level, i.e., the state×year level (80 cluster). Standard errors in all other models are adjusted for clustering on person identifiers and are always in parentheses. All models have 42,019 observations expect for Panel A column (2) (33,975 obs.) and column (3) (16,935) as well as Panel B column (1) (30,625) and column (3) (38,962). For more details about the different model specifications and the interpretation of the results, please see main text.

Table 6: Placebo Reform Estimates

	Model 1: Net effect		Model 2: Copayment effect		Model 3: Cut in paid leave	
Variable	**Probit**	**OLS**	**Probit**	**OLS**	**Probit**	**OLS**
DiD95	0.0015	0.0004	0.0052	0.0058	-0.0029	-0.0058
	(0.0055)	(0.0068)	(0.0076)	(0.0068)	(0.0026)	(0.0039)
DiD94	-0.0004	-0.0008	0.0013	0.0045	-0.0004	-0.0037
	(0.0050)	(0.0073)	(0.0069)	(0.0071)	(0.0026)	(0.0040)
Educational covariates	yes	yes	yes	yes	yes	yes
Job covariates	yes	yes	yes	yes	yes	yes
Personal covariates	yes	yes	yes	yes	yes	yes
Regional unempl. rate	yes	yes	yes	yes	yes	yes
State dummies	yes	yes	yes	yes	yes	yes
Year dummies	yes	yes	yes	yes	yes	yes

* p<0.1, ** p<0.05, *** p<0.01; in columns (1), (3), and (5), marginal effects are displayed; they are calculated at the means of the covariates except for *T1 (T2, T3)*(=1) and *DiD94* (*DiD95*)(=1). All columns but (2), (4), and (6) estimate OLS models. The dependent variable is *convalescent care* and measures the incidence of convalescent care programs (see Section 3). Every cell represents one regression model. Columns (1) and (2) use *T1*, columns (3) and (4) use *T2*, and columns (4) and (5) use *T3* (see Section 3 for further details). Each model in columns (1) and (2) has 27,791 observations; each model in columns (3) and (4) has 42,019 observations and columns (5) and (6) are based upon 61,288 observations. All models compare the same groups of (pseudo) treated and (pseudo) non-treated respondents than the non-placebo models. *DiD94* (*DiD95*) is an interaction term between the treatment indicator (*T1*, *T2*, or *T3*) and the year 1994 (1995). Standard errors in parentheses are adjusted for clustering on person identifiers.

Appendix A

Table 7: Descriptive Statistics for the Working Sample

Variable	Mean	Std. Dev.	Min.	Max.	N
Dependent variable					
Convalescent care	0.0393	0.1943	0	1	65,549
Covariates					
Treatment Indicators					
T1	0.8467	0.3603	0	1	27,791
T2	0.8986	0.3019	0	1	42,019
T3	0.3839	0.4863	0	1	61,288
Personal characteristics					
Female	0.5195	0.4996	0	1	65,549
Age	44	17	18	99	65,549
Age squared	2,236	1,625	324	9,801	65,549
Immigrant	0.1811	0.3851	0	1	65,549
East Germany	0.2751	0.4465	0	1	65,549
Partner	0.7214	0.4483	0	1	65,549
Children	0.4045	0.4908	0	1	65,549
Good health (best 2 of 5 categories)	0.521	0.4996	0	1	65,549
Bad health (worst 2 of 5 categories)	0.1607	0.3672	0	1	65,549
Educational characteristics					
Drop out	0.062	0.2412	0	1	65,549
8 years of completed schooling	0.3905	0.4879	0	1	65,549
10 years of completed schooling	0.2878	0.4528	0	1	65,549
12 years of completed schooling	0.0298	0.1701	0	1	65,549
13 years of completed schooling	0.1305	0.3369	0	1	65,549
Other certificate	0.0878	0.2829	0	1	65,549
Job characteristics					
Full-time employed	0.455	0.498	0	1	65,549
Part-time employed	0.0767	0.2662	0	1	65,549
Marginally employed	0.0204	0.1414	0	1	65,549
Gross wage per month	1,162	1,301	0	51,129	65,549
Regional unemployment rate	12.0	3.9	7	21.7	65,549