u^b

b
**UNIVERSITÄT
BERN**

Faculty of Business, Economics
and Social Sciences

Department of Economics

**Does full insurance increase
the demand for health care?**

Stefan Boes
Michael Gerfin

13-05

October 2013

DISCUSSION PAPERS

**Schanzeneckstrasse 1
Postfach 8573
CH-3001 Bern, Switzerland**
http://www.vwi.unibe.ch

Does full insurance increase the demand for health care?

Stefan Boes[*]
University of Lucerne

Michael Gerfin[†]
University of Bern and IZA

October 3, 2013

Abstract

We estimate the causal effect of having full health insurance on health care expenditures. We take advantage of a unique quasi-experimental setup in which deductibles and co-payments were zero in a managed care plan, and non-zero in regular insurance, until a policy change forced all individuals with an active plan to cover a minimum amount of their expenses. Using panel data and a non-linear difference-in-differences strategy, we find a demand elasticity of about -0.14 comparing full insurance with the cost-sharing model, and a significant upward shift in the likelihood to generate costs.

JEL Classification: I11, C14

Keywords: Deductibles, health cost, quasi experiment, changes-in-changes.

---

[*]University of Lucerne, Department of Health Sciences and Health Policy, Frohburgstrasse 3, CH-6002 Lucerne, Switzerland, phone: +41 41 229 4959, fax: +41 41 229 5635, email: stefan.boes@unilu.ch.

[†]University of Bern, Department of Economics, Schanzeneckstrasse 1, CH-3001 Bern, Switzerland, phone: +41 31 631 4092, fax: +41 31 631 3992, email: michael.gerfin@vwi.unibe.ch.

1 Introduction

Health insurance plans usually contain some form of patient cost-sharing such as deductibles or co-payments in order to mitigate moral hazard effects. The common measure of moral hazard effects in health insurance is the price sensitivity of health care demand, conditional on health status (Pauly, 1968; Cutler and Zeckhauser, 2000). The first estimates of this price sensitivity were obtained in the RAND Health Insurance Experiment (HIE), which was run during the 1970s. Aron-Dine et al. (2013) provide an account of the HIE and a re-analysis of the experimental data within the modern causal analysis framework. Their estimates of the demand elasticity comparing full insurance with several plans containing different degrees of cost-sharing are in the range of -0.1 to -0.2. This of course corresponds to the well-known benchmark estimate of roughly -0.2 reported in Keeler and Rolph (1988).

A major challenge in identifying the price elasticity is the definition of the price. Demand is usually measured by expenditures for health care, not units of health care services. The price of an additional dollar spent on health care is then the share of that dollar which the patient has to pay out-of-pocket. Given the non-linear cost-sharing schedules in most health insurance plans caused by deductibles, co-payments and stop-losses, this marginal price changes as health care expenditures accumulate during the year. Keeler et al. (1977) derived a model of medical consumption decisions in the presence of deductibles using a dynamic programming approach. They demonstrated that the correct price for consumers to use when making health consumption decisions is the shadow (or effective) price. They also showed that using the "wrong" price (e.g., the marginal instead of the effective price) leads to biased estimates of price responsiveness. Ellis (1986) provides evidence that the expected end-of-year price is a good proxy for the shadow price. However, due to the complexity of the model, the effective price has not been used in applied work until recently.[1] For example, the RAND elasticity of -0.2 is calculated assuming individuals respond only to the spot price. More recent stud-

---

[1]Ellis (1986) is a notable exception.

ies have assumed that individuals respond to the actual (realized) end-of-year price (Eichner, 1998; Kowalski, 2009; Marsh, 2012). Aron-Dine et al. (2012) provide evidence that the insured are not myopic and indeed respond to the expected end-of-year price.

In this paper, we sidestep the problem of defining the correct price taking advantage of panel data, provided by a large health insurance company in Switzerland, combined with a unique quasi-experimental setup. For a group of individuals (those in the HMO plan), there was a limited period of time without any cost-sharing. We estimate the behavioral changes induced by going from some (unknown) positive effective price to price zero, implying a price reduction of 100%. As a control group, we use individuals within the regular health care plan of the same insurance company living in the same cities.

The analysis is carried out within the potential outcomes framework. We look at the average causal effect of having full health insurance, as opposed to cost-sharing. This effect is defined as the difference between the actual average outcomes (health care expenditures or incidence of zero expenditures) with full health insurance and the counterfactual expected outcomes in absence of full insurance. In addition, we estimate the causal effects at different quantiles of the health care expenditures distribution. The econometric analysis uses the standard difference-in-differences (DID) estimator and the more recent changes-in-changes (CIC) estimator proposed by Athey and Imbens (2006). Note that in our case, treatment (full insurance) occurred in the first observation period (2002), and the no-treatment periods took place in the following years 2003 and 2004. We interpret this as a "reverse" DID model. Having two no-treatment periods allows us to examine our identification strategy because we should not find any effect in the comparison of the years 2003 and 2004.

Our results suggest a small but significant positive effect of full insurance on expected expenditures within the HMO group. The estimated effect translates into an elasticity of -0.14. The probability of having positive health care expenditures is increased by about 7%-points (from 0.76 to 0.83). This corresponds to an increase by roughly 10%.

Methodologically, the paper most related to ours is Borah et al. (2011). They analyze a

setup in which one employer in the US changed its policy such that only a high deductible plan with markedly larger deductibles than before was available for the employees. There was no such change in the plan offered by the control employer. Both employers operate in similar communities in close geographic proximity. They estimate both average and quantile treatment effects of having the high deductible plan on total health care expenditures using the DID and the CIC estimators. Their estimated causal effect is roughly -10%, but the estimate is only significant when they exclude the top percentile. From the information given in the paper it is not possible to compute a price elasticity.

The remainder of the paper is organized as follows. Section 2 briefly gives some institutional background, Section 3 describes the data, Section 4 provides the econometric framework, and Section 5 discusses the empirical results. Section 6 concludes.

2 Institutional context

Since the reform of the Swiss health insurance law in 1996, a basic health insurance is mandatory for all people living or working in Switzerland (with few exceptions, e.g., staff of international organizations or diplomats). Coverage is for a rather comprehensive set of medical services and pharmaceuticals, offered by about 80 private, not-for-profit insurers competing in a regulated market. Free consumer choice of plan is a distinctive feature of the system. There is no pre-selection of plans by employers or government agencies. Insurers are obliged to accept all applicants during annual open enrollment periods. The contracts have a duration of one year. Premiums are community-rated, not risk-rated. In the baseline contract (referred to as the regular health care plan), insured individuals enjoy unlimited access to all licensed physicians and most hospitals in their region of residence. In managed care type plans, access to health care is through a gatekeeper.

In the regular plan, there was a minimum annual deductible of CHF 230 in the years 2002 and 2003, which was increased to CHF 300 since 2004, and a co-payment rate of 10 percent

4

up to a stop-loss at CHF 700 (plus the deductible) per year. The insured can opt for higher deductibles, which were regulated at 400, 600, 1200 or 1500 CHF during the analysis period. These are offered with premium reductions, which are fixed percentages of the base premium or 80 percent of the additional financial risk taken by the consumer (deductible minus 230), whichever is less. The same deductible levels are typically applied to managed-care plans, although the health insurance law allows insurers to set zero cost-sharing in those plans. We will further explore this exemption below. Since all contracts are on an individual basis, there are no family-related shared deductibles.

Physicians in independent practice are reimbursed fee-for-service according to an administered fee schedule that is collectively bargained between the providers' and the insurers' associations. Hospitals receive per diems for patients treated (the nation-wide introduction of a DRG system was introduced in 2012). In addition to the mandatory basic insurance, there are several voluntary supplementary insurance types available. By contrast, these are risk-rated, and we do not consider them in our analysis.

3 Health insurance data

To make the HMO plan more attractive, the health insurance company CSS offered it without deductibles and co-payments from 1996 until 2002. In other words, the HMO plan provided full health insurance. In 2003, after complaints by the HMO physicians regarding above average health costs generated by CSS-insured patients (HMO practices also treated patients from other insurance companies), CSS introduced the same cost-sharing instruments as in their other plans. This introduction of cost-sharing provides a possibility to estimate the *causal* effect of full health insurance on health care demand.

CSS provided data for four large Swiss cities (Zurich, Basel, Berne, and Lucerne) where the HMO plan was on offer. The data cover the years 2002 to 2004. Prior to 2002 the data were not available in comparable quality due to changes in the electronic billing system. Hence, we

confine ourselves to the last year in which the HMO plan provided full insurance and the two following years in which all insured faced the same cost-sharing options. In addition, we have data for all insured in regular contracts living in the same cities during the same time period. These insured will serve as our control group.

The data are very reliable in the outcome variable *effective health cost* (which for simplicity we treat synonymously to health care expenditures) per year and the chosen insurance contract. In addition to the total cost, we also use a discretized version indicating the extensive margin, i.e., the incidence of *zero cost*. As is often the case with administrative data, there are only a few control variables available. In the present case, we have age, gender and city of residence. Table 1 shows descriptive statistics of the variables. The sample consists of all insured in the regular or the HMO plan observed over the full three years. This yields a total of 85,626 observations (28,542 insured times 3 years). The fraction of HMO is about 12.5 percent. As one would expect, the HMO insured generate significantly smaller health costs with about CHF 1700 per year (fraction of 22% with no cost) as opposed to about CHF 4300 on average (16% with no cost) for the regularly insured.

— Insert Table 1 about here —

The background characteristics also indicate that the HMO group differs from the regularly insured, in terms of age, gender, and regional composition. Nevertheless, we will argue below that the latter group can serve as a suitable control when looking at trends, not levels. Table 2 gives an overview of these trends in mean outcomes by group (HMO vs. regular) and by year for observation period. Between 2002 and 2003, average health care expenditures for the HMO group slightly decreased (from CHF 1678 to CHF 1563) whereas that for the regularly insured increased by almost CHF 400. The fraction of zero cost increased by about 9%-points for HMO, and it stayed about constant for the regularly insured.

— Insert Table 2 about here —

The 2002/03 changes for HMO possibly reflect the switch from full insurance to cost-sharing. On the other hand, the 2003/04 comparison can help us to evaluate our identification strategy, because all insured faced the same cost-sharing instruments. We observe that the fraction of zero cost is higher for HMO as opposed to regular in these two years, as one would expect, but the fraction stays about constant for both groups from 2003 to 2004. And even though the cost increase for the regularly insured is almost twice as large as for the HMO insured, in terms of growth rates we observe similar changes (+23.8% for HMO as opposed to +22.5% for regular).[2] Figure 1 illustrates these developments in terms of the entire expenditure distribution. For better readability, we show the distribution of the log of expenditures plus one.[3] For the regularly insured, we observe small movements in the upper half of the distribution (which explain the positive shift in means). For the HMO insured, the graphs look very similar comparing 2003 and 2004. The striking difference for HMO arises in the 2002 to 2003 comparison with the shift at the extensive margin.

— Insert Figure 1 about here —

4   Econometrics

To formalize the comparisons (and underlying assumptions) of the previous section in a potential outcomes framework, we now introduce our econometric model. It is based on variants of the difference-in-differences (DID) estimator. The basic idea of DID is that the effect of full health insurance for HMO can be estimated by a comparison of the 2002 and 2003 outcomes for that group, net of the time trend in a suitable control group (see Imbens and Wooldridge (2009) for an overview of DID methods). We have to deal with two complications, however. First, in

---

[2]Re-analyzing the CSS data for the time period 2003-2006 as used in Trottmann et al. (2012) we find strong evidence for the common growth assumption. The difference between log mean expenditures of regular insurance and HMO insurance is roughly constant over the years, while the difference between mean expenditures is increasing (conditional on staying within the initially chosen health plan).

[3]Using the log transform here is without consequence for our later analysis because Figure 1 is purely descriptive and for the purpose of illustration of the changes in the outcome distribution.

reaction to the changes in cost-sharing for HMO, some insured may either have changed their plan, or left CSS for another insurance company. This behavior may affect the composition of the treatment and control groups, which would invalidate the DID identification assumptions. We address this potential selection problem in Section 5.1.

Second, there is the issue of the distribution of health cost, which has a large probability mass at zero. The traditional approach in health economics is to use a two-part model in which the discrete and the continuous part of the distribution are analyzed separately (Mullahy, 1998; Buntin and Zaslavsky, 2004). However, Angrist (2001) argues that in the second part (conditional on positive cost) the causal effect is not well-defined, essentially because conditioning on an outcome is not valid in the potential outcomes framework. Furthermore, in a difference-in-differences context, it is unclear how to apply the two-part model because both treated and control observations may switch between the two parts across time periods. For this reason we model the extensive margin (incidence of zero expenditures) as well as the mean and quantiles of the entire outcome distribution.

4.1 The standard DID model

We consider the standard model in which there are two time periods $T \in \{0, 1\}$ and two groups $G \in \{0, 1\}$. The potential outcome without treatment is denoted as Y^0, the potential outcome with treatment as Y^1. The observed outcome for individual i is given by

$$Y_i = Y_i^0 \cdot (1 - D_i) + Y_i^1 \cdot D_i, \tag{1}$$

where $D_i = G_i \cdot T_i$ is the treatment indicator. The treatment effect is defined as $\tau_i = Y_i^1 - Y_i^0$. In the standard DID model, the potential outcome without treatment is given by

$$Y_i^0 = \alpha + \beta \cdot T_i + \gamma \cdot G_i + \varepsilon_i, \tag{2}$$

where β measures the time effect, γ the group (or selection) effect, and ε_i represents unobserved characteristics independent of G_i and T_i. The standard DID estimand, τ^{DID}, is

$$
\begin{aligned}
\tau^{DID} = & \left\{ E[Y_i|G_i = 1, T_i = 1] - [E[Y_i|G_i = 1, T_i = 0]] \right\} \\
& - \left\{ E[Y_i|G_i = 0, T_i = 1] - [E[Y_i|G_i = 0, T_i = 0]] \right\}
\end{aligned}
\tag{3}
$$

which identifies the average effect of the treatment on the treated (ATT), formally $E[\tau_i|D_i = 1]$. The four expectations are easily estimated by their sample analogues or by regressing the observed outcome Y_i on T_i, G_i and D_i. The main identifying assumption is that in absence of treatment the time trend of Y is the same for the two groups defined by G. It is important to note that this assumption cannot hold for nonlinear transformations of Y. In the present context this means that the common trend assumption may either hold in levels or in logs (growth), but not in both. We come back to this point in Section 4.3.

4.2 The CIC model

Athey and Imbens (2006) propose to generalize the standard DID model. First, they write the outcome in absence of treatment as

$$
Y_i^0 = h(U_i, T_i)
\tag{4}
$$

where $h(u, t)$ is an increasing function in u. The random variable U represents unobservable characteristics (e.g., risk aversion, health status). The distribution of U is allowed to vary across groups, but not over time, i.e., $U_i \perp T_i|G_i$. Hence, in absence of treatment, an individual with $U_i = u$ will have the same outcome in a given time period t irrespective of her group membership. The standard DID model is nested in eq. (4) under linearity assumptions imposed on $U_i = \alpha + \gamma \cdot G_i + \varepsilon_i$ and on $h(u, t) = u + \beta \cdot t$.

To derive the CIC estimator, denote the four observed outcomes of Y as Y_{gt}. For example, Y_{11} is the observed outcome of Y for individuals with $g = 1$ and $t = 1$. Assuming these are the observations receiving the treatment, we have $Y_{11} = Y_{11}^1$, i.e., the observed outcome corresponds to the potential outcome. The other three observed outcomes, Y_{00}, Y_{01} and Y_{10}

are used to estimate the counterfactual outcome Y_{11}^0. The cumulative distribution functions of the observed outcomes Y_{gt} are denoted by F_{gt}.

The assumptions required for the identification of the CIC model for continuous outcomes include eq. (4) combined with the time invariance assumption $U_i \perp T_i | G_i$. To illustrate the idea of the CIC model, consider an individual in the treatment group ($g = 1$) prior to the treatment ($t = 0$), whose observed outcome is y. Let $F_{10}(y) = q$. Given the CIC assumptions, this individual would have the same y if it were in the control group instead. Hence, the rank of this individual in the distribution of Y_{00} is given by $F_{00}(y) = q'$. Because the distribution of U is independent of time, the second period outcome of an individual with $U_i = u$ in absence of treatment is the q'-quantile of F_{01}. Formally,

$$y' = F_{01}^{-1}\left(F_{00}(y)\right) \tag{5}$$

y' is the estimate of the counterfactual outcome without treatment for a treatment group observation with pre-treatment outcome y. The counterfactual distribution of Y_{11}^0, denoted by $F_{11}^0(y)$, is given by

$$F_{11}^0(y) = F_{10}\left(F_{00}^{-1}(F_{01}(y))\right) \tag{6}$$

Now, the CIC estimand for the average treatment effect τ^{CIC} can be written as

$$\tau^{CIC} = E[Y_{11}^1] - E[Y_{11}^0] = E[Y_{11}] - E[(F_{01}^{-1}(F_{00}(Y_{10}))] \tag{7}$$

The corresponding quantile treatment effects at quantile q are given by

$$\tau_q^{CIC} = F_{11}^{1\,-1}(q) - F_{11}^{0\,-1}(q) = F_{11}^{-1}(q) - F_{01}^{-1}(F_{00}(F_{10}^{-1}(q))) \tag{8}$$

For additional details see Athey and Imbens (2006). Note that the described estimator is for continuous outcomes and requires that $h(u, t)$ is strictly increasing in u. Treatment effects for discrete outcomes are are not point identified in the CIC framework. Athey and Imbens (2006) derive bounds for these cases. They show that the estimator described above corresponds to the lower bound of the treatment effect in the case of discrete outcomes.

4.3 Estimation model

In our baseline model, we focus on individuals who remain in their insurance plan throughout the observation period. For this reason, we do not need to condition on individual characteristics like age, gender and region of residence. Their effects are absorbed in the group effect. The DID estimates are obtained from the following general specification

$$E[Y_i|G_i, T_i^{02}, T_i^{04}] = m\left(\alpha + \beta_1 T_i^{02} + \beta_2 T_i^{04} + \gamma G_i + \delta_1 G_i \cdot T_i^{02} + \delta_2 G_i \cdot T_i^{04}\right) \tag{9}$$

where $G_i = 1$ denotes membership in the HMO plan, T^{02} and T^{04} are dummies for 2002 and 2004, and 2003 is used as the base year. Note that this setup forms a "reversed" difference-in-differences design in which the treatment (full insurance) takes place in the first period (2002). In periods 2 and 3, treated and control face identical cost-sharing incentives.

The average effect of the treatment on the treated (ATT) is calculated as

$$ATT^{2002-2003} = m(\alpha + \beta_1 + \gamma + \delta_1) - m(\alpha + \beta_1 + \gamma) \tag{10}$$

This gives the average causal effect of having full health insurance on Y_i (health expenditures and incidence of zero cost) for the HMO group. The placebo effect (PE) is calculated as

$$PE^{2004-2003} = m(\alpha + \beta_2 + \gamma + \delta_2) - m(\alpha + \beta_2 + \gamma) \tag{11}$$

The placebo effect serves as a test of our identification strategy. Since both the treatment and the control group face identical cost-sharing options in 2003 and 2004, the placebo effect should be zero, which can be formally tested.

The specification of $m(\cdot)$ depends on the outcome considered. For the zero cost indicator, the linear index is fully saturated, and hence we estimate the model by OLS with $m(\cdot)$ the identity function. The left-hand side of eq. (9) corresponds to the probability of having zero health costs in this case (denoted by $P(0)$ below), and the coefficients on the interaction terms, δ_1 and δ_2, are the ATT of having full insurance and the placebo effect, respectively.

For the effect of full health insurance on total expenditures (including the zeros) the specification of $m(\cdot)$ is more subtle. Given that the common trend assumption cannot be fulfilled

in say levels and growth rates, the specification of $m(\cdot)$ is critical for the validity of the assumption. In a first step, we use the traditional specification for $m(\cdot)$, the identity function, which yields the common linear DID model. In the second specification, and to avoid the log transformation of the dependent variable (which would require some arbitrary assumptions for the observations with zero expenditures), we choose a generalized linear model (GLM), in which $m(\cdot) = \exp(\cdot)$. In this case, we estimate eq. (9) by quasi maximum likelihood assuming a Poisson distribution. Gourieroux et al. (1984) have shown that this estimator is consistent if the mean function is correctly specified (see also Wooldridge (2010)). We use the placebo effect as a device to decide upon the preferred specification of $m(\cdot)$.

Estimation of the CIC model does not require the specification of $m(\cdot)$, and the quantities in eqs. (7) and (8) can be estimated fully nonparametrically. The CIC model can therefore be used as an additional check for the specification of $m(\cdot)$ in the parametric alternatives.

5   Results

5.1   Selection out of HMO

Restricting the baseline sample to stayers in their plan over the full three years implies that the group effect accounts for *all* individual time-constant background. Common changes over time are captured by the time effect. Moreover, within the stayer sample the independence assumption $U \perp T | G$ in the CIC model seems very plausible due to the short panel structure of the data, i.e., the distribution of unobserved factors like health status or preferences likely remains constant within group (HMO or regular insurance) over the three years. The restriction to stayers affected 1.96% of the total sample, or 1,686 out of the 85,626 observations.[4] Even though this fraction is relatively low, we evaluate whether selection out of HMO may pose a problem to the external validity of our results, i.e., whether the selection is in any way related to the introduction of cost-sharing in HMO.

---

[4]Table A1 in the appendix shows the summary statistics for the restricted stayer sample. The statistics are very similar to those in the full sample in Table 1.

Table 3 shows that out of the 4,909 insured in HMO in 2002, 84% stayed in HMO, 6.7% switched to the regular plan, and 9.3% switched out of CSS. Similar numbers are observed for the switching from 2003 to 2004. We then estimated a standard multinomial logit (MNL) model for the outcomes *stay in HMO*, *switch to regular plan*, and *switch out of CSS* for the 2002/03 and 2003/04 samples. The covariates are gender, age, city of residence, and lagged health costs. Table 3 shows the marginal/discrete probability effects for the three outcomes. The numbers show by how much the probability of observing any of the outcomes changes with a unit change in the regressor. For example, the estimate -0.107 for Basel can be interpreted as a 10.7%-points smaller probability of staying in HMO for those living in region Basel as opposed to the base region Zurich. Since many of the probability effects are statistically significant, the results suggest an endogenous selection out of HMO. However, important for our identification strategy is whether the pattern of effects changed from 2002/03 to 2003/04. Comparing the MNL model results in the two samples, there is little evidence that the selection is systematically related to the introduction of cost-sharing in HMO, i.e., the probability effects in Table 3 do not significantly differ between the 2002/03 and 2003/04 samples.

— Insert Table 3 about here —

Hence, we conclude that overall there is little evidence that would suggest that the introduction of cost-sharing in HMO caused different selection mechanisms than we would observe otherwise. For our DID/CIC analysis this implies that restricting the sample to stayers within their plans is a reasonable choice for the baseline sample. Any impact that we find in the 2002/03 comparison can then be attributed to having full insurance in HMO.

5.2 Impact of full insurance on health cost

Table 4 displays the main results of the paper. Consider first the estimates of the effect of full insurance on total expected health cost in the DID models (columns 1 and 2). The linear model specification predicts an increase in health cost of about CHF 440 per year for

the HMO insured when having full health insurance. By contrast, the GLM model with an exponential specification only predicts a change of about CHF 230. The difference in these two predictions is statistically significant, and the question is which model, if any, provides an unbiased estimate of the causal effect of full insurance.

— Insert Table 4 about here —

Ultimately, this depends on whether the common trend assumption seems more plausible for one or the other specification. In judging that we use the "control" year 2004 and evaluate the estimates of the placebo effect. In the linear specification, we find a significant negative placebo effect. In the exponential model, the placebo effect is close to zero and statistically insignificant. The latter is what we would expect, a priori, because both groups (HMO and regularly insured) face identical cost-sharing options in 2003 and 2004. Moreover, the simple descriptive comparisons of Table 2 suggest that a constant growth rate in health cost by group is more reasonable in our data than a linear trend. Hence, the results support that the GLM model gives an unbiased estimate of the causal effect of full insurance, while the linear DID model likely does not fulfill the common trend assumption. This conclusion is also supported by the results of the CIC model (which is invariant to monotonic transformations of the cost variable) and the results for the extensive margin.

The mean effect in the CIC model (column 4) is very close to the DID estimate in the GLM model. In addition, we confirm the zero placebo effect. Regarding the extensive margin (column 3), the placebo effect is zero, too. The effect on the likelihood of generating no cost is about -0.073, which implies that having full insurance increases the probability of having positive cost by about 7.3%-points for the HMO insured. The effect is both statistically significant and economically important. Given that the fraction of positive cost in the HMO group in 2003 is about 75%, the effect corresponds to an increase of almost 10%.

The results in Table 4 focus on the average effects of having full health insurance on mean expenditures and on the extensive margin of the HMO group. In addition to that, the

14

CIC model allows us to estimate the effects of full health insurance on the entire outcome distribution by looking at quantiles. Figure 2 shows the CIC effects for the 0.05 to 0.85 quantiles for the impact of full insurance on the HMO group (left graph), and the placebo effect (right graph). The placebo effect is zero over the entire distribution, as we would expect. There is no effect of having full insurance below the 0.2 quantile (because that part of the outcome distribution does not change). The effect is positive and almost constant in absolute terms at around CHF 100 to 150 above the 0.2 quantile. In relative terms, the CIC effects are decreasing over the outcome distribution. The CIC effects become insignificant above the 0.7 quantile and are rather noisy for the upper tail of the distribution.

— Insert Figure 2 about here —

To put the effects of having full insurance into perspective, we calculate price elasticities from the numbers obtained above. The usual formula for the coefficient of price elasticity relates the change in demand (here measured by health expenditures) to a marginal change in the price. Since having full insurance implies a zero price, the traditional formula is not applicable. Instead, we use the concept of an arc elasticity where the percentage change in demand is related to the percentage change in the price:[5]

$$E_d = \frac{\% \text{ change in } Q}{\% \text{ change in } P} = \frac{\frac{Q_{full} - Q_{cost}}{Q_{cost}}}{\frac{P_{full} - P_{cost}}{P_{cost}}} = -\frac{Q_{full} - Q_{cost}}{Q_{cost}} \tag{12}$$

Q_{full} denotes health care expenditures under full insurance, Q_{cost} denotes expenditures under the cost-sharing regime (analogous for prices P_{full} and P_{cost} in the denominator). Numbers in the numerator can be extracted from our estimates in Table 4, e.g., for the DID effect in the GLM model we can calculate $232.6/1563.9 \approx 0.148$, which is relative to the 2003 average in the HMO group (see Table 2).[6] The idea then is to compare the zero price under full insurance to the price P_{cost} under cost-sharing. This yields a denominator of -1, independent of the value

---

[5]We do not use the arc elasticity with respect to the midpoint, as in Aron-Dine et al. (2013), but to average expenditures and prices under the cost-sharing regime, which we deem more reasonable in our case. Using arc elasticities the denominator would be doubled to -2.

[6]In fact, the GLM model allows us to approximate the numerator directly by extracting the estimate of δ_1 in eq. (9) from the estimation output.

of P_{cost} , i.e., the price elasticity is -0.148. Regarding the mean effect in the CIC model, we obtain a price elasticity of about -0.124, close to the elasticity in the GLM model.

For the quantile CIC effects, the elasticities range from an almost unit elasticity for the 0.3 quantile of the cost distribution to an elasticity of about -0.1 for the 0.7 quantile. Intuitively, this is what we would expect because introducing a cost-sharing element to basic health insurance will most likely affect individuals in the lower ranges of the cost distribution, who possibly save an optional check-up, but hardly influences those in the upper range.

In a secondary analysis (detailed results are available upon request), we estimated GLM models conditional on low (CHF 230) and high (CHF 1200 or 1500) deductible choices in 2003. The level of the chosen deductible may be interpreted as an indicator for health status. This yields an elasticity of -0.11 for the low deductible and -0.57 for the high deductible, consistent with the CIC results. Since deductible choice is likely endogenous to health care utilization, however, we do not want to place too much emphasis on these conditional GLM results.

5.3 Sensitivity analysis

The results of the previous section are remarkably robust. A first concern could be the restriction of the sample to stayers in their health plan over the full observation period. Table 5 shows that when including the switchers in the analysis, the effects remain stable. Since we have panel data we include individual fixed effects in the DID models to control for time-constant characteristics, beyond those available in the data (gender, age, and city of residence). Unfortunately, allowing for fixed effects is not straightforward in the CIC model, but still the estimates in the full sample are very close to those in the stayer sample. This holds for the mean as well as the quantile CIC effects (see Figure A1 in the appendix).

— Insert Table 5 about here —

A second concern could be that the mean effects are driven by outliers, i.e., those with large health costs, or severe health problems. We have information about whether a hospital

16

stay occurred in the past year. Excluding those from the sample hardly affects our results. In the same way, excluding the top 1% of the cost distribution has little impact.

A last concern from the results in Table 3 could be that there is regional heterogeneity. In particular, we find some weakly significant differences (p-value 0.074) between the 2002/03 and 2003/04 samples in the MNL models that are driven by the switching out of HMO behavior in Berne. When we exclude those insured from the sample, then the estimated effects are a bit smaller, but still in the order of Table 4. Overall, we conclude that our results are not very sensitive to these changes in the baseline sample, and that our DID/CIC estimates provide robust evidence of the effect of full insurance on health expenditures.

6   Discussion and conclusion

Ellis (2012) in his recent presidential address to the *American Society of Health Economists* asked five questions. The final one was concerned with the behavioral aspects of economics, and the psychological effect of a zero price in particular. He argues that the incentives played by a zero price are fundamentally different from any positive price. Having full health insurance implies a zero price of utilizing health care.[7] This in turn, as Ellis (2012) puts it, "liberates your mind to not even think about cost [p. 231]". Our paper contributes to the literature by comparing a managed-care plan with temporary full insurance to a cost-sharing alternative with deductibles, co-payments, and a stop-loss. We find that having full insurance has a significant impact on the extensive margin, i.e., on the likelihood to generate costs. Thus, having to pay for health care services, from the very first visit in the year, has a strong influence on the consumers' threshold to generate costs.

Health care demand is more sensitive for the lower cost ranges, with elasticities estimated about -1.0 at low quantiles of the health cost distribution to -0.1 at higher quantiles. This finding suggests that primarily low cost patients react to the price change, often reducing

---

[7]Of course, there can still be monetary costs of going to the doctor (e.g., cost for transport, lost work time, etc.), and there are non-monetary costs as well. However, we abstract from these here because we expect that consumers care most about the price they have to pay for physician services.

their demand to zero. For the average cost, the elasticity is around -0.14. This estimate is well in the range of the RAND health insurance experiment and other previous findings. A recent study for Switzerland using data of the same insurance company is Trottmann et al. (2012). Transforming their estimates of the moral hazard effects (Table 5 of their paper) into elasticities gives an estimate -0.33 when individuals with regular contract and high deductible were to have the basic deductible instead.[8]

Our results are informative regarding optimal insurance designs. As discussed in Cutler and Zeckhauser (2000) health insurances are confronted with an inherent trade-off between risk-sharing and moral hazard. Our results can be interpreted in this light. Permitting consumers the access to health services without any cost will increase utilization significantly, especially at the extensive margin. Hence, we provide further evidence that cost-sharing instruments are an effective tool to reduce overconsumption of medical care (e.g., Newhouse et al., 1993). However, given that the cost-saving effects are concentrated at the bottom of the health cost distribution, the overall impact on cost containment may be limited.

Acknowledgments

---

[8]To compute the elasticity we used the realized end-of-year price of 1 if the deductible is high (mean expenditure = 1100 < deductible) and of 0.1 if the deductible is low (the counterfactual mean expenditure = 1422 > deductible). This corresponds to a percentage change of -0.9. The percentage increase in mean expenditure estimated by Trottmann et al. (2012) is 0.29.

References

Angrist, Joshua D, "Estimation of limited dependent variable models with dummy endoge- nous regressors," *Journal of Business & Economic Statistics*, 2001, *19* (1).

Aron-Dine, Aviva, Liran Einav, Amy Finkelstein, and Mark R Cullen, "Moral hazard in health insurance: How important is forward looking behavior?," Working Paper 17802, National Bureau of Economic Research 2012.

_ , _ , **and** _ , "The RAND Health Insurance Experiment, Three Decades Later," *Journal of Economic Perspectives*, 2013, *27* (1), 197–222.

Athey, Susan and Guido W Imbens, "Identification and Inference in Nonlinear Difference- in-Differences Models," *Econometrica*, 2006, *74* (2), 431–497.

Borah, Bijan J, Marguerite E Burns, and Nilay D Shah, "Assessing the impact of high deductible health plans on health-care utilization and cost: a changes-in-changes approach," *Health Economics*, 2011, *20* (9), 1025–1042.

Buntin, B. M. and A.M. Zaslavsky, "Too much ado about two-part models: Comparing methods of modeling Medicare expenditures," *Journal of Health Economics*, 2004, *23*, 525– 542.

Cutler, David M and Richard J Zeckhauser, "The anatomy of health insurance," *Hand- book of Health Economics*, 2000, *1*, 563–643.

Eichner, Matthew J, "The demand for medical care: What people pay does matter," *The American Economic Review*, 1998, *88* (2), 117–121.

Ellis, Randall, "Five questions for health economists," *International Journal of Health Care Finance and Economics*, September 2012, *12* (3), 217–233.

Ellis, Randall P, "Rational behavior in the presence of coverage ceilings and deductibles," *The RAND Journal of Economics*, 1986, pp. 158–175.

Gourieroux, Christian, Alain Monfort, and Alain Trognon, "Pseudo maximum likelihood methods: Theory," *Econometrica: Journal of the Econometric Society*, 1984, pp. 681–700.

Imbens, Guido W and Jeffrey M Wooldridge, "Recent developments in the econometrics of program evaluation," *Journal of Economic Literature*, 2009, *47* (1), 5–86.

Keeler, Emmett B and John E Rolph, "The demand for episodes of treatment in the health insurance experiment," *Journal of Health Economics*, 1988, *7* (4), 337–367.

_ , **Joseph P Newhouse, and Charles E Phelps**, "Deductibles and the demand for medical care services: The theory of a consumer facing a variable price schedule under uncertainty," *Econometrica: Journal of the Econometric Society*, 1977, pp. 641–655.

Kowalski, Amanda, "Censored Quantile Instrumental Variable Estimates of the Price Elasticity of Expenditure on Medical Care," Working Paper 15085, National Bureau of Economic Research 2009.

Marsh, Christina, "Estimating Health Expenditure Elasticities Using Nonlinear Reimbursement," Technical Report, University of Georgia 2012.

Mullahy, John, "Much Ado about Two: Reconsidering Retransformation and Two-Part Model in Health Economics," *Journal of Health Economics*, 1998, *17*, 247–281.

Pauly, Mark V, "The economics of moral hazard: comment," *The American Economic Review*, 1968, *58* (3), 531–537.

Trottmann, Maria, Peter Zweifel, and Konstantin Beck, "Supply-side and demand-side cost sharing in deregulated social health insurance: Which is more effective?," *Journal of Health Economics*, 2012, *31* (1), 231–242.

Wooldridge, Jeffrey M, *Econometric Analysis of Cross Section and Panel Data (2. ed.)*, The MIT Press, 2010.

Tables and Figures

Table 1: Summary statistics of CSS panel

	HMO Mean	SD	Regular Mean	SD
A. Outcomes				
Effective health costs	1723.5	4322.3	4292.5	7854.1
Fraction of zero cost	22.0%		16.1%	
B. Background characteristics				
Age in 2002	46.9	16.0	56.8	18.3
Female	47.1%		60.9%	
Basel	10.5%		14.0%	
Berne	14.6%		8.8%	
Lucerne	62.7%		32.0%	
Zurich	12.2%		45.2%	
Number of observations	10,713		74,913	

Source: CSS panel, own calculations. *Notes:* Statistics shown for all individuals observed over full period 2002-2004. HMO as opposed to regular health care plan. Effective health costs in Swiss Francs per year. SD = standard deviation.

Table 2: Comparison of health costs by plan and year

	Year 2002 (1)	Year 2003 (2)	Year 2004 (3)
A. Effective health costs			
HMO	1678.6	1563.9	1936.6
	(62.1)	(62.1)	(90.7)
Regular	3741.3	4102.8	5026.7
	(43.4)	(46.2)	(57.7)
B. Fraction of zero cost			
HMO	0.165	0.250	0.249
	(0.006)	(0.007)	(0.007)
Regular	0.162	0.161	0.160
	(0.002)	(0.002)	(0.002)
Number of observations	28,542	28,542	28,542
Fraction HMO	12.9%	12.5%	12.1%

Source: CSS panel, own calculations. *Notes:* See notes Table 1. Reported numbers are means, standard errors in parentheses.

Table 3: Switching out of HMO - Results of multinomial logit model

	2002/03 (1)			2003/04 (2)		
	Stay	Regular	Out	Stay	Regular	Out
Age in 2002	0.002**	-0.0004	-0.002**	0.002**	-0.001*	-0.001**
	(0.0003)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0002)
Female	-0.001	0.009	-0.008	-0.006	0.011	-0.005
	(0.010)	(0.007)	(0.008)	(0.009)	(0.007)	(0.007)
Basel	-0.107**	0.095**	0.012	-0.064**	0.051*	0.013
	(0.026)	(0.024)	(0.013)	(0.024)	(0.021)	(0.012)
Berne	0.047**	-0.023	-0.024*	0.041**	0.002	-0.043*
	(0.016)	(0.013)	(0.011)	(0.015)	(0.013)	(0.007)
Lucerne	0.063**	0.019	-0.082**	0.043**	0.002	-0.045**
	(0.016)	(0.011)	(0.012)	(0.014)	(0.010)	(0.011)
Log cost previous year	-0.005**	0.003	0.002	-0.003*	0.002	0.001
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Number of observations	4,124	330	455	3,824	214	265
χ^2-statistic (1)=(2)	19.7 (p-value 0.074)					
χ^2-statistic (1)=(2)	2.26 (p-value 0.994), excluding region Berne					

Source: CSS panel, own calculations. *Notes:* 2002/03 looks at all individuals in HMO in 2002 and their choice of plan in 2003 (stay in HMO, switch to regular plan, leave CSS insurance), analogous for 2003/04. Reported numbers are probability effects. Standard errors in parentheses. ** $p < 0.01$ * $p < 0.05$.

Table 4: Difference-in-differences results for HMO versus Regular

| | DID | | | CIC |
	Linear	GLM	P(0)	Mean
$ATT^{2002-2003}$	436.9**	232.6**	-0.073**	193.8**
	(80.6)	(77.7)	(0.009)	(70.0)
$PE^{2004-2003}$	-548.5**	30.2	-0.000	43.0
	(90.7)	(85.5)	(0.009)	(64.9)
Number of observations	55,960			

Source: CSS panel, own calculations. *Notes:* 2002-2003 shows impact of full insurance on health cost using DID/CIC comparisons of cost trends in HMO (no deductible in 2002, deductible in 2003) with cost trends in regular plan (deductibles in both years). 2004-2003 shows placebo effect. Results conditional on not switching plans (between HMO and regular). Cluster adjusted/bootstrapped standard errors in parentheses (bootstrap based on 500 replications).   ** $p < 0.01$ * $p < 0.05$.

Table 5: Difference-in-differences results – full sample

| | DID | | | CIC |
	Linear	GLM	P(0)	Mean
$ATT^{2002-2003}$	430.5**	219.4**	-0.086**	239.7**
	(80.6)	(76.1)	(0.008)	(84.1)
$PE^{2002-2003}$	-552.7**	33.3	-0.000	56.8
	(90.7)	(82.5)	(0.009)	(117.2)
Number of observations	57,084			

Source: CSS panel, own calculations. *Notes:* See notes Table 4. Results including switchers between health plans. DID models include individual fixed effects. Cluster adjusted/bootstrapped standard errors in parentheses.   ** $p < 0.01$ * $p < 0.05$.

Table 6: Difference-in-differences sensitivity checks

|	DID			CIC
	Linear	GLM	P(0)	Mean
Exlude hospital stayers[1]	384.3**	254.8**	-0.073**	224.8**
	(77.3)	(74.3)	(0.009)	(67.3)
Exlude top 1% costs[2]	414.3**	268.0**	-0.073**	225.1**
	(58.0)	(55.9)	(0.009)	(49.4)
Exlude Berne[3]	402.7**	199.9**	-0.060**	163.4**
	(88.9)	(86.3)	(0.009)	(78.3)
Number of observations	54,608[1]			
	54,704[2]			
	50,718[3]			

Source: CSS panel, own calculations. *Notes:* See notes Table 4. DID/CIC effects for $ATT^{2002-2003}$. Results conditional on not switching plans. Cluster adjusted/boot-strapped standard errors in parentheses. ** $p < 0.01$ * $p < 0.05$.

Figure 1: Cumulative distribution of log health costs by plan and year
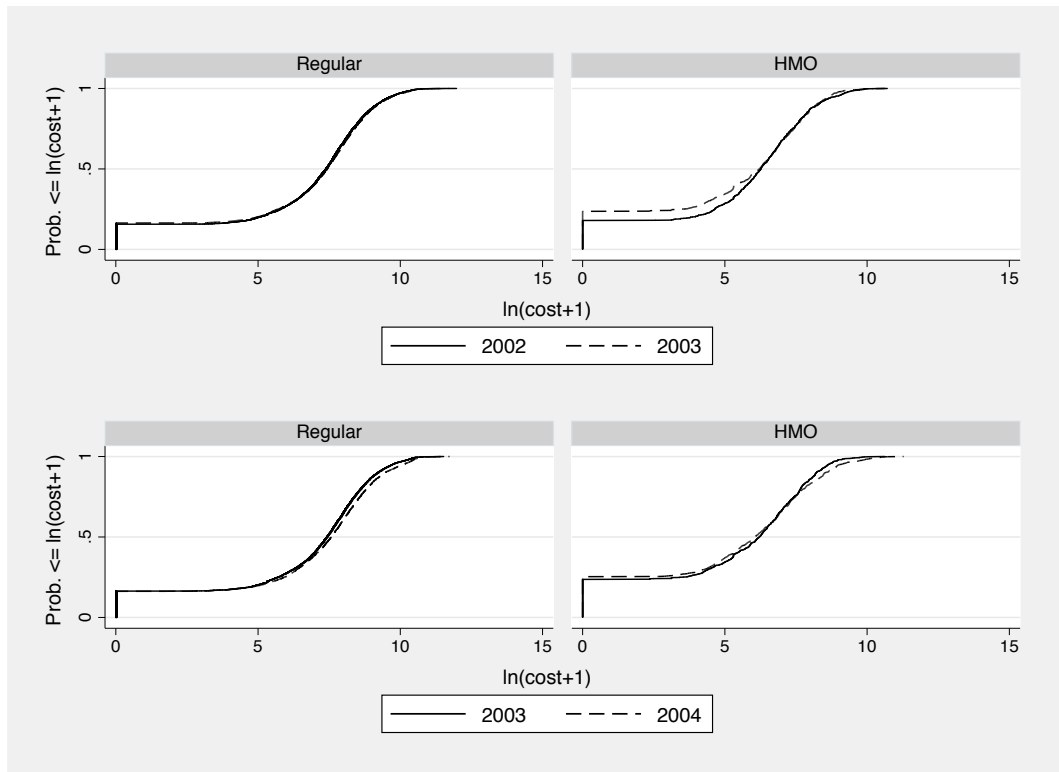
Source: CSS panel, own calculations. *Notes:* Zero deductibles in HMO in 2002, non-zero deductibles in HMO in 2003/04, and in regular plans for all years.

Figure 2: Changes-in-changes results for different percentiles
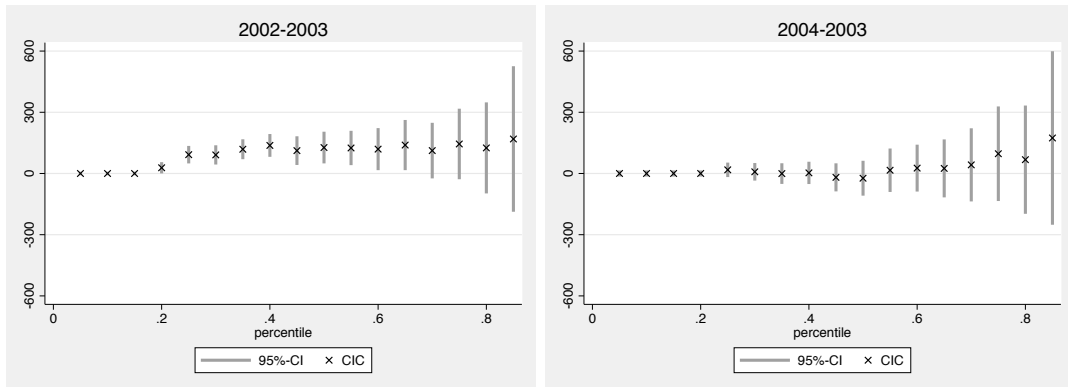
Source: CSS panel, own calculations. *Notes:* Horizontal axis: percentile of effective health costs; vertical axis: CIC estimate of treatment effect. Results conditional on not switching plans (between HMO and regular). Shaded bars show 95% confidence intervals (based on bootstrapped standard errors, 500 replications).

Appendix

Table A1: Summary statistics - stayers only

	HMO Mean	SD	Regular Mean	SD
A. Outcomes				
Effective health costs	1749.4	4403.3	4317.5	7876.4
Fraction of zero cost	21.9%		16.0%	
B. Background characteristics				
Age in 2002	47.1	16.1	56.9	18.3
Female	46.8%		61.0%	
Basel	9.7%		13.9%	
Berne	13.9%		8.8%	
Lucerne	63.7%		31.7%	
Zurich	12.7%		45.6%	
Number of observations	9,852		74,088	

Source: CSS panel, own calculations. *Notes:* See notes Table 1.

Figure A1: Changes-in-changes results – full sample
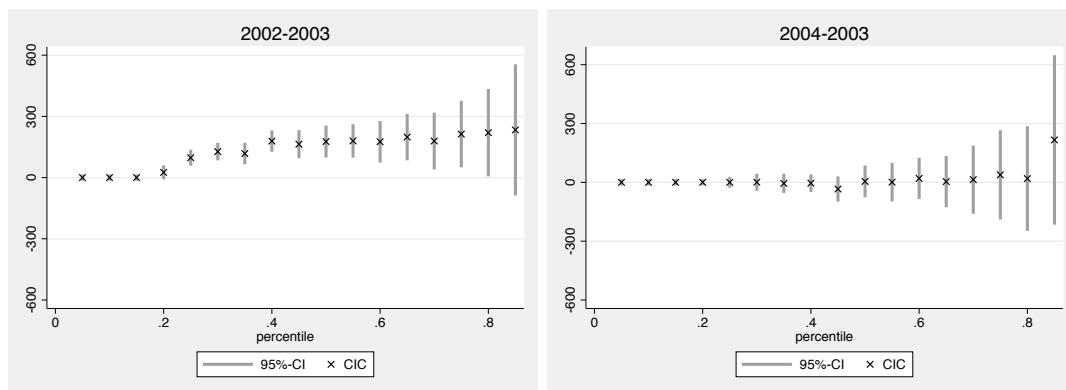
Source: CSS panel, own calculations. *Notes:* See notes Figure 2. Results are shown for full sample, i.e., including switchers between health plans.