

HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

THE UNIVERSITY *of York*

WP 14/22

Health care demand in the presence of discrete price changes

Michael Gerfin, Boris Kaiser & Christian Schmid

August 2014

<http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/>

Health Care Demand in the Presence of Discrete Price Changes

Michael Gerfin^a Boris Kaiser^b Christian Schmid^b

May 30, 2014

Abstract

Deductibles in health insurance generate nonlinear budget sets and dynamic incentives. This paper uses detailed individual claims data from a large Swiss insurance company to estimate the response in health care demand to the discrete price increase that is generated by resetting the deductible at the start of each calendar year. We use a regression discontinuity type framework based on daily data to estimate the change in health care demand right before and right after the turn of the year. We find that for individuals with high deductibles health care demand drops by 27%, which translates into an elasticity of $-.21$. The decrease is most pronounced for inpatient care and prescription drugs. By contrast, for individuals with low deductibles there is no significant change in health care demand (except for prescription drugs). A remaining open question is whether the observed behavioral responses can be attributed to intertemporal substitution or whether they constitute a classic moral hazard effect.

JEL Classification: C31, D12, I13

Keywords: Health care demand; nonlinear pricing; dynamic incentives; health insurance.

Acknowledgements: We thank Blaise Melly and participants at the Laax Labor conference and the annual congress of the Swiss Society of Economics and Statistics for helpful discussions.

^aDepartment of Economics, University of Bern, Switzerland. E-mail: michael.gerfin@vwi.unibe.ch.

^bDepartment of Economics and Center for Regional Economic Development (CRED), University of Bern

1 Introduction

The microeconomic analysis of the demand for health care is important and interesting for at least two reasons. First, from an efficiency point of view, there is the question of how price sensitive this demand is conditional on health, because health insurance reduces the price of health care for patients and may induce moral hazard. Second, most health insurance plans feature non-linear price schedules due to deductibles and stop-losses. The analysis of consumers' behavior in the presence of non-linear pricing is an important and challenging topic for applied work. Apart from health care demand, examples include electricity (Borenstein, 2009; Ito, forthcoming) and water demand (Ito, 2013). These papers usually find that individuals seem to respond to average rather than to marginal prices, if they respond at all. Also closely related is the analysis of labor supply in the presence of non-linear taxation (Hausman, 1985; Saez, 2010; Chetty et al., 2011).

In health economics, there is a large literature on estimating the price sensitivity of health care demand. The first reliable 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). Their analysis is based on the assumption that individuals are myopic and base their health care demand on the spot price, ignoring the non-linearity of the budget constraint created by a stop-loss.¹

More recent approaches explicitly take the non-linear pricing into account, but assume perfect forward-looking behavior. Examples include Kowalski (2012), Marsh (2013) and Einav et al. (2013b). Aron-Dine et al. (2012) are the first to explicitly test for forward-looking behavior. They exploit that individuals entering a health plan during a calendar year have less time to accumulate health care expenditure and therefore have a smaller probability of reaching the deductible. Hence these individuals face a higher expected end-of-year price compared to those who already were in the health plan at the beginning of the year. They

¹To be precise, Keeler and Rolph (1988) assume that agents have perfect foresight regarding medical care within an illness episode, but do not consider the possibility of future health shocks.

find that initial demand is lower for individuals who face the same marginal price but higher expected end-of-year prices. Einav et al. (2013a) apply the same basic idea to drug expenditure in Medicare Part D. They find significant bunching of annual spending around the convex kink in the budget set created by the famous Part D “donut hole”, which is evidence for a static price response. They also document a dynamic price response in the sense that initial drug use is lower for individuals in the same contract who face the same initial spot price of drugs but higher expected end-of-year prices.

For Switzerland, there is a series of papers that analyze the effect of deductibles on health care demand. Examples include Gerfin and Schellhorn (2006) and Trottmann et al. (2012). However, to our knowledge, only Boes and Gerfin (2013) explicitly estimate an elasticity of health care demand for Switzerland. Their analysis is based on a quasi-experimental setup where part of the clients of a Swiss insurance provider were subject to full insurance for a limited time. Their baseline estimate of the price elasticity is $-.14$.

Our analysis exploits the institutional regulation that deductibles apply to a calendar year and reset to the initial level at the start of the new year. This generates a discrete jump in the price of medical care for individuals who have hit their deductible in the old year. We use a regression discontinuity type framework based on daily insurance-claims data to estimate the change in health care demand right before and right after the turn of the year. We find that for individuals with high deductibles health care demand drops by 27%, which translates into an elasticity of $-.21$. The decrease is most pronounced for inpatient care and prescription drugs. By contrast, for individuals with low deductibles, there is no significant change in health care demand, with the exception of prescription drugs.

Common to virtually all empirical estimates of the elasticity of health care demand is that annual outcomes are studied. By contrast, our analysis is based on daily health care expenditure and daily prices. This allows us to identify the behavioral change at the moment of the price change. This is our major contribution to the literature. Of course, this begs the question of comparability with the literature. Our results are informative about behavioral responses in the presence of price changes caused by the institutional setting. However, it turns out to be difficult to separate the observed responses into a true moral hazard effect (health care that would not have been consumed if the price was higher) and an intertemporal substitution effect (health care that would have been consumed next year instead of this year). In this sense, our results may be

difficult to compare with the literature.

The remainder of this paper is organized as follows. Section 2 provides a theoretical framework describing individuals' behavioral response to a discrete jump in the price for medical care at the end of the year. In Section 3, we summarize the institutional background. Section 4 contains the empirical analysis: we explain the construction of the dataset, present descriptive statistics, discuss the evolution of the spot price and health care costs, describe the empirical framework, and discuss the main results (in Section 4.5). Section 5 contains some concluding remarks.

2 Theoretical Framework

Consider a two-period framework, where each period (e.g. year) consists of T subperiods (e.g. days) denoted by $t = 1, 2, \dots, T$. Hence, there are $2T$ subperiods, with $T+1$ the first subperiod in the second period. Health care demand decisions are made within each subperiod. Agents choose a health insurance plan before the start of each period. For simplicity, assume that there are only two plans, one with a low deductible and one with a high deductible and a lower premium. Common to both plans is a co-payment rate γ , which applies to all health care expenditure exceeding the deductible (i.e. there is no stop-loss). We do not model plan choice, but assume that agents choose the plan that maximizes expected utility during the deductible period.

Assume that for an agent with the high deductible the only reason for exceeding the deductible is a health shock that requires health care expenditures above the deductible.² Denote the ex ante probability of this shock with $\Pi = \sum_{t=1}^T \pi_t$, where π_t is the shock probability in subperiod t . If a shock occurs, there are two regimes in which the agent can be during the period, denoted as $s \in \{b, a\}$ (*before* and *after* the shock). The shock subperiod itself is not explicitly modelled.³

Individuals derive utility from medical care m and a composite consumption good c . Suppressing all subscripts for the moment, the per-subperiod utility function is given by

$$U(m, c) = u(m) + c, \tag{1}$$

where we assume positive but diminishing marginal utility of medical care. Units

²Agents who choose the low deductible are assumed to exceed their deductible with probability one. For this reason we focus on agents who choose the high deductible .

³In the shock period, health care expenditures are exogenously determined by the necessity of treating the shock.

of health care are defined such that the marginal cost of providing one unit of health care is normalized to unity. In addition, the composite consumption good has price one.

We now characterize the agents' optimal behavior in this simple framework. At the beginning of each subperiod, individuals learn whether a health shock has been realized or not, and then choose the consumption for this subperiod accordingly. In periods except the shock-period, optimal demand of medical care is determined by the following first order condition

$$\frac{\partial u(m_{st})}{\partial m_{st}} = p_{st}, \quad (2)$$

where p_{st} is the price of medical care that is relevant to the individual in regime s and period t . Due to the quasi-linear utility in (1), health care demand before and after the shock is only a function of the price, i.e. $m_{st} = m^*(p_{st})$. This demand function is decreasing in the price p_{st} . After hitting the deductible, the price is $p_{at} = \gamma$. However, the price before hitting the deductible, p_{bt} , depends on whether the individual is myopic or forward-looking.

If individuals are myopic, they only respond to the marginal (or spot) price, which is the price they would pay for the next additional unit of medical care. In the present case, myopia implies that $p_{bt} = 1$. If, on the other hand, individuals are forward-looking, they understand the dynamic incentives created by the deductible. In our setup, the expected end-of-year price in period t , p_{bt}^{eoy} , can be approximated by $(1 - \Pi_t) \cdot 1 + \Pi_t \cdot \gamma$ which is obviously smaller than 1 if $\Pi_t > 0$. The probability $\Pi_t = \sum_{j=t+1}^T \pi_j$ is the cumulated probability that the agent hits the deductible after period t . The intuition behind the difference in the two prices is that a forward-looking individual who has not received a shock yet anticipates that with probability Π_t a shock may occur in subsequent subperiods, which would lower his price in all periods after the shock.

Consider now the situation at the change of year, i.e. the subperiods T and $T + 1$. For an individual who experiences a shock at some point during the first year, her relevant price in period T , p_T , equals γ . At the beginning of the second year, the deductible resets, inducing a change in the price which is now given by

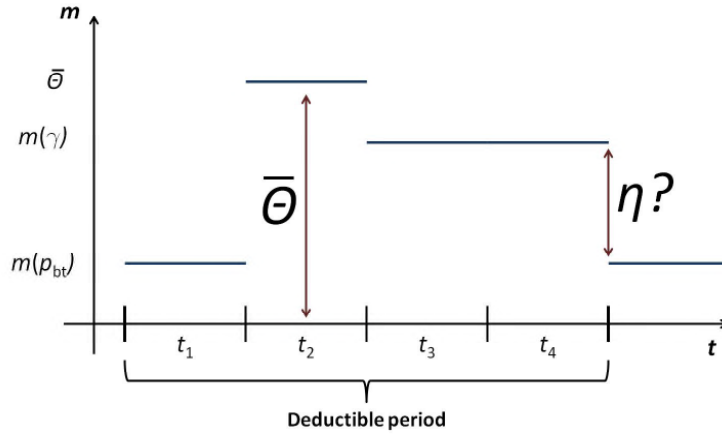
$$p_{T+1} = \begin{cases} 1 & \text{if agent is myopic} \\ (1 - \Pi_{T+1}) + \Pi_{T+1} \cdot \gamma & \text{if agent is forward looking} \end{cases}$$

where $\Pi_{T+1} = \sum_{j=T+2}^{2T} \pi_j$. If Π_{T+1} is close to 1, p_{T+1} is close to γ if the agent

is forward looking. This is the case for less healthy individuals. If, on the other hand, Π_{T+1} is significantly below 1 (healthy individuals), there is a significant increase in the price of health care, even if agents are forward looking. Note that $\Pi_{T+1} \approx \Pi$ if the number of subperiods is large.

We use this setup to estimate the price responsiveness of the demand for medical care at the change of year. For an agent with a health shock in the first year and assuming that she does not experience a health shock in the first subperiod of the second year, the observed change in health care consumption before and after the change of year is solely due to the price change. Figure 1 illustrates the dynamics of health care demand during the first period with a health care shock of magnitude $\bar{\theta}$ in subperiod 2. The price effect we aim to estimate is denoted by η .

Figure 1: Health care demand before, during, and after the health shock



It is clear that the elasticity depends on the price which individuals consider for their health care consumption. We have seen that $p^{coy} \leq p^{spot}$ as forward-looking individuals anticipate that they might exceed the deductible later in the year. This implies that the potential price jump at the change of year is smaller for forward-looking individuals than for myopic individuals. Likewise, the potential demand response is also smaller because demand is decreasing in the price. We therefore do not know whether the elasticity is smaller or larger for forward-looking individuals compared to myopic individuals. However, if we impose the assumption that all patients are either completely myopic or completely forward-looking, we can determine the two corresponding price elasticities unambiguously. Since the patients' foresight is likely to lie in-between these two polar cases (as shown in Aron-Dine et al., 2012), we are able to provide a lower

and upper bound to the underlying price elasticity of health care demand.

If we apply the insights from our theoretical considerations to a practical context, it implies that (i) we should study the consumption dynamics at the change of year, (ii) we should consider those patients who hit the deductible during the year, and (iii) we must exclude those patients who hit the deductible late in the year such that observed end-of-year consumption is no longer influenced by the health shock that initially pushed them over the deductible level.

3 Institutional Background

The Swiss health insurance system can broadly be classified as “managed competition”. It consists of mandatory insurance, which is regulated by government, and supplementary insurance, which is free-market oriented. The former dominates the health insurance system because it covers an extensive range of medical services and pharmaceutical products. Mandatory insurance plans are offered by about 60 private insurance companies, but are subject to strong government regulation. Insurers may not make profits on mandatory insurance plans and are obliged to accept all patients who wish to enroll regardless of health status, age, sex, etc. Apart from the baseline contract, most insurers offer several forms of managed-care plans (e.g. PPO, HMO) which, for example, restrict the patients’ choice of physicians in exchange for lower premiums. Premiums may vary by region and across the three age groups <18 , $18-25$ and ≥ 25 years, but not *within* these age groups.⁴ Adult patients can freely choose among several deductible levels (normally CHF 300, 500, 1000, 1500, 2000 and 2500).⁵ All contracts are on an individual basis, hence there are no family-related shared deductibles. While premiums are decreasing in the deductible level, the maximum reduction relative to the premium associated with the baseline deductible of CHF 300 is subject to regulation.⁶ For medical costs exceeding the deductible, all individuals face a co-payment rate of 10%, which drops to zero after the co-payments exceed the stop-loss amount of CHF 700. As an important feature of the system, individuals can change insurance company, switch to another plan and/or adjust the deductible level only at the beginning of every new calendar year, provided that

⁴The exact age-group cut-offs are determined by birthday dates. According to federal regulations, premiums usually rise on the patient’s 18th and 25th birthday.

⁵The deductible levels for children (age 18 and younger) are lower.

⁶The maximum premium reduction is 80% of the difference in deductibles.

these changes are applied for by November 30th in the preceding year.⁷ As a consequence, the insurance period corresponds to the calendar year, which also means that the deductible resets at the beginning of every new year. This feature implies that people exceeding the deductible in a given year experience a discontinuous rise in the price of medical care at the beginning of the subsequent year.

4 Empirical Analysis

4.1 Construction of the Dataset

The data is drawn from the records of a large Swiss health insurance company with roughly half a million customers and covers the time period from 2008 to 2011. It includes all individual insurance claims from patients enrolled in mandatory health insurance plans. For each claim, we observe medical costs, the number of visits, beginning and end of treatment, the type of service (inpatient or outpatient treatment, laboratory, drugs, physiotherapy, etc.) and the type of provider (hospital, physician, pharmacy, etc.). For each patient, we have information on sex, age, region of residence, choice of insurance plan and deductible, and the periods of enrollment.

In the analysis below, we will study the evolution of health care utilization across time using day-level data. We therefore need a sensible mechanism of allocating the costs of a treatment across the observed treatment spell because we have no information as to how and when costs are incurred *within* the treatment spell. Of course, if the treatment spell is only a single day, the observed costs are simply allocated to this day. (This is the case for about 64% of all treatments.) However, if the treatment spell involves more than one day, we decide to spread the costs uniformly across the days of the treatment spell. For example, if a patient obtained outpatient care from September 1st to September 20th, which involved several consultations and a total bill of CHF 500, we allocate costs of CHF 25 to each of the 20 days of the treatment spell. This procedure also has the advantage that undesired “calendar effects” arising from the billing behavior of health care providers are eliminated: certain lengthy or recurring treatments (e.g. old-age care, physiotherapy, treatments for chronically ill patients) are often found to be terminated on the last day of the month, but a new treatment

⁷Exceptionally, an *increase* in the deductible can be filed until December 31st of the preceding year.

spell re-starts on the next day. Health care providers do this most likely for reasons of accounting.⁸ As a result, we observe substantial heaping in the number of recorded ends of treatment on certain days (typically on the last day of the month, and more markedly, on the last day of the quarter). By spreading health care costs uniformly across spells ensures that such accounting effects are eliminated.

From the patient pool, we only select patients who were enrolled in mandatory health insurance during the entire two-year period under consideration (because we require two consecutive years for the analysis). This excludes patients who died, emigrated or immigrated, as well as those who were enrolled at another insurance company in one of the two years. Moreover, we only include patients who are at least 18 years of age. In addition, we exclude women who had childbirth-related medical costs because these are exempt from the deductible. These selection criteria produce our baseline sample which comprises about 360,000 people in the two-year period 2009–2010. (The sample sizes are comparable in the other periods 2008–2009 and 2010–2011.)

For the main part of the empirical analysis, we focus on the subset of patients in the baseline sample who exceed their deductible level during the first year of the two-year period. These individuals are henceforth referred to as “crossers”, whereas those who remain below the deductible are referred to as “non-crossers”. In accordance with the theoretical discussion above, we restrict the sample of crossers in two ways. First, we exclude patients who hit the deductible late in the year (October to December), which applies to 8% of the sample. This ensures that a comparison of health care expenditure before and after the turn of the year is not confounded by shocks that occur at the end of the year. Otherwise, declining costs around the change of year due to a gradual recovery from a health shocks would be mistakenly attributed to the effect of the price change. Appendix A presents some suggestive evidence that supports this notion. Second, to make sure that patients who experience a permanent deterioration in their latent health status do not confound the analysis, we exclude patients who alter their deductible level in the second year (4% of the sample). This procedure is consistent with our theoretical framework because deductible choice is informative on health status (i.e. the probability of health care shocks).

⁸Note that all treatment spells *must* be terminated on December 31st at the latest because the deductible period ends on that day. Therefore, there are no treatment spells in the data that span over two calendar years.

4.2 Descriptive Statistics

For the analysis, we split the sample in three groups according to the deductible level and henceforth define these groups as follows: high (at least CHF 1500), medium (CHF 500 and 1000) and low (CHF 300).⁹ Table 1 summarizes de-

Table 1: Descriptive Statistics: Whole Sample (Year 2009)

deductible group	high		medium		low	
	mean	st.dev.	mean	st.dev.	mean	st.dev.
annual medical costs (CHF)	1077	(3973)	4677	(9312)	5395	(10147)
annual # visits	3.62	(8.04)	11.39	(15.73)	13.71	(17.26)
age in years	41.95	(15.78)	57.76	(19.01)	53.76	(21.52)
female	0.41		0.56		0.60	
share of crossers	0.14		0.68		0.83	
# observations	85,838		149,576		124,113	

Notes: The sample is the entire adult population (18 and above) who is enrolled for mandatory health insurance in the period 2009–2010.

scriptive statistics from our baseline sample in 2009 across the three groups. As explained above, this sample includes adults who are enrolled in the period 2009–2010 and contains thus both crossers and non-crossers. It can be clearly seen that average health care costs and utilization is higher, the lower the deductible level. This confirms our expectations and is consistent with the notion that deductible choice depends on health status.¹⁰ For instance, annual medical costs in the medium-deductible and low-deductible groups are about 5 to 6 times higher than in the high-deductible group. Correspondingly, the share of individuals hitting their deductible level is very low in the latter group compared to the two former groups. In addition, we observe that younger people as well as males are markedly more likely to choose a high deductible level.

Since our empirical analysis focuses only on individuals who hit their deductible during the calendar year, we now present descriptive statistics only for this subsample (the crossers) in Table 2.¹¹ We note an interesting difference to

⁹In the high-deductible group, 58% have the CHF 1500 deductible, 12% have the CHF 2000 deductible and 30% have the CHF 2500 deductible. In the medium-deductible group, the deductible level CHF 500 is predominant with a share of 84%.

¹⁰Several empirical studies for the Swiss case have shown that roughly 75% of the health care expenditure differences across deductible groups can be explained by selection effects (Gerfin and Schellhorn, 2006; Trottmann et al., 2012).

¹¹Note that we draw random subsamples of size 30,000 from the medium-deductible and low-deductible crossers, respectively (*before* applying the sample selection criteria). The reason is that we have to construct day-level time series panels in the empirical analysis below. An analysis of the full samples would result in prohibitively large datasets.

Table 2: Descriptive Statistics: Only Crossers (Year 2009)

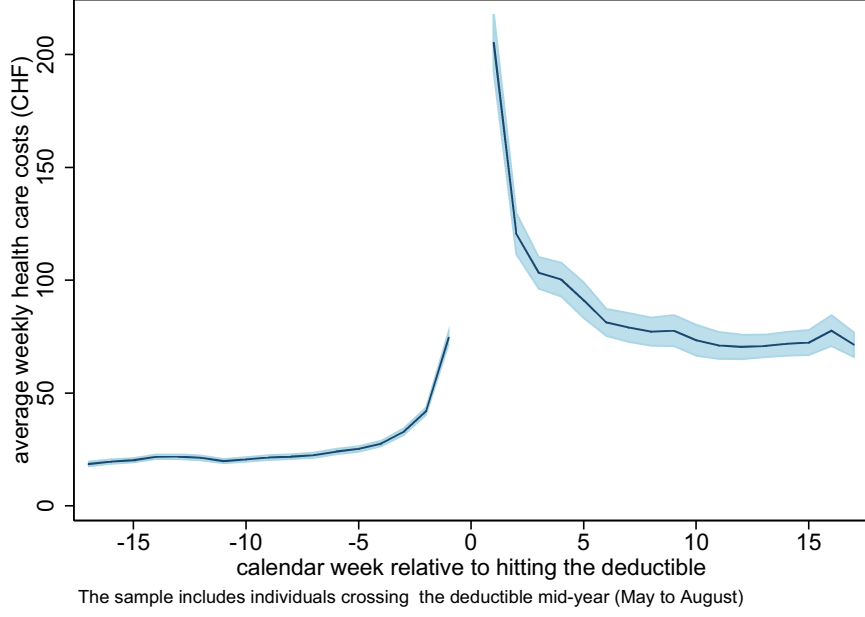
deductible group	high		medium		low	
	mean	st.dev.	mean	st.dev.	mean	st.dev.
annual medical costs (CHF)	7196	(9416)	7364	(11365)	6519	(11131)
annual # visits	18.37	(15.91)	17.03	(17.31)	16.72	(17.79)
age	53.13	(17.89)	62.53	(17.99)	53.86	(20.19)
female	0.49		0.60		0.62	
calendar day of hitting deductible	141.47	(75.14)	89.89	(70.06)	74.58	(65.72)
# observations	6,976		24,989		26,425	

Notes: This table reports statistics for the subsample of crossers in 2009, i.e. those who hit the deductible. In addition, those who change the deductible level and those who exceed the deductible late in the year (from October to December) are excluded.

the (unconditional) baseline sample in Table 1: among those who exceed their deductible, annual expenditure and visits are similar across groups. This suggests that conditional on hitting the deductible, the average severity of health shocks does not depend much on the chosen deductible. Another interesting (though not very surprising) finding for the high-deductible group is that crossers are much older compared to the unconditional sample. The crossers are therefore not a random sample from the baseline sample.

To illustrate the dynamics of health care costs in the context of crossing the deductible level, we compute average health care expenditure relative to the point in time when the deductible level was hit. Figure 2 shows average costs per calendar week relative to the week in which the deductible is hit in the year 2009. The sample consists of crossers who exceed the deductible mid-year (May to August) such that we can analyze the before-and-after dynamics *within* a calendar year. As we can see, costs remain on a low level until a few weeks prior to hitting the deductible level and rise sharply one week before the deductible level is reached. After hitting the deductible, costs begin to decline rapidly. This is evidence that the health shocks are, at least on average, transitory in nature. Note also that costs remain on a higher level than before the shock. Overall, the average dynamics observed here are consistent with the implied demand functions in equation (2): First, costs are low when the price is high and no health shock occurs. Second, when a shock occurs, costs jump up sharply. And third, costs decline quickly after the shock, but remain on a higher level than before the shock, which is consistent with a higher level of health care consumption in the face of a lower price.

Figure 2: Average Costs Before and After Exceeding the Deductible (Year 2009)



4.3 Evolution of Spot Prices and Health Care Costs

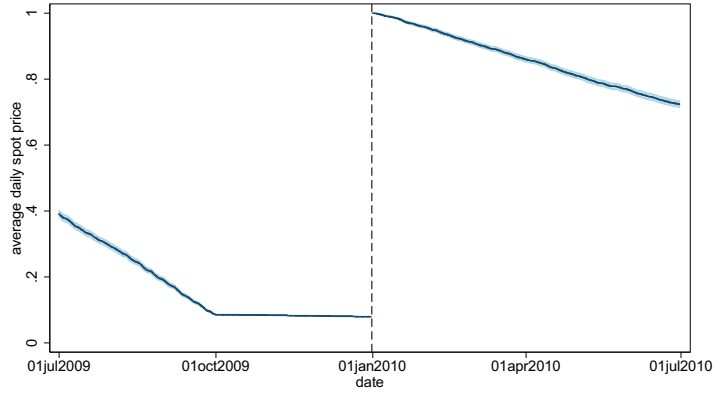
When moving to the next calendar year, crossers experience a sharp and discontinuous rise in the spot price of health care. The price either jumps from the co-payment rate (10%) or the stop-loss rate (0%) back to the full out-of-pocket price (100%). Note that we calculate the daily spot price for each individual based on day-level consumption of medical care. Due to the discrete-time setting, this variable is defined as either the realized price of medical care (if costs are positive) or the hypothetical price (if costs are zero) on each day.¹²

Figure 3 plots the evolution of average daily spot prices of medical care before and after the change of year. In general, mean daily prices are monotonically

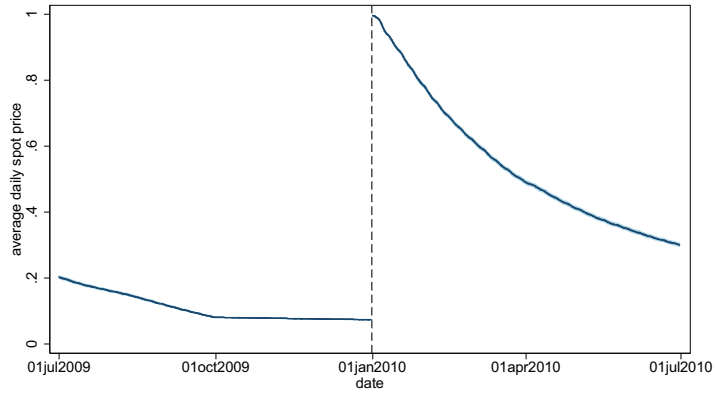
¹²For each individual i and each day t , we calculate the spot price as the *realized* price of health care consumption in case of positive consumption on this day and as the *marginal* price in case of no consumption on this day. Denote d_i the deductible, sl_i the stop-loss, y_{it} the health care expenditure on day t , and y_i^t the *cumulated* health care expenditures up to day t . The spot price is then defined as follows:

$$p_{it} = \begin{cases} 1 & \text{if } y_i^t < d_i \\ 1 \cdot \frac{d_i - y_i^{t-1}}{y_{it}} + 0.1 \cdot \frac{y_i^t - d_i}{y_{it}} & \text{if } y_i^{t-1} < d_i \text{ and } y_i^t > d_i \\ 0.1 & \text{if } y_i^{t-1} \geq d_i \text{ and } y_i^t < sl_i \\ 0.1 \cdot \frac{sl_i - y_i^{t-1}}{y_{it}} + 0 \cdot \frac{y_i^t - sl_i}{y_{it}} & \text{if } y_i^{t-1} < sl_i \text{ and } y_i^t > sl_i \\ 0 & \text{if } y_i^{t-1} \geq sl_i \end{cases}$$

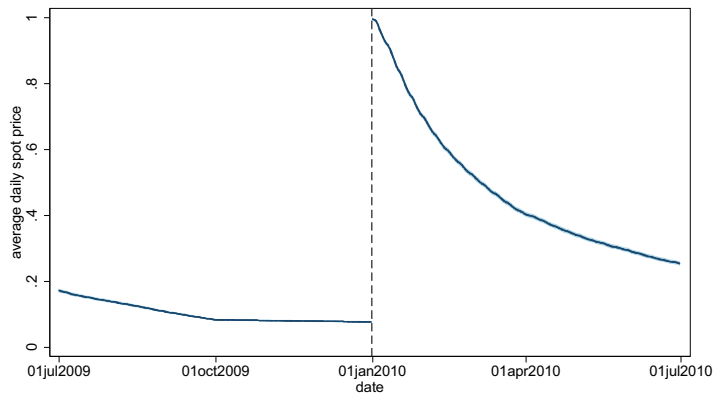
Figure 3: Evolution of Average Daily Spot Prices



(a) high deductible



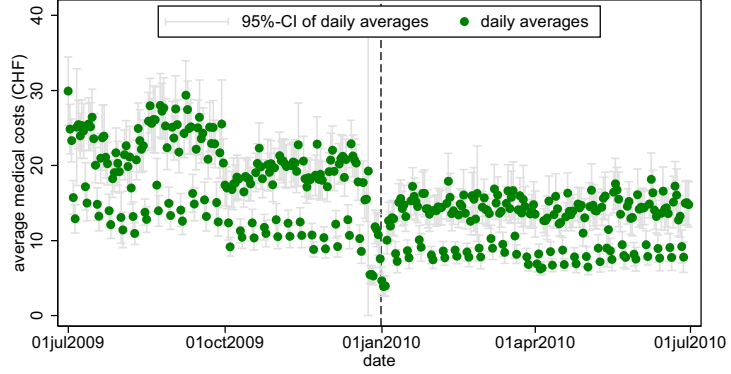
(b) medium deductible



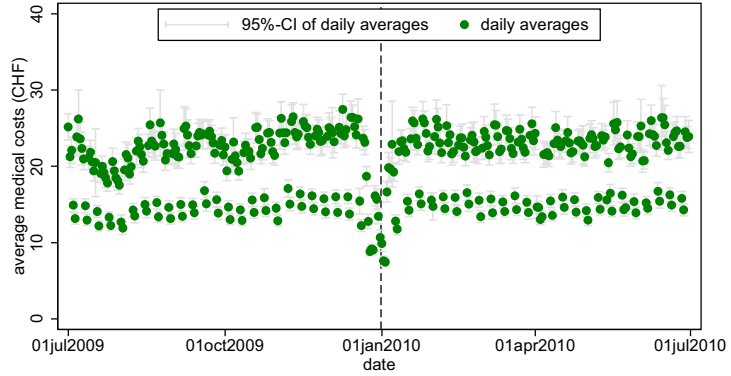
(c) low deductible

Notes: This figure shows average daily spot prices across deductible groups. The sample includes crossers who hit the deductible before October 2009 and keep the same deductible level in the two years.

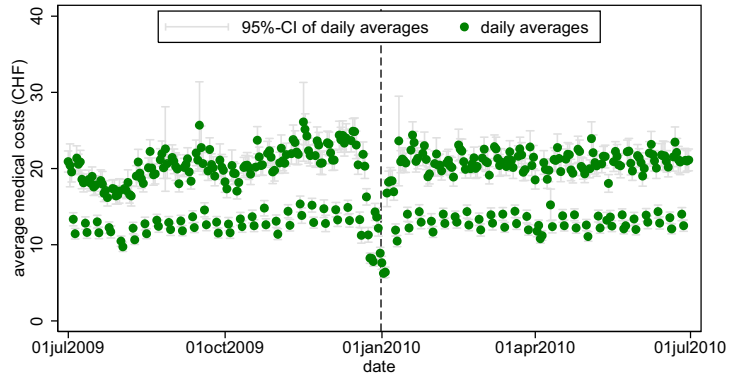
Figure 4: Evolution of Average Health Care Costs



(a) high deductible



(b) medium deductible



(c) low deductible

Notes: This figure shows average daily health care costs across deductible groups. The sample includes crossers who hit the deductible before October 2009 and keep the same deductible level in the two years.

decreasing *during* the year because on each day some additional individuals experience health shocks that push their health costs past the deductible level and consequently lower the average spot price of medical care. Note that, in all graphs, the kink on October 1st is *by design* because we have excluded all patients who hit the deductible from October to December (see discussion above). As the calendar year changes, the spot price jumps back to unity because the deductible resets for everyone. Note that the magnitude of the jump in the spot price differs very little across deductible groups. However, the dynamics during the year differ significantly; the average spot price in the medium-deductible and low-deductible groups decline at a much faster rate than in the high-deductible group. First, this is because lower deductibles are hit more quickly, everything else being equal. The second reason is that people with high deductibles experience fewer (or less severe) health shocks that push them over the deductible level.

Figure 4 plots average daily costs across time for each of the three deductible groups. The graphs reveal some important features of our time-series panel data. First, the daily averages on a markedly lower level are Saturdays and Sundays and reflect the fact that health care utilization on weekends is lower due to lower availability relative to weekdays. Hence, there is a large short-term variation in the data. Second, the data before and after the change of year exhibits a funnel-shaped pattern with a cluster of extremely low daily health care costs. This effect is due to the Christmas and New Year break during which accessibility and availability of non-emergency health care services is limited. Third, there are some seasonality patterns common to all deductible groups with troughs around August 1st and October 1st. This is most likely due to vacation periods. (Note that in the high-deductible group, the level shift around October 1st is again, at least partly, by design due to the exclusion of people who hit the deductible from October to December.) In the first half-year of 2010, by contrast, average daily costs do generally not exhibit such seasonality patterns and thus fluctuate much less. Finally, comparing costs before and after the change of year, it appears that average costs drop to a lower level in the high-deductible group, while this effect is much less pronounced or absent in the other groups. This can be taken as *prima facie* evidence that patients with high deductibles respond more strongly to the changes in the price of health care.

4.4 Empirical Framework

The aim of the empirical analysis is to measure the response of individuals' health care consumption to changes in the price. The setting we exploit is a type of sharp regression discontinuity design (SRD) often used in the treatment effects literature. However, there is an important difference to the standard SRD: In our setup, the running variable is *time* such that we observe all individuals at each value of the running variable. In the standard setup, an individual is observed at one value of the running variable. This generates several advantages. First, our data is a large time-series panel with many observations at each value of the running variable such that we can estimate local averages with reasonable precision. Second, causal effects are estimated for the entire sample under study and not only for those who are near the threshold. The drawback is that we have to deal with serial correlation and the fact that the lumpiness of health care expenditure generates a substantive amount of dispersion in combination with a large fraction of zeros. Another important difference is that we do not use the data that is very close to the threshold due to the effect of the Christmas and new-year break on health care utilization documented in Figure 4. Such a framework has been termed a “donut-RD” design by Barreca et al. (2011) who study the effect of heaping in the running variable on the estimation of treatment effects. Shigeoka (2013), for example, uses this approach to perform robustness checks.

To formalize the empirical framework, we begin by considering two subsequent years, with r indicating the change of year and T the number of subperiods (days) per year. Subperiods are of equal length, $t = 1, 2, \dots, 2T$, and subperiods with $t < r$ belong to the first year and subperiods with $t > r$ belong to the second year. The outcome variable Y_{it} is the health care consumption (as measured by health care costs) incurred by individual i on day t . We define this outcome by the function $Y_{it} = m(p_{it}, a_i, \varepsilon_{it})$, where p_{it} is the relevant price of health care, a_i is an unobserved time-constant effect, for example due to demographics or education, and ε_{it} is a random unobserved shock that varies over time.

In our context, the central question is how $m(p_{it}, a_i, \varepsilon_{it})$ responds to a change in p_{it} . Since health care utilization during the holiday season is strongly affected by the above described calendar effects (as illustrated in Figure 4), we decide to exclude this period in the before-and-after comparison of health care costs. Consequently, denote the (hypothetical) prices that we wish to study by p_{i,r_0} and p_{i,r_1} , respectively, where r_0 denotes the last day in the first year before the Christmas break and r_1 the first day in the new year after the Christmas break.

The period between r_0 and r_1 is excluded from the analysis.

As a result, the relevant individual-specific demand response to the change in prices is

$$\Delta_i = m(p_{i,r_1}, a_i, \varepsilon_{i,r_1}) - m(p_{i,r_0}, a_i, \varepsilon_{i,r_0}) \quad (3)$$

If we integrate over the population in a given subperiod, a_i is automatically integrated out and thus plays no role in estimating average demand responses. In other words, controlling for time-constant covariates is superfluous because the same individuals are compared before and after the change of year. However, we must assume that, for some given price p_i , changes in ε_{it} do not affect the average demand between subperiods r_0 and r_1 :

$$E[m(p_i, \varepsilon_{i,r_1}) - m(p_i, \varepsilon_{i,r_0})] = 0 \quad (4)$$

Under this assumption, the average demand response to the price change from p_{i,r_0} to p_{i,r_1} is simply identified by

$$E[\Delta_i] = E[Y_{i,r_1}] - E[Y_{i,r_0}] \quad (5)$$

To estimate these quantities in practice, it must be borne in mind that health care costs are characterized by a large degree of lumpiness and short-term (within-week) variation. Given a finite sample, simple one-day averages would result in very noisy estimates of (5). Instead, it appears more sensible to use smoothing techniques. We therefore use local linear regression (LLR), which has good bias properties relative to standard kernel estimation at boundary values of the covariate support (Hahn et al., 2001; Porter, 2003). Considering the samples of the two years of data separately, the local linear regression estimator for some value $t = x$ solves

$$\begin{aligned} \min_{\alpha_0, \beta_0} \sum_{i=1}^N \sum_{t=1}^{r_0} \frac{1}{h} K\left(\frac{t-x}{h}\right) (Y_{it} - \alpha_0(x) - (t-x)\beta_0(x))^2 \text{ for } t \leq r_0 \\ \min_{\alpha_1, \beta_1} \sum_{i=1}^N \sum_{t=r_1}^{2T} \frac{1}{h} K\left(\frac{t-x}{h}\right) (Y_{it} - \alpha_1(x) - (t-x)\beta_1(x))^2 \text{ for } t \geq r_1 \end{aligned} \quad (6)$$

where $K(\cdot)$ is a kernel function and h is a bandwidth to be specified. Given the above minimization problem, $\hat{\tau} = \hat{\alpha}_1(r_1) - \hat{\alpha}_0(r_0)$ is the nonparametric estimator of interest.

In conducting inference, it is crucial to take into account the potential serial correlation (state dependence) of patients' health care consumption patterns

across time. Therefore, we use a block bootstrap procedure to estimate standard errors.¹³ This allows for arbitrary serial correlation patterns of observations *within* panels (individuals).

4.5 Results

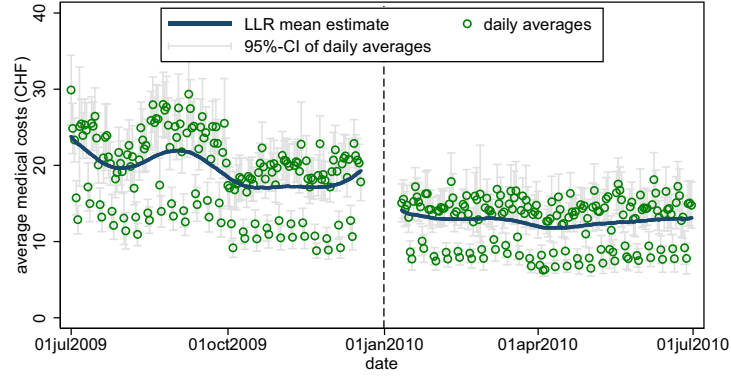
Figure 5 provides a graphical illustration of the evolution of average costs before and after the change of calendar year when average costs are smoothed using nonparametric estimation. As discussed above, we have excluded the time spell affected by the Christmas break, during which health care utilization is markedly different than in any other time of year. For this reason, r_0 is set to be the last Friday prior to Christmas and r_1 is the Monday of the second complete calendar week of January. We include 180 days of data before and after the change of year, respectively, such that the estimation window depicted is one year. The local linear regression is estimated with a triangular kernel.¹⁴ Daily average costs are included for comparison and can be interpreted as nonparametric estimates with $h \rightarrow 0$. We see that the LLR estimator is effective in “smoothing out” the short-term variation in the data, since the regression function is, for instance, not affected by the lower utilization at weekends. Panel (a) shows that, in the high-deductible group, the change of calendar year coincides with a discernable drop in health care costs. In contrast, the drop is less distinct in the groups with lower deductibles (Panels b and c). The figures for the two-year periods before (2008/2009) and after (2010/2011) are included in Appendix B and present a very similar picture.

In a next step, we produce numerical estimates of average prices and average costs of medical care before and after the change of year. In accordance with the previous discussion, we estimate both spot prices and end-of-year prices. To measure average spot prices on both sides of the threshold, we simply compute sample averages on the days defined by r_0 and r_1 . The end-of-year price on the left of the threshold (i.e. at the end of the first year) is equivalent to the spot price. The question is as to how to estimate the expected end-of-year price on the right of the threshold, i.e. at the beginning of the second year. Since we only require an average estimate, we take the *realized* end-of-year price as a proxy for

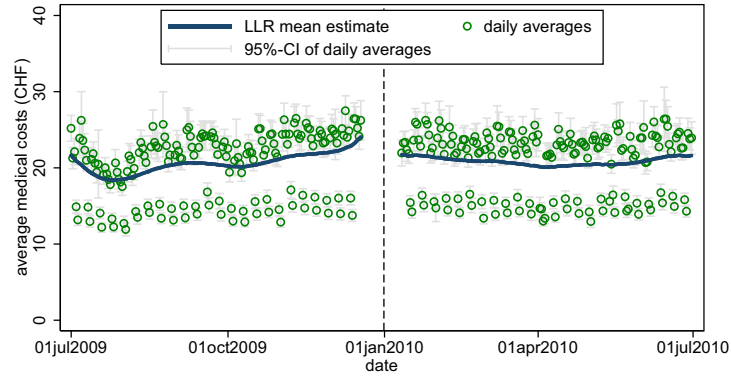
¹³Note that the serial correlation problem disappears asymptotically because $h \rightarrow 0$ as $N \rightarrow \infty$, but in a finite sample it appears important to correct for clustering.

¹⁴The bandwidth is set to $h = 30$ in *all* regressions. The same bandwidth is used to facilitate the comparison of results across samples. The value has been chosen on the basis of a series of optimal bandwidth calculations using the plug-in method, which suggested bandwidth values in a range around 30.

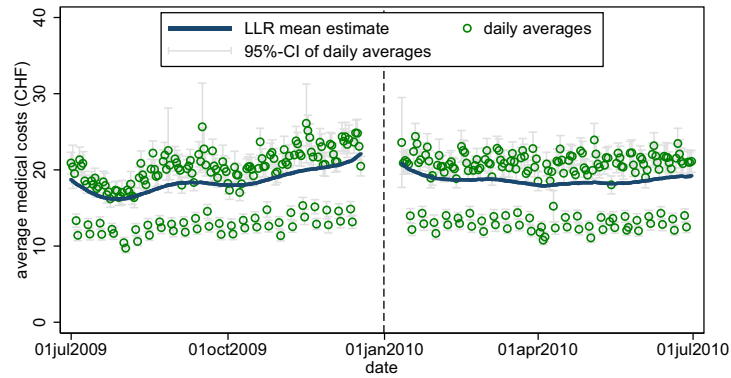
Figure 5: Average Health Care Costs At the Change of Year (2009/2010)



(a) high deductible



(b) medium deductible



(c) low deductible

Notes: This figure shows average costs estimated from daily means (green dots) and LLR (blue line) using a triangular kernel and bandwidth 30. The sample excludes the Christmas break.

the expected end-of-year price.¹⁵ To estimate average costs at the threshold, we use the LLR estimator specified in (6).

4.5.1 Change of Year Estimates

Table 3 presents numerical estimates of average costs and average prices of medical care on both sides of the threshold 2009/2010 in the sample of crossers. Standard errors are estimated with a nonparametric block bootstrap based on 1,000 replications. We see that the average spot price is around 0.08 in all three deductible groups at the end of the year which suggests that about 80% of the patients in each sample are in the co-payment region with a spot price of 0.1 and 20% are beyond the stop-loss with a spot price of zero. The average spot price is close to unity after the change of year such that the spot price changes by 0.8–0.9. This translates into a relative change of about 11 for each deductible group. Looking at the change in the end-of-year price, however, we find stark differences across deductible groups: In the high-deductible group, the change is 0.46 and thus much larger than in the lower deductible groups (0.08–0.10). Consequently, the relative change of the end-of-year price is about 5 times larger for the high deductible group (and about half the size of the relative change of the spot price).

¹⁵The realized end-of-year price is the spot price at the end of the year. Thus, the estimate is based on the average spot price on day r_1 in the second year.

Table 3: Estimation Results at Change of Year (2009/2010)

deductible: high	below (r_0)	above (r_1)	abs. change	rel. change
spot price	0.080 (0.000)	0.990 (0.001)	0.909** (0.001)	11.302** (0.074)
end-of-year price	0.080 (0.000)	0.541 (0.006)	0.461** (0.006)	5.725** (0.073)
total costs	19.278 (0.937)	14.095 (0.935)	-5.184** (1.193)	-0.269** (0.054)
outpatient costs	6.586 (0.335)	5.116 (0.313)	-1.470** (0.432)	-0.223** (0.059)
inpatient costs	6.187 (0.765)	3.671 (0.506)	-2.516** (0.883)	-0.407** (0.103)
drug costs	4.425 (0.270)	2.831 (0.234)	-1.594** (0.306)	-0.360** (0.056)
number of patients (N): 6,976				
deductible: medium	below (r_0)	above (r_1)	abs. change	rel. change
spot price	0.075 (0.000)	0.931 (0.001)	0.856** (0.001)	11.431** (0.047)
end-of-year price	0.075 (0.000)	0.175 (0.002)	0.100** (0.002)	1.336** (0.025)
total costs	24.046 (0.515)	21.723 (0.474)	-2.323** (0.637)	-0.097** (0.025)
outpatient costs	6.408 (0.168)	6.206 (0.150)	-0.202 (0.206)	-0.031 (0.032)
inpatient costs	7.339 (0.386)	7.555 (0.351)	0.215 (0.493)	0.029 (0.066)
drug costs	7.583 (0.269)	5.255 (0.177)	-2.328** (0.284)	-0.307** (0.029)
number of patients (N): 24,989				
deductible: low	below (r_0)	above (r_1)	abs. change	rel. change
spot price	0.078 (0.000)	0.906 (0.002)	0.829** (0.002)	10.660** (0.036)
end-of-year price	0.078 (0.000)	0.154 (0.002)	0.076** (0.002)	0.982** (0.020)
total costs	22.121 (0.529)	20.849 (0.728)	-1.272 (0.773)	-0.058 (0.034)
outpatient costs	6.100 (0.141)	6.154 (0.167)	0.054 (0.203)	0.009 (0.033)
inpatient costs	6.560 (0.385)	6.918 (0.769)	0.358 (0.815)	0.055 (0.124)
drug costs	6.498 (0.201)	4.961 (0.185)	-1.537** (0.246)	-0.237** (0.033)
number of patients (N): 26,425				

Notes: The table contains nonparametric estimates of average outcomes and prices before and after the change of year excluding the Christmas break. Average price estimates are based on the data from the corresponding day. Outcome estimates are based on LLR with a triangular kernel and $h = 30$. Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

Table 3 also presents estimates of total health care costs and its three most important sub-components: outpatient care, inpatient care and drugs. Recall that the outcomes are measured in daily costs per person. A number of interesting findings emerge. First, the absolute and relative response in demand to the change of year (as measured by total health care costs) is increasing in the deductible. In the high-deductible group, for instance, average health costs drop from 19 to 14 CHF, which corresponds to a decline of 27%. The economic significance of this number is considerable: it implies that high-deductible patients adjust their health care consumption quite strongly in response to the jump in the price. In comparison, the drop only amounts to 10% in the medium-deductible group and is insignificant in the low-deductible group. Second, drug costs decrease significantly in all deductible groups (they are the only driver of the significant negative effect for total costs in the medium deductible group). In relative terms, drug costs are reduced by 24% (low deductible group) to 36% (high deductible group). Third, outpatient and inpatient costs are reduced significantly in the high deductible group. Interestingly, this reduction is most pronounced for inpatient costs (it amounts to half of the reduction of total costs). We come back to this result below.

Tables A.1 and A.2 in Appendix B contain the same analysis for the episodes 2008/2009 and 2010/2011, respectively. The findings appear very similar and do not change any of the conclusions drawn from the sample of 2009/2010 that we have discussed above. The only notable exception is a significant reduction of total costs in the low deductible group in 2010/2011.

4.5.2 Elasticity Calculations

In the next step, we compute elasticities with respect to the two prices: the spot price and the end-of-year price. We use arc-elasticities, which are defined as

$$\eta = \frac{(m_1 - m_0)/(\frac{1}{2}[m_1 + m_0])}{(p_1 - p_0)/(\frac{1}{2}[p_1 + p_0])}, \quad (7)$$

rather than the ordinary elasticity formula, for two reasons. First, we have a setting with large discrete changes in the price in combination with a base point close to zero. Taken together, a percentage change in the price from, say, 0.08 to 1 equals 1150%, while a change in the opposite direction (from 1 to 0.08) would only equal 92%. Arc-elasticities evaluate the changes relative to the midpoint of the two points and are thus base independent. In the above example, the price change relative to the midpoint is 170% regardless of the direction of the

Table 4: Estimated Elasticities (2009/2010)

deductible:	high		medium		low	
	est.	st.err.	est.	st.err.	est.	st.err.
total costs						
implied arc-elasticity w.r.t. the						
spot price	-0.183**	(0.043)	-0.060**	(0.016)	-0.035	(0.022)
E-O-Y price	-0.210**	(0.049)	-0.127**	(0.034)	-0.090	(0.055)
outpatient costs						
implied arc-elasticity w.r.t. the						
spot price	-0.148**	(0.043)	-0.019	(0.019)	0.005	(0.020)
E-O-Y price	-0.170**	(0.050)	-0.040	(0.040)	0.013	(0.050)
inpatient costs						
implied arc-elasticity w.r.t. the						
spot price	-0.300**	(0.098)	0.017	(0.039)	0.032	(0.068)
E-O-Y price	-0.344**	(0.113)	0.036	(0.083)	0.081	(0.175)
drug costs						
implied arc-elasticity w.r.t. the						
spot price	-0.259**	(0.049)	-0.213**	(0.024)	-0.159**	(0.025)
E-O-Y price	-0.296**	(0.056)	-0.453**	(0.051)	-0.407**	(0.064)

Notes: Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

change. Second, arc-elasticities are widely used in other studies, most notably in the analysis of the RAND health insurance experiment (Keeler and Rolph, 1988). Using arc-elasticities thus facilitates the comparison of our results with the RAND benchmark.¹⁶

Generally, the interpretation of the following elasticity estimates is not straightforward because the observed demand response can occur for two reasons. First, if the price falls to a lower level due to hitting the deductible, the individual consumes more health care. This behavior relates to the traditional concept of the price elasticity of demand which suggests that patients consume health care they would not have consumed if the price had not dropped. Second, the dynamic aspect of our setting introduces the possibility of intertemporal substitution of health care consumption. If the price is low this year and high next year, it may seem optimal for patients to increase present consumption relative to future consumption. In other words, the additional consumption this year would have been consumed next year if the price would not have changed. In our setting, both effects are likely to play a role and it does not appear to be possible to identify the two effects separately. Elasticity estimates for the baseline episode 2009/2010 are presented in Table 4. Results are produced for

¹⁶Aron-Dine et al. (2013) demonstrate that elasticity estimates can vary substantially depending on whether the traditional definition or the arc-elasticity are estimated.

both the spot price and the (realized) end-of-year price. If an individual is completely myopic, he will only consider the spot price as the basis for his health care consumption. On the other hand, if an individual is forward-looking, he will respond to the (expected) end-of-year price. It is, of course, likely that the typical patient behavior lies somewhere in-between these two polar cases. In this sense, the elasticity with respect to the spot price can be seen as a lower bound and the elasticity with respect to the end-of-year price as an upper bound to the relevant elasticity of health care demand.

We first discuss the results for total health care costs given in the top panel. For the high-deductible group, we estimate the elasticity of health care costs with respect to the two price concepts at about -0.2 . Interestingly, the elasticities with respect to the spot price and the end-of-year price are close to each other. This means that a high-deductible individual reduces his health care demand by about 20% if the relevant price doubles irrespective of whether he is myopic or forward-looking. The estimated elasticities are considerably smaller for the medium-deductible group, and not significantly different from zero for the low-deductible group. Hence, the responsiveness of health care demand to changes in prices is stronger, the higher the chosen deductible level and thus the better the health status. This might be interpreted as some evidence for (partially) forward-looking behavior. If individuals reacted to spot prices, we would expect to see similar responses across all deductible groups, because as demonstrated in Table 3, spot price changes are similar across deductible groups.

Looking at the individual components of health care costs, we find similar results for the demand of outpatient care as well as for inpatient care. Elasticities are only significantly different from zero for the high-deductible group. Of course, the elasticities mirror the estimated percentage changes discussed above. The rather large elasticity of inpatient care is somewhat surprising. Figure A.5 in the Appendix shows the estimation result for inpatient care graphically. For the high-deductible group, there is a large increase in inpatient costs in December. This may be an indication for at least partial intertemporal substitution. Because inpatient care is relatively cheap after hitting the deductible, some patients may be able to bring forward operations that would otherwise have taken place at a later point (i.e. operations that are not urgent and can be planned). This case is not moral hazard in the classic sense because the operations would have been performed in any case. However, patients are able to shift the costs of these operations to the insurance company and hence to society. Furthermore, this type of intertemporal substitution keeps the high deductible attractive in the

future.

The demand for prescription drugs is more elastic than the demand for total health care: the estimated elasticities are large and significantly different from zero across all deductible groups. This finding is surprising for the low deductible group given that for them the end-of-year price hardly changes. If individuals in the low deductible group were to react to the spot price instead, there should be demand responses in the other categories of health care as well. It is possible that patients with long-term prescriptions have these prescriptions filled as long as the price is low. However, it seems unlikely that this completely explains the result. The graphical analysis in Figure A.6 suggests some intertemporal substitution for the low and medium deductible group because there is a steep increase in prescription drug expenditures during December. On the other hand, there is no visible drop in January compared to later months in 2010. This is in contrast to the findings in Einav et al. (2013a), who document a shift in drug expenditure from December to January among those with cumulated drug expenditure close to the donut hole gap in Medicare Part D. A further analysis of this topic is left for future research.¹⁷

4.5.3 Effect Heterogeneity and Sensitivity Analysis

In this section, we provide additional results to explore effect heterogeneity in the patient population and to test the sensitivity of our results to the sample restriction criteria. First, we estimate the elasticities of total health care costs with respect to prices separately by age groups and gender. We define two age groups denoted young ($\text{age} < 55$ years) and old ($\text{age} \geq 55$ years). The results from the period 2009/2010 are presented in Table 5. It is interesting to observe that men respond more strongly than women. This heterogeneity across gender generates a significant response for men with low deductibles. This could be an indication that men are more prone to moral hazard than women or that they respond more strongly to the dynamic incentives. Comparing estimates between age groups, the evidence is more mixed; while there is no discernable difference in the high-deductible group, we find a significant response of younger patients, but a close-to-zero response of older patients.

¹⁷Informal evidence for intertemporal substitution is provided by an interview with the manager of a large pharmacy. His observation is that there is an above-average demand for prescription drugs in December even for long-term prescriptions that are valid beyond December 31. According to him, one reason for this is that patients think that the drugs are cheaper for them before the deductible resets. This corresponds to strong myopia in individuals' behavior.

Table 5: Estimated Elasticities by Age and Gender (2009/2010)

deductible:	high		medium		low	
	est.	st.err.	est.	st.err.	est.	st.err.
young (age < 55)						
implied arc-elasticity w.r.t. the						
spot price	-0.189*	(0.076)	-0.109**	(0.036)	-0.104**	(0.030)
E-O-Y price	-0.207*	(0.083)	-0.164**	(0.053)	-0.201**	(0.057)
old (age ≥ 55)						
implied arc-elasticity w.r.t. the						
spot price	-0.178**	(0.048)	-0.047**	(0.018)	0.003	(0.031)
E-O-Y price	-0.219**	(0.059)	-0.160**	(0.062)	0.017	(0.194)
men						
implied arc-elasticity w.r.t. the						
spot price	-0.225**	(0.068)	-0.071*	(0.031)	-0.105**	(0.031)
E-O-Y price	-0.254**	(0.077)	-0.137*	(0.059)	-0.229**	(0.068)
women						
implied arc-elasticity w.r.t. the						
spot price	-0.134**	(0.048)	-0.052**	(0.019)	0.005	(0.031)
E-O-Y price	-0.156**	(0.056)	-0.121**	(0.044)	0.015	(0.090)

Notes: Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$.

Second, we test whether our results are sensitive to the sample restriction of excluding individuals who alter their deductible level for the second year. For this purpose, we re-estimate the change in health care demand including these individuals. Table 6 displays the results based on the period 2009/2010. The sample of the low-deductible group is nearly the same as before which is indicative of the fact that they are unlikely to alter their deductible. By contrast, the sample of the high-deductible group is now considerably larger than before. Comparing the results between Table 6 and Table 3, we find no qualitative differences. In addition, the point estimates are remarkably similar with nearly all of the relative changes being within one percentage point of the original estimates. The only exception is the change in drug costs in the high-deductible group, which is 5 percentage points lower than before. However, the change in total costs in the high-deductible group is barely affected (26% compared to 27% before). Thus, our overall findings are not sensitive to the exclusion of patients who adjust their deductible level.

Finally, we estimate the model using aggregated monthly data.¹⁸ In order to estimate the individual monthly health care expenditures we did not exclude the weeks of the Christmas break. In general, the estimated effects are somewhat

¹⁸These results are available on request.

Table 6: Estimation Results Including Deductible Changers (2009/2010)

deductible: high	below (r_0)	above (r_1)	abs. change	rel. change
total costs	20.616 (0.821)	15.285 (0.838)	-5.332** (0.992)	-0.259** (0.043)
outpatient costs	6.869 (0.304)	5.294 (0.266)	-1.575** (0.388)	-0.229** (0.050)
inpatient costs	6.269 (0.728)	3.810 (0.479)	-2.459** (0.856)	-0.392** (0.104)
drug costs	5.056 (0.310)	3.477 (0.303)	-1.579** (0.347)	-0.312** (0.060)
number of patients (N): 8,554				
deductible: medium	below (r_0)	above (r_1)	abs. change	rel. change
total costs	23.978 (0.559)	21.813 (0.459)	-2.165** (0.632)	-0.090** (0.025)
outpatient costs	6.360 (0.193)	6.248 (0.163)	-0.112 (0.234)	-0.018 (0.036)
inpatient costs	7.390 (0.391)	7.704 (0.432)	0.314 (0.526)	0.042 (0.072)
drug costs	7.503 (0.272)	5.156 (0.186)	-2.347** (0.294)	-0.313** (0.031)
number of patients (N): 26,717				
deductible: low	below (r_0)	above (r_1)	abs. change	rel. change
total costs	21.838 (0.510)	20.542 (0.751)	-1.296 (0.793)	-0.059 (0.036)
outpatient costs	6.048 (0.148)	6.079 (0.149)	0.032 (0.197)	0.005 (0.032)
inpatient costs	6.450 (0.404)	6.802 (0.692)	0.352 (0.750)	0.055 (0.120)
drug costs	6.415 (0.215)	4.886 (0.164)	-1.529** (0.255)	-0.238** (0.033)
number of patients (N): 26,870				

Notes: Individuals altering the deductible level from 2009 to 2010 are included. Outcome estimates are based on LLR with a triangular kernel and $h = 30$. Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

smaller in absolute terms. One major difference to the above results is that in the high deductible group the effect on inpatient care becomes much weaker and statistically insignificant. One possible explanation for this finding is that part of the observed increase in daily expenditures on inpatient care in December is simply due to a shift of routine operations to early December. This effect cancels out if we look at monthly data. The same mechanism may be at play in the other cost categories as well, albeit to a lesser degree. The drawback of the monthly analysis is that it stretches the identification assumptions for our RD-type design to some extent.

5 Conclusions

This paper has analyzed health care demand in the presence of nonlinear pricing which in turn generates dynamic incentives. Nonlinear price schedules are the result of deductibles and a maximum out-of-pocket payment in the Swiss health insurance system. Our analysis exploits the institutional regulation that deductibles apply to a calendar year and reset to the initial level at the start of the new year. This generates a discrete jump in the price of medical care for people who have exceeded their deductible level.

We use a regression discontinuity type framework based on daily data to estimate the change in health care demand right before and right after the turn of the year. We find that for individuals with high deductibles total health care demand drops by 27%, which translates into an elasticity of -0.21 . The decrease is most pronounced for inpatient care and prescription drugs. By contrast, for individuals with low deductibles the estimated elasticities are almost zero, with the exception of prescription drugs. These results may be interpreted as evidence for forward-looking behavior, because the expected end-of-year price does change substantially in the high deductible group, but not in the low deductible group. However, the results for prescription drugs in the low deductible group do not support the forward-looking hypothesis.

Our results are informative about behavioral responses caused by price changes induced by the institutional setting. If individuals understand their health insurance plans, they know that the deductible will reset on January 1st and should thus be aware of the price change. If they anticipate a higher price in the next year, or if they are myopic and do not consider the development of the price during the year, they may bring forward medical care (routine check-ups, operations that are not urgent). For this reason, it is difficult to separate

the observed responses into a true moral hazard effect (health care that would not have been consumed if the price was higher) and an intertemporal substitution effect (health care that would have been consumed next year instead of this year). In this sense, our results are difficult to compare with the literature, especially the well-known RAND results. The separation of the moral hazard and the substitution effect is left to future research.

References

- 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 February 2012.
- , —, and —, “The RAND Health Insurance Experiment, Three Decades Later,” *Journal of Economic Perspectives*, 2013, 27 (1), 197–222.
- Barreca, Alan I., Jason M. Lindo, and Glen R. Waddell**, “Heaping-Induced Bias in Regression-Discontinuity Designs,” Working Paper 17408, National Bureau of Economic Research September 2011.
- Boes, Stefan and Michael Gerfin**, “Does full insurance increase the demand for health care?,” Discussion Paper 7689, Institute for the Study of Labor (IZA) 2013.
- Borenstein, Severin**, “To what electricity price do consumers respond? Residential demand elasticity under increasing-block pricing,” Working Paper, UC Berkeley 2009.
- Chetty, Raj, John N Friedman, Tore Olsen, and Luigi Pistaferri**, “Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records,” *The quarterly journal of economics*, 2011, 126 (2), 749–804.
- Einav, Liran, Amy Finkelstein, and Paul Schrimpf**, “The Response of Drug Expenditures to Non-Linear Contract Design: Evidence from Medicare Part D,” Working Paper 19393, National Bureau of Economic Research August 2013.
- , —, and **Stephen P Ryan**, “Selection on Moral Hazard in Health Insurance,” *American Economic Review*, 2013, 103 (1), 178–219.
- Gerfin, Michael and Martin Schellhorn**, “Nonparametric bounds on the effect of deductibles in health care insurance on doctor visits—Swiss evidence,” *Health Economics*, 2006, 15 (9), 1011–1020.
- Hahn, Jinyong, Petra Todd, and Wilbert Van der Klaauw**, “Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design,” *Econometrica*, 2001, 69 (1), pp. 201–209.
- Hausman, Jerry A**, “The econometrics of nonlinear budget sets,” *Econometrica: Journal of the Econometric Society*, 1985, pp. 1255–1282.
- Ito, Koichiro**, “How Do Consumers Respond to Nonlinear Pricing? Evidence from Household Water Demand,” Working Paper, Stanford University 2013.

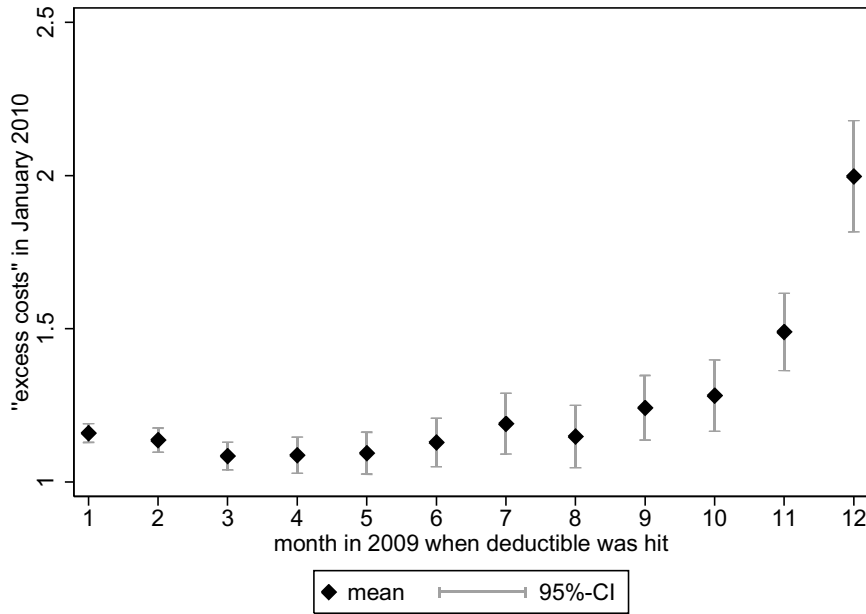
- , “Do Consumers Respond to Marginal or Average Price? Evidence from Nonlinear Electricity Pricing,” *American Economic Review*, forthcoming.
- Keeler, Emmett B and John E Rolph**, “The demand for episodes of treatment in the health insurance experiment,” *Journal of Health Economics*, 1988, 7 (4), 337–367.
- Kowalski, Amanda E.**, “Estimating the Tradeoff Between Risk Protection and Moral Hazard with a Nonlinear Budget Set Model of Health Insurance,” Working Paper 18108, National Bureau of Economic Research May 2012.
- Marsh, Christina**, “Estimating Demand Elasticities Using Nonlinear Pricing,” Working Paper, University of Georgia 2013.
- Porter, Jack**, “Estimation in the regression discontinuity model,” 2003. Working Paper, University of Wisconsin at Madison.
- Saez, Emmanuel**, “Do Taxpayers Bunch at Kink Points?,” *American Economic Journal: Economic Policy*, 2010, 2 (3), 180–212.
- Shigeoka, Hitoshi**, “The Effect of Patient Cost Sharing on Utilization, Health, and Risk Protection,” Working Paper 19726, National Bureau of Economic Research December 2013.
- 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.

Appendix

A Excess Costs in January

In this section, we investigate whether individuals who cross their deductible late in the year are still affected by health shocks in the beginning of the subsequent year. To do this, we use a similar strategy as Einav et al. (2013a). We consider all crossers from 2009 and compute the ratio of their health care costs in January 2010 relative to the monthly average from February to December 2010. This ratio is informative on the health care consumption in January relative to the rest of the year. Taking the ratio thus “controls” for the heterogeneity in the level of health care expenditures across individuals. Figure A.1 shows the average ratio as a function of the calendar month in 2009 in which patients hit their deductible level (“crossing month”). On average, costs in January are higher relative to the

Figure A.1: Excess Costs in January Relative to Rest of Year



rest of the year which is due to more intensive health care utilization in the winter. Indeed, it is very suggestive to compare these “excess January costs” across the crossing month in 2009. We see clearly that patients who hit their deductible in December (and November) 2009 have very high costs in January 2010 *relative* to the rest of the year 2010. In contrast, the ratio is not significantly different between crossing months from January to October 2009. This provides

evidence that health shocks occurring at the end of year have lasting consequences for health care expenditures at the beginning of the subsequent year. For this reason, it is sensible to exclude from the analysis those who cross their deductible late in the year because their consumption in January still appears to be affected by health shocks.

B Additional Estimation Results (2008/2009 and 2010/2011)

Table A.1: Estimation Results at Change of Year (2008/2009)

deductible: high	below (r_0)	above (r_1)	abs. change	rel. change
spo55t price	0.080 (0.000)	0.988 (0.001)	0.908** (0.001)	11.314** (0.072)
end-of-year price	0.080 (0.000)	0.518 (0.006)	0.438** (0.006)	5.457** (0.069)
total costs	18.706 (0.804)	15.074 (0.749)	-3.632** (0.950)	-0.194** (0.046)
outpatient costs	7.044 (0.315)	5.275 (0.297)	-1.769** (0.364)	-0.251** (0.046)
inpatient costs	4.431 (0.515)	4.523 (0.652)	0.093 (0.853)	0.021 (0.206)
drug costs	4.811 (0.261)	3.087 (0.177)	-1.724** (0.268)	-0.358** (0.042)
number of patients (N): 6,683				
deductible: medium	below (r_0)	above (r_1)	abs. change	rel. change
spot price	0.074 (0.000)	0.922 (0.001)	0.848** (0.001)	11.442** (0.041)
end-of-year price	0.074 (0.000)	0.153 (0.002)	0.079** (0.002)	1.069** (0.021)
total costs	25.001 (0.515)	23.733 (0.577)	-1.268 (0.730)	-0.051 (0.029)
outpatient costs	6.368 (0.176)	7.074 (0.396)	0.706 (0.439)	0.111 (0.071)
inpatient costs	7.955 (0.407)	8.439 (0.450)	0.484 (0.584)	0.061 (0.074)
drug costs	7.918 (0.295)	5.547 (0.132)	-2.371** (0.302)	-0.299** (0.028)
number of patients (N): 26,294				
deductible: low	below (r_0)	above (r_1)	abs. change	rel. change
spot price	0.075 (0.000)	0.888 (0.002)	0.813** (0.002)	10.844** (0.039)
end-of-year price	0.075 (0.000)	0.132 (0.001)	0.057** (0.001)	0.762** (0.018)

continued on next page

Table A.1 – *Continued from previous page*

total costs	22.889 (0.475)	22.055 (0.433)	−0.834 (0.607)	−0.036 (0.026)
outpatient costs	6.285 (0.179)	6.717 (0.173)	0.432 (0.233)	0.069 (0.038)
inpatient costs	6.960 (0.346)	7.122 (0.354)	0.161 (0.449)	0.023 (0.064)
drug costs	7.056 (0.224)	5.432 (0.174)	−1.625** (0.265)	−0.230** (0.032)
number of patients (N): 27,480				

Notes: The table contains nonparametric estimates of average outcomes and prices before and after the change of year excluding the Christmas break. Average price estimates are based on the data from the corresponding day. Outcome estimates are based on LLR with a triangular kernel and $h = 30$. Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

Table A.2: Estimation Results at Change of Year (2010/2011)

	deductible: high	below (r_0)	above (r_1)	abs. change	rel. change
spot price		0.080 (0.001)	0.989 (0.001)	0.909** (0.001)	11.362** (0.077)
end-of-year price		0.080 (0.001)	0.515 (0.006)	0.435** (0.006)	5.439** (0.072)
total costs		20.063 (1.115)	14.365 (0.761)	−5.699** (1.193)	−0.284** (0.048)
outpatient costs		6.864 (0.341)	5.605 (0.307)	−1.260** (0.400)	−0.184** (0.052)
inpatient costs		6.308 (0.950)	3.572 (0.593)	−2.736** (1.042)	−0.434** (0.123)
drug costs		4.656 (0.275)	2.958 (0.224)	−1.698** (0.331)	−0.365** (0.059)
number of patients (N): 6,881					

	deductible: medium	below (r_0)	above (r_1)	abs. change	rel. change
spot price		0.074 (0.000)	0.934 (0.001)	0.860** (0.001)	11.657** (0.043)
end-of-year price		0.074 (0.000)	0.161 (0.002)	0.087** (0.002)	1.182** (0.022)
total costs		25.862 (0.656)	23.131 (0.613)	−2.732** (0.792)	−0.106** (0.029)
outpatient costs		7.067 (0.188)	6.819 (0.223)	−0.249 (0.274)	−0.035 (0.039)
inpatient costs		8.629 (0.571)	8.053 (0.417)	−0.575 (0.693)	−0.067 (0.077)
drug costs		7.020 (0.234)	5.274 (0.214)	−1.746** (0.293)	−0.249** (0.036)

number of patients (N): 25,427

	deductible: low	below (r_0)	above (r_1)	abs. change	rel. change
--	-----------------	-----------------	-----------------	-------------	-------------

continued on next page

Table A.2 – Continued from previous page

spot price	0.077 (0.000)	0.912 (0.002)	0.835** (0.001)	10.861** (0.039)
end-of-year price	0.077 (0.000)	0.152 (0.002)	0.075** (0.002)	0.979** (0.020)
total costs	22.529 (0.499)	20.248 (0.450)	−2.281** (0.575)	−0.101** (0.024)
outpatient costs	6.634 (0.209)	6.370 (0.179)	−0.264 (0.235)	−0.040 (0.034)
inpatient costs	6.812 (0.359)	6.446 (0.364)	−0.366 (0.474)	−0.054 (0.070)
drug costs	6.295 (0.199)	4.779 (0.219)	−1.516** (0.280)	−0.241** (0.039)

number of patients (N): 26,522

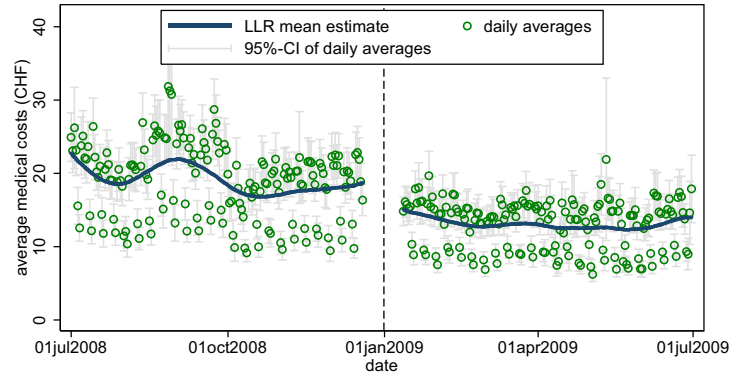
Notes: The table contains nonparametric estimates of average outcomes and prices before and after the change of year excluding the Christmas break. Average price estimates are based on the data from the corresponding day. Outcome estimates are based on LLR with a triangular kernel and $h = 30$. Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

Table A.3: Estimated Elasticities (2008/2009)

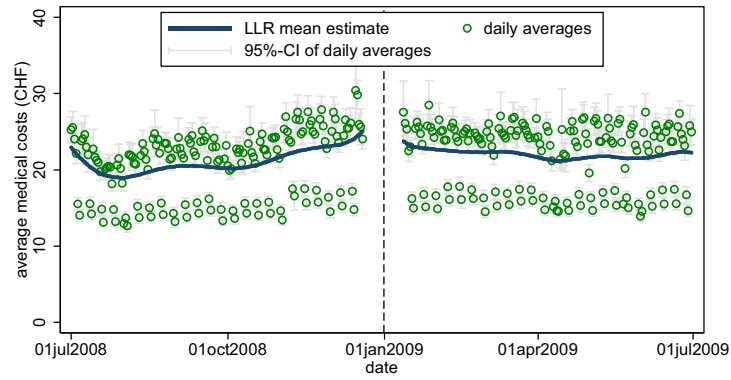
deductible:	high		medium		low	
	est.	st.err.	est.	st.err.	est.	st.err.
total costs						
implied arc-elasticity w.r.t. the						
spot price	−0.127**	(0.033)	−0.031	(0.018)	−0.022	(0.016)
E-O-Y price	−0.147**	(0.039)	−0.075	(0.043)	−0.067	(0.049)
outpatient costs						
implied arc-elasticity w.r.t. the						
spot price	−0.169**	(0.035)	0.062	(0.037)	0.039	(0.021)
E-O-Y price	−0.196**	(0.041)	0.151	(0.089)	0.120	(0.065)
inpatient costs						
implied arc-elasticity w.r.t. the						
spot price	0.012	(0.114)	0.035	(0.041)	0.014	(0.038)
E-O-Y price	0.014	(0.133)	0.085	(0.101)	0.042	(0.116)
drug costs						
implied arc-elasticity w.r.t. the						
spot price	−0.257**	(0.037)	−0.207**	(0.023)	−0.154**	(0.024)
E-O-Y price	−0.298**	(0.043)	−0.505**	(0.058)	−0.471**	(0.074)

Notes: Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

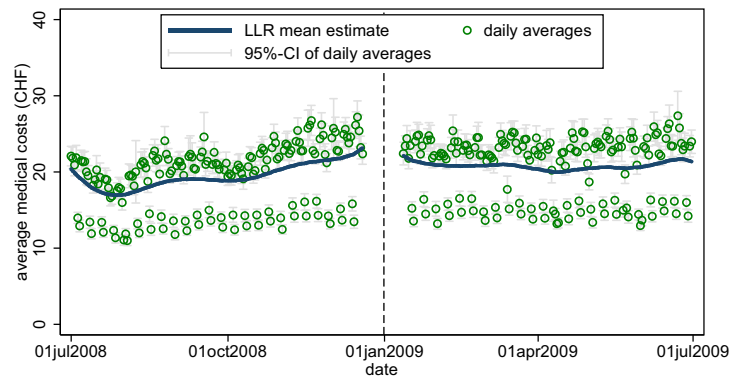
Figure A.2: Average Health Care Costs At the Change of Year (2008/2009)



(a) high deductible

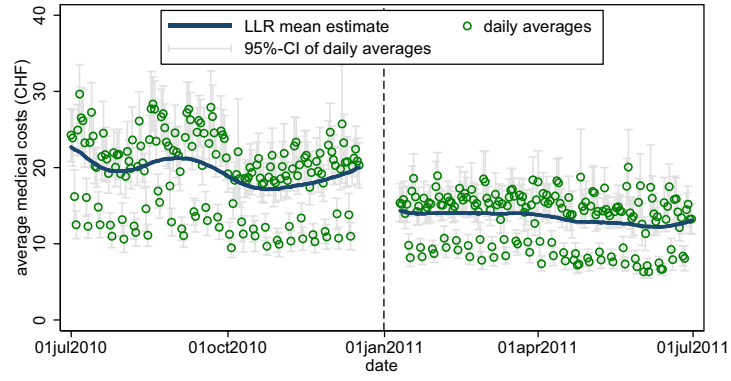


(b) medium deductible

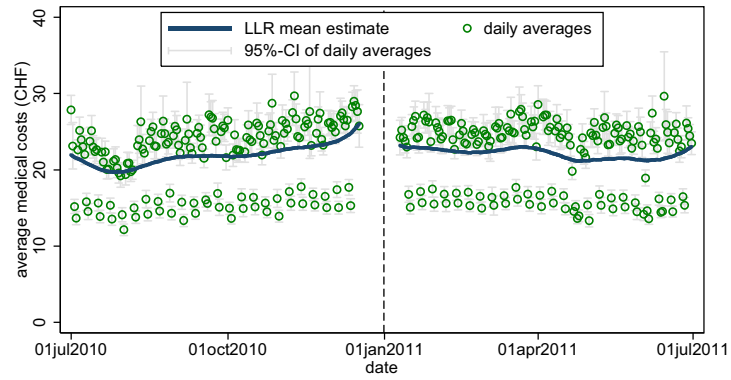


(c) low deductible

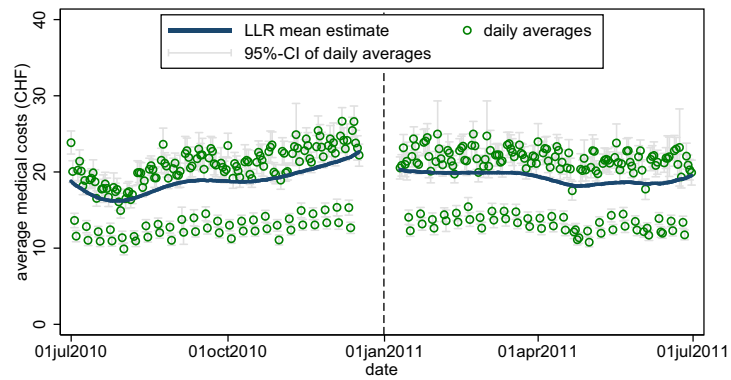
Figure A.3: Average Health Care Costs At the Change of Year (2010/2011)



(a) high deductible



(b) medium deductible



(c) low deductible

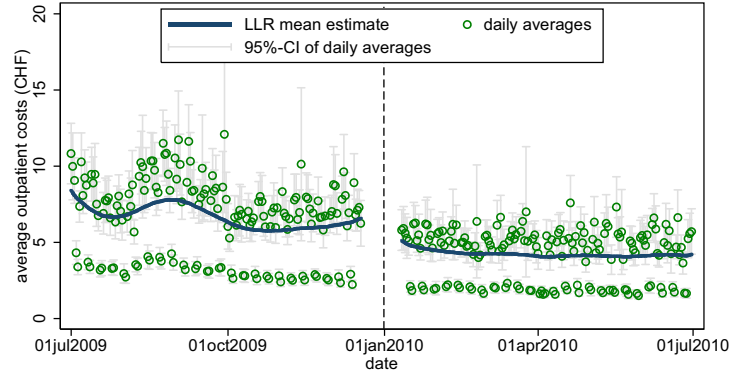
Table A.4: Estimated Elasticities (2010/2011)

deductible:	high		medium		low	
	est.	st.err.	est.	st.err.	est.	st.err.
total costs						
implied arc-elasticity w.r.t. the						
spot price	-0.195**	(0.038)	-0.065**	(0.019)	-0.063**	(0.016)
E-O-Y price	-0.226**	(0.045)	-0.150**	(0.044)	-0.162**	(0.041)
outpatient costs						
implied arc-elasticity w.r.t. the						
spot price	-0.119**	(0.037)	-0.021	(0.023)	-0.024	(0.021)
E-O-Y price	-0.138**	(0.043)	-0.048	(0.054)	-0.062	(0.054)
inpatient costs						
implied arc-elasticity w.r.t. the						
spot price	-0.326**	(0.112)	-0.040	(0.048)	-0.033	(0.042)
E-O-Y price	-0.379**	(0.130)	-0.094	(0.112)	-0.084	(0.109)
drug costs						
implied arc-elasticity w.r.t. the						
spot price	-0.262**	(0.051)	-0.166**	(0.028)	-0.162**	(0.031)
E-O-Y price	-0.305**	(0.059)	-0.382**	(0.064)	-0.416**	(0.079)

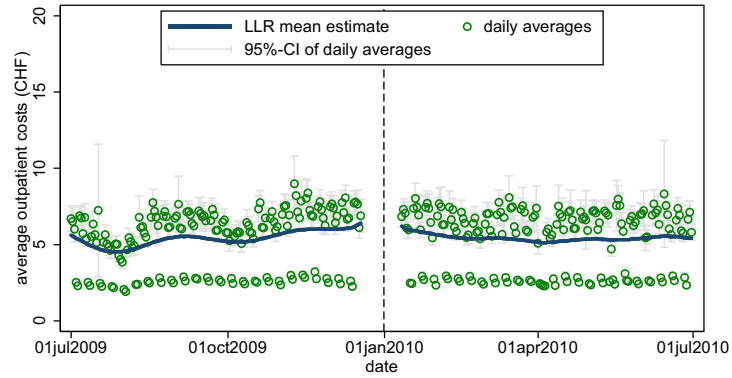
Notes: Cluster-robust standard errors in parentheses are block bootstrapped using 1,000 replications. Significance levels: ** $p < 0.01$, * $p < 0.05$

C Estimation for Cost Categories

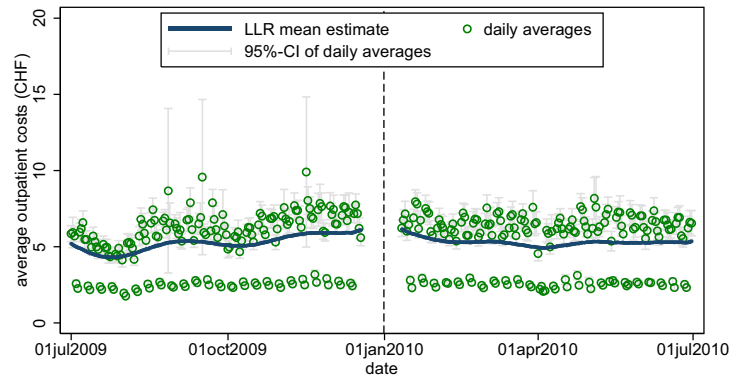
Figure A.4: Average Outpatient Costs At the Change of Year (2009/2010)



(a) high deductible



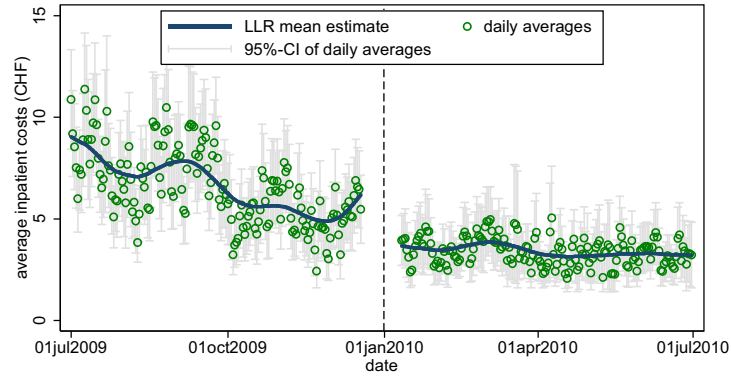
(b) medium deductible



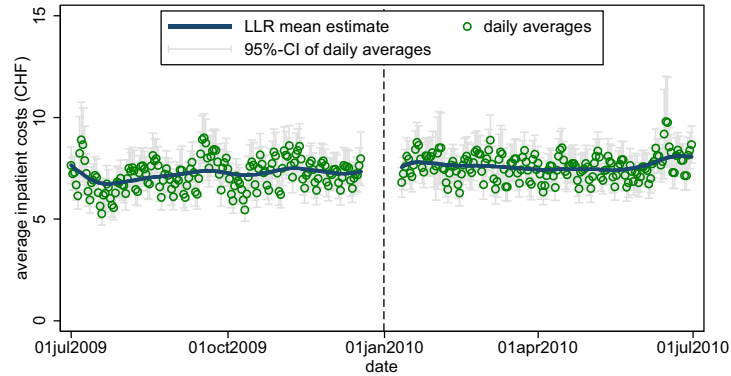
(c) low deductible

Notes: This figure shows average outpatient costs estimated from daily means (green dots) and LLR (blue line) using a triangular kernel and bandwidth 30. The sample excludes the Christmas break.

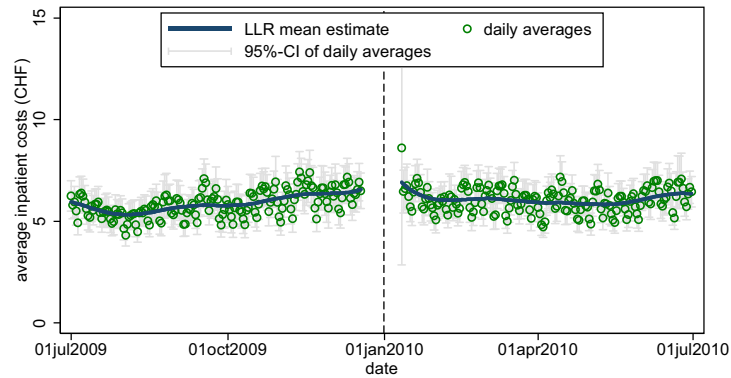
Figure A.5: Average Inpatient Costs At the Change of Year (2009/2010)



(a) high deductible



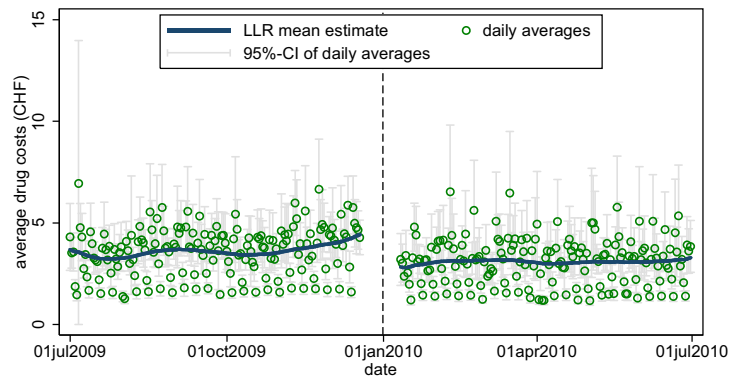
(b) medium deductible



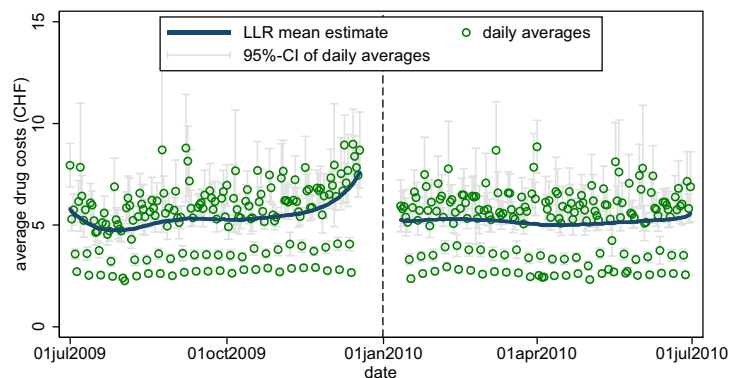
(c) low deductible

Notes: This figure shows average inpatient costs estimated from daily means (green dots) and LLR (blue line) using a triangular kernel and bandwidth 30. The sample excludes the Christmas break.

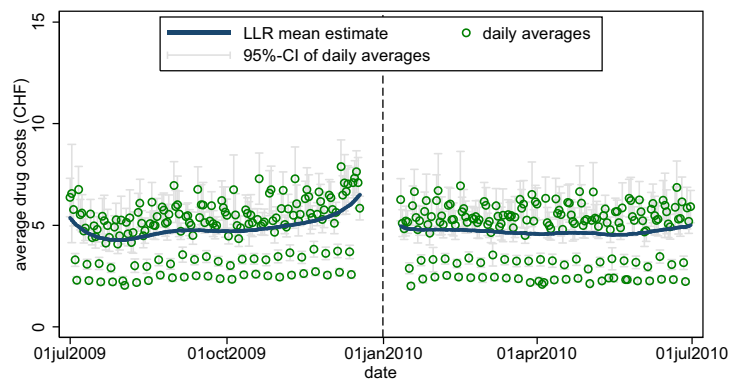
Figure A.6: Average Drug Costs At the Change of Year (2009/2010)



(a) high deductible



(b) medium deductible



(c) low deductible

Notes: This figure shows average drug costs estimated from daily means (green dots) and LLR (blue line) using a triangular kernel and bandwidth 30. The sample excludes the Christmas break.