CINCH competence in competition + health series

#2019/04

Arthur P. Hayen, Tobias J. Klein, and Martin Salm

Does the framing of patient cost-sharing incentives matter?

The effects of deductibles vs. no-claim refunds

Imprint

CINCH SERIES

CINCH – Health Economics Research Center
Weststadttürme, Berliner Platz 6-8
45127 Essen

www.cinch.uni-due.de
cinchseries@cinch-essen.de
Phone +49 (0) 201 183 - 3679
Fax +49 (0) 201 183 - 3716

ISSN 2199-8744 (online)

competence in competition + health

series

#2019/04

Arthur P. Hayen, Tobias J. Klein, and Martin Salm

Does the framing of patient cost-sharing incentives matter?

The effects of deductibles vs. no-claim refunds

Bundesministerium
für Bildung
und Forschung

Arthur P. Hayen, Tobias J. Klein, and Martin Salm [*]

Does the framing of patient cost-sharing incentives matter?[†]

The effects of deductibles vs. no-claim refunds

Abstract

Understanding how health care utilization responds to cost-sharing incentives is of central importance for providing high quality care and limiting the growth of costs. While there is compelling evidence that patients react to financial incentives, it is less well understood how and why specific aspects of the design of contracts shape the size of this reaction. In this paper, we focus on the question whether the framing of cost-sharing incentives has an effect on health care utilization. To study this, we make use of a policy change that occurred in the Netherlands. Until 2007, patients received a no-claim refund if they consumed little or no health care; from 2008 onward there was a deductible instead. This means that very similar economic incentives were first framed in terms of smaller gains and later as losses. We use claims-level data for a broad sample from the Dutch population to estimate whether the reaction to economic incentives was affected by this. Our empirical approach is to exploit within-year variation using an instrumental variables approach while controlling for differences across years. Our central finding is that patients react to incentives much more strongly when they are framed in terms of losses. Simulations based on our estimates show that the effect on yearly spending is 8.6 percent. This suggests that discussions on the optimal design of cost-sharing incentives should not only involve coinsurance rates and cost-sharing limits, but also how these are presented to patients.

Keywords: incentives contracts, altruism, dynamic panels, healthcare

JEL Classification Codes: C22, C23, D21, D22, I18

* *Hayen: Tranzo, Tilburg University; Klein: Netspar, CentER, Tilburg University; e-mail: T.J.Klein@uvt.nl; Salm: CentER, Tilburg University.*

† *First version: January 2017. We are grateful for the numerous insightful discussions we had with Johan Polder. We would also like to thank Viola Angelini, Arthur Attema, Jan Boone, Friedrich Breyer, Rudy Douven, Peter Eibich, Gautam Gowrisankaran, Hans van Kippersluis, Jürgen Maurer, Minke Remmerswaal, Pietro Tebaldi, Valentina Tonei, AlbertWong, conference participants at the RIVM/Celsus symposium in The Hague, the 2017 Netspar International Pension Workshop in Leiden, the workshop on Health Economics and Health Policy in Heidelberg, the LOLA Health Economics Study Group meeting in Rotterdam, the IHEA congress in Boston, the ASHECON conference in Atlanta, the European-American Health Economics Study Group in Boston, the EUHEA conference in Maastricht, and the 2018 CEAR/MRIC Behavioral InsuranceWorkshop in Munich, as well as seminar participants at Erasmus University Rotterdam, Tilburg University, the Hamburg Center for Health Economics, the Technical University in Zurich, the University of Lausanne, Trier University, Imperial College, and Queen Mary for many helpful comments and suggestions. Funding of the PACOMED project by the Netherlands National Institute for Public Health and the Environment (RIVM) is gratefully acknowledged. Disclosure statement: After starting this project Arthur Hayen has taken up employment at a Dutch health insurer. This did not give rise to a conflict of interest.*

1 Introduction

High and increasing health care expenditures are a major policy concern in many countries. One of the main policy tools to reduce medical spending is to introduce or expand patient cost-sharing schemes that let patients pay for a portion of health care expenditures not covered by health insurance. Across OECD countries, patients now pay on average around 20 percent of total health care expenditures out-of-pocket (OECD Indicators, 2015). Evidence from both field experiments (Manning et al., 1987; Finkelstein et al., 2012) and natural experiments (Chandra et al., 2010; Brot-Goldberg et al., 2017) shows that individuals reduce medical spending in response to economic incentives that follow from cost-sharing schemes such as deductibles and co-payments.

While there is compelling evidence that patients react to financial incentives, it is less well understood how and why details of the design of contracts shape the size of this reaction. In this paper, we focus on the question whether the framing of cost-sharing incentives affects the reaction to those incentives. For this, we make use of a policy change that occurred in the Netherlands. Until 2007, patients received a no-claim refund if they consumed little or no health care. The refund was the lower the more care they had consumed. From 2008 onward, there was a deductible. This means that economic incentives were first framed in terms of smaller gains and later as losses. Our central hypothesis is that patients are loss-averse and form their reference point in a way so that they react more strongly to incentives when they are presented in terms of losses. This implies that patients react stronger to deductibles than to no-claim refunds.

For several reasons, the Dutch setup provides a favorable setting to study this question. First, the policy change we exploit affected all individuals above the age of 18. Insurance contracts were required by law to feature a refund in 2006 and 2007 and a deductible from 2008 onward. At the same time, there were no major changes on the supply side. Second, the Netherlands has a system of universal health insurance where insurance status is not tied to labor market status, employer, or age. This allows us to study the effect for a broad population of insured individuals.

Our analysis is based on administrative data for the years 2006 to 2015. The data are from a Dutch health insurer whose population of insured individuals is broadly representative for the Dutch population. They include claim-level information on health care utilization, which we aggregate to around 9 million person-month observations, as well as background variables such as neighborhood income, age, and gender.

To measure the effects framing has on the reaction to cost-sharing incentives, we must overcome two empirical challenges. The first challenge is that we have a before-after design: the framing of cost-sharing incentives changed for all individuals at the beginning of 2008. While there were no other big reforms that became effective at the same time, we need to control for smaller changes that occur over time, including changing prices. The second challenge is that within a year the incentives patients face are endogenous: individuals who have reached the cost-sharing limit at a given point in time and therefore do not have to pay out-of-pocket anymore are on average sicker than individuals who have not reached the cost-sharing limit yet. Our empirical approach to overcome these two challenges is to exploit differences in the within-year variation of incentives across individuals using an instrumental variables approach. As we explain in detail in Section 6 this can be seen as a triple differences-in-differences grouping estimator. The groups consist of individuals with a similar risk score, age, and gender. The first difference is within groups and year and across months; the second difference is across groups; and the third difference is across policy regimes. Implementing this by means of a instrumental variables approach with a simulated instrument allows us to also account for differences in seasonal trends across groups within the year and autocorrelation of health care expenditures across months.

Our central finding is that patients react to incentives much more strongly when they are framed in terms of losses. In particular, we find that cost-sharing decreases health care expenditures under both a no-claim refund and a deductible policy, but the decrease is much larger under a deductible policy. When patients have to fully pay out-of-pocket monthly health care expenditures are reduced by 16.9 percent under a no-claim refund and by 38.7 percent under a deductible. While these effect sizes might seem large, our estimates of price effects for the deductible scheme are similar to the results Brot-Goldberg et al. (2017) report for a sample of employees of a large firm in the U.S.

Simulations based on our estimates show that the overall effect that is due to framing is a reduction in yearly spending of 8.6 percent. This suggests that discussions on the optimal design of cost-sharing incentives should not only involve coinsurance rates and cost-sharing limits, but also how these are presented to patients.

In addition, we provide a more in-depth analysis. First, we investigate whether the average effects we estimate are driven by particular groups or whether framing has an effect across all groups. For this we leverage the fact that our sample is for a broad population and that we observe background variables such as income in the neighborhood. Perhaps surprisingly we do not find significant differences in the size of the effect by age and by income in the neighborhood, but we do find that women react stronger to prices than men and that effects are stronger for individuals with lower risk scores than for individuals with higher risk scores.

Second, we turn to the question why changing from a no-claim refund to a standard deductible had such a large effect. Our preferred explanation is that individuals are loss-averse (Kahneman and Tversky, 1979) and that they perceive their reference point in a way so that they see a deductible payment as a loss and no-claim refund as a gain; as a consequence they will react stronger to a deductible than to a no-claim refund. This hypothesis was developed by Johnson et al. (1993) in a theoretical contribution. We discuss why this is our preferred explanation and show that our findings are likely not driven by differences in the timing of payments, liquidity constraints, differences in information about cost-sharing incentives between both schemes, and salience.

This paper contributes to the literature on the effect of patient cost-sharing on health care expenditures. A focus of recent studies is a better understanding of the behavioral underpinnings of patients' responses to cost-sharing schemes, e.g. to what degree patients respond to the dynamic incentives provided by cost-sharing schemes (Einav et al., 2015; Abaluck et al., 2018; Dalton et al., ming), whether co-payments cause patients to shop for cheaper providers (Brot-Goldberg et al., 2017), or how responses vary by type of medical service (Ellis et al., 2017).[1] In our study, we expand the knowledge about patients' responses to cost-sharing schemes in a new direction by showing that the framing of patient cost-sharing—as a deductible or as a no-claim refund—can have a large impact on health care utilization.

In contemporaneous independent research, Remmerswaal et al. (2017) show that the difference in annual health care expenditures in the Netherlands between 19 and 17 year olds is bigger in years in which a deductible was in place than in years when a no-claim refund was in place. Our study has several advantages compared to Remmerswaal et al. (2017). First, our data and identification strategy allow us to estimate the effect of patient cost-sharing across all age groups and not just for people around the age of 18, which adds to the external validity of our study. Second, we do not assume that annual health care expenditures are a linear function of the amount of deductibles and no-claim refunds. Instead, we examine individual responses to whether or not they have exceeded cost-sharing limits at a given point of time within a year. Third, we are able to say more about the underlying mechanism.

More generally, we contribute to the literature on the effect of patient cost-sharing on health care expenditures in the Netherlands. Recent contributions to this literature include Ecorys (2011) and Lambregts and van Vliet (2017).[2]

In the following, Section 2 describes the institutional background. A conceptual framework based on loss aversion is proposed in Section 3. Section 4 provides details on the data. Section 5 presents descriptive evidence. Section 6 describes the empirical approach. Section 7 presents the main results. Section 8 discusses potential alternative mechanisms other than loss aversion, and Section 9 presents robustness checks. Section 10 discusses policy implications and concludes.

2 Institutional background

The Dutch health care system lends itself well for separating framing effects from other determinants of demand. Insurance plans with a no-claim refund in the years 2006 and 2007 and insurance plans with a deductible in the years thereafter were similar in terms of coverage and

[1] More broadly, it contributes to the behavioral literature in health economics. See survey by Chandra et al. (2019).

[2] A survey of the literature on patient cost-sharing in the Netherlands can be found in Remmerswaal et al. (2015).

services for which there was cost-sharing. These plans also provided similar economic incentives because the deductible limit was similar in size to the no-claim refund limit. At the same time, there were no substantial changes on the supply side.

Next, we provide some background information on the health care system in the Netherlands, the two cost-sharing schemes, and the way in which providers are paid.[3]

2.1 Health insurance in the Netherlands

The Netherlands has a system of universal, mandatory health insurance coverage. Since a reform of the health insurance system in the year 2006, insurance is funded in about equal proportions by income-dependent employer contributions and premiums paid by the insured. At the same time, there is a risk equalization scheme between insurance providers.

Every resident of the Netherlands is obliged to purchase a basic, nationally standardized insurance plan on the market, from one of several competing private health insurers. The premiums are community-rated, meaning that the insurance premium is similar across patients, irrespective of characteristics like age, gender, and health status. Premiums may only be different for individuals who choose an additional voluntary deductible or enroll via a collective bargaining agreement (in which case they can receive a discount of up to a 10% of their monthly insurance premium). Furthermore, health insurers are not allowed to deny basic health insurance coverage.

The basic health insurance plan includes a wide range of services, such as general practitioner (GP) services, hospital care, prescription drugs, mental health care, geriatric care, and medical devices such as hearing aids and prostheses.[4] On top of this basic insurance package, individuals can buy additional coverage on the market for supplemental health insurance, for example for dental care.

Private insurers compete with one another in terms of policy conditions such as monthly premiums and service quality. Insurers negotiate contracts with providers that typically contain agreements on prices, volumes, and quality (Krabbe-Alkemade et al., 2016). At the start of each year, insurees are allowed to switch to another health insurer. Between 2006 and 2012, the annual switching rates ranged between 4 and 6 percent (NZA, 2012).

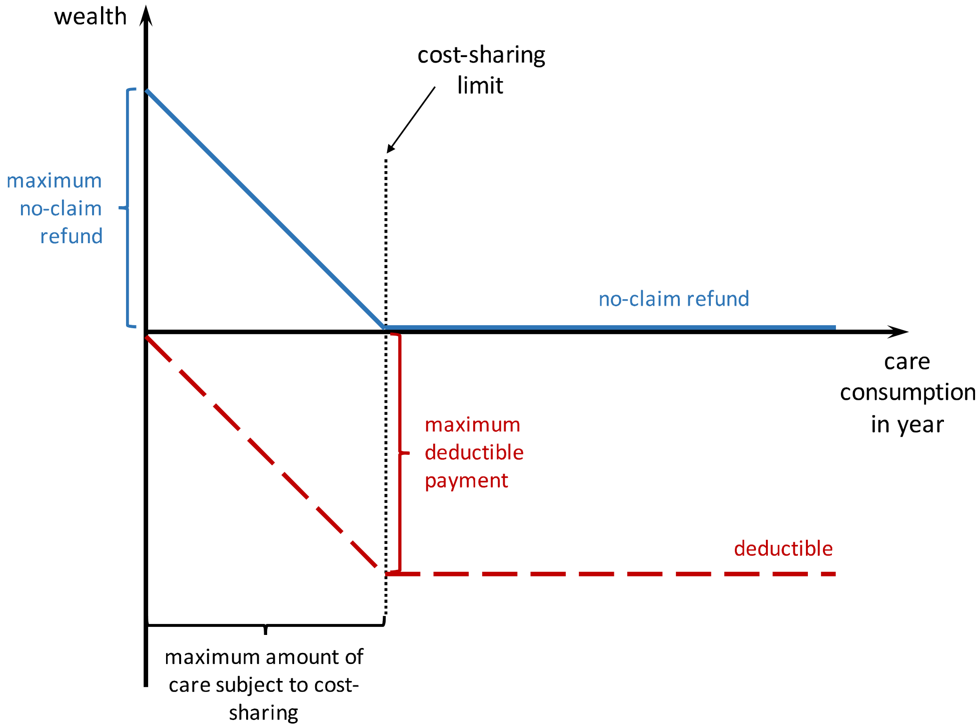
2.2 Patient cost-sharing

In this paper we make use of the fact that two alternative cost-sharing schemes have been in place at different points in time. Both schemes were mandatory and uniform across health insurance plans. In the years 2006 and 2007, health insurance plans had to feature a no-claim

[3]See Schäfer et al. (2010) for additional details.

[4]Services included in the basic health insurance plan changed slightly over the years. The most notable change was the inclusion of mental health care services and geriatric care in the year 2008. As a robustness check we estimate a specification in which we exclude mental health care and geriatric care.

Figure 1: No-claim refund and deductible



Notes: This figure shows financial implications of care consumption under a no-claim refund and a deductible. The cost-sharing limit is the same for both.

refund, and since 2008 health insurance plans have to feature an annual deductible.

Under the no-claim refund enrollees receive a refund at the end of the year if their annual health care spending during the year was below the no-claim refund limit. The amount of the refund is equal to the difference between the no-claim refund limit and annual health care spending. Under the deductible individuals pay out-of-pocket for all care up to the deductible limit. Health care utilization beyond the deductible limit is free from the patient's perspective.

Financial implications of care consumption from the patient perspective are illustrated in Figure 1. The horizontal axis depicts annual health care spending. The vertical axis shows wealth relative to the situation after paying insurance premiums: under a no-claim refund, patients receive a weakly positive payment; under a deductible they make payments up to the maximum deductible when consuming care. In the figure, the no-claim refund limit is set equal to the deductible limit. The two curves for no-claim refunds and for deductibles have the same shape, but they are shifted vertically. The identical shape of the two curves reflects that the "price" that patients have to pay for an additional euro of health care utilization is identical under both schemes for any level of annual health care spending. The vertical shift between the two curves does not imply that health care is cheaper under a no-claim refund than under a deductible. Instead, the vertical shift will be offset by different levels of health insurance premiums, which will be higher under the no-claim refund than under the deductible.[5]

The no-claim refund limit in the years 2006 and 2007 was 255 euros. The deductible limit started at 150 euros in the year 2008, and it was subsequently increased in several steps to 375 euros in the year 2015.[6]

Minors up to the age of 18 years are excluded from patient cost-sharing under both the no-claim refund and the deductible scheme (a feature we make use of in Section 9.1 to conduct a placebo test). Furthermore, certain types of medical services such as GP care and maternity care are also excluded from patient cost-sharing (GPs are a patient's first point of contact when in need for health care, and they have a gatekeeping function in the Dutch health care system, see Schäfer et al., 2010).

Under both regimes, individuals could opt for a voluntary, additional deductible (in quantities of 100 euros, with a maximum of 500 euros).[7] In return, they received a discount on their monthly premium.

Even though cost-sharing rules apply uniformly to all patients, people with a chronic illness can apply for a subsidy that partially compensates them for their higher expected deductible payments. Eligibility is based on whether or not people have received treatment for chronic illnesses in the previous two years. It does not depend on current year's spending. Despite the fact that many people qualify for this subsidy, a substantial share of the eligible population is not aware of the rule (Reitsma-van Rooijen and Jong, 2009).

2.3 Payment of providers

In the Netherlands, providers are predominantly paid for bundles of services (Schäfer et al., 2010). For specialist care, these bundles are called "diagnosis-treatment combinations (DTCs)" (Hasaart, 2011; Van de Ven and Schut, 2009). The specialist assigns the patient to a DTC upon first contact, but is allowed to adjust this choice later should there be changes in the treatment process (see Hasaart, 2011).

In general, providers bill the insurance company directly, even if the patients have to pay out-of-pocket (as opposed to billing patients who then ask for reimbursement). Billing is usually done with a substantial delay. Figure C.1 in the Online Appendix shows that the delay between the first contact and the date the bill was sent to the insurer depends on the size of the claim. It is 90 days for claims between 100 and 200 euros and more than 150 days for claims above 200 euros.

When processing incoming bills, health insurers calculate how much patients will have to pay out-of-pocket and, if applicable, send a bill to patients. If patients were unable to pay the

[5]This means that patients could also perceive the two schemes as equivalent. We further discuss this in Section 3.2.

[6]The deductible limits were 155 euros in 2009, 165 euros in 2010, 170 euros in 2011, 220 euros in 2012, 350 euros in 2013, and 360 euros in 2014.

[7]The number of individuals doing so was almost negligible. We exclude them from our analysis. See Section 4 for details.

bill immediately then it was usually possible to pay in installments. The no-claim refund for a given year was paid out at the end of the first quarter in the following year. This means that there was a delay for both deductible payments insurers received from patients and no-claim refunds patients received from the insurers. The delay was generally shorter for deductible payments.[8]

2.4 Patients' information

Survey evidence shows that knowledge about the existence of both no-claim refunds and deductibles was close to universal in the Netherlands during our study period (Ecorys, 2011; Goudriaan et al., 2007; Van der Maat and de Jong, 2010).[9] An exception was knowledge about the existence of a deductible immediately after its introduction in the year 2008 when only 42% of respondents to a representative survey knew about the deductible. However, already one year later this share increased to 95%.

Under the no-claim refund policy the health insurer sent a letter to insurees that informed them as soon as they had exceeded the no-claim refund limit. Under the deductible policy, patients receive invoices for deductible payments during the course of the year. From these invoices, they can see whether they still have to pay for the next treatment. Importantly, in either case patients were informed about exceeding the cost-sharing limit as soon as the health insurer had received the relevant invoices from the providers. In addition, patients knew which treatments they received and could use this knowledge to inform their beliefs about whether or not they had exceeded the cost sharing limit in a given year. Arguably, given that the cost sharing limit was very low, patients had a fairly good idea about whether or not they had reached it at a given point in time. For instance, if they had received a big treatment that for example involved an overnight hospital stay, then they could reasonably expect that they had exceeded the cost-sharing limit even if they did not know the exact cost of the treatment.

3 Conceptual framework

3.1 Non-linear price schedules

Our study examines how patients react to cost-sharing under a deductible and no-claim refund policy, respectively. Defining the price of care from the patients' perspective is not straightforward.

Under both policies individuals make decisions throughout the year. The price of care is a function of past health care consumption in the same year, and it can thus vary over the course of a year. Forward-looking decision-makers should take this into account when making choices about health care utilization earlier in the year. In two seminal theory articles, Keeler et al.

[8] In Section 8.1 we investigate whether this could be an alternative explanation for our findings.

[9] At the same time, knowledge about specific details was less than universal. For instance, individuals were less aware of exemptions from cost sharing for maternity care.

(1977) and Keeler and Rolph (1988) present a set of assumptions including risk neutrality and separability of the utility function in health and money under which the expected end-of-year price—this is equal to one minus the probability of hitting the cost-sharing limit by the end the year—is the only relevant price. Ellis (1986) refines this by also incorporating risk aversion and derives the relevant shadow price. Based on this, one could proceed and spell out a structural model and estimate it using our data.

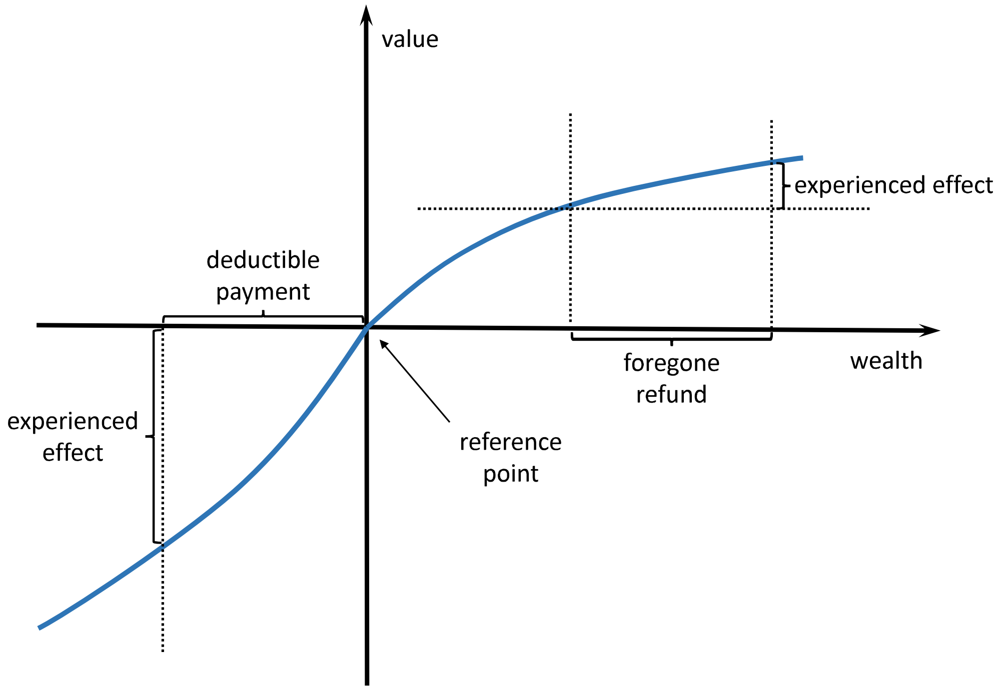
Here, we instead follow Brot-Goldberg et al. (2017) and Ellis et al. (2017) and say that the price is one if an individual has to pay for the first euro of a treatment herself and zero if a treatment is fully covered by the insurance. At the beginning of the year the price of care is 1. After patients have exceeded their annual cost-sharing limit the price they pay for additional care is 0. Thereby, we focus on the question of how health care consumption depends on the current price of health care at the point in time at which the decision is made and the question how this reaction depends on the framing of cost-sharing. We do so for two related reasons. First, recent empirical studies find evidence for substantial myopia in patients' decisions about health care use in the presence of non-linear price schedules. In particular, Brot-Goldberg et al. (2017) show that patients react to current prices even after controlling for expected end-of-year prices.[10] Second, health care consumption is lumpy and deductibles are relatively small in the Netherlands. Both the current price and the expected end-of-year ("future") price are functions of past health care consumption, and because of the lumpiness and the small cost-sharing limits both are highly correlated. In our estimation sample, this correlation is 0.8.

We intentionally do not control for the *evolution* of the future price over the course of the year, as it is not the focus of this paper whether the reaction we observe is a reaction to the future price or the current price.[11] We will however control for differences across individuals by controlling for the future price at the beginning of the year. Further details are provided in Section 6.

[10] Also Abaluck et al. (2018) and Einav et al. (2015) find evidence for a high degree of myopia for Medicare Part D patients. Abaluck et al. (2018) find that a model with substantial myopia fits the data well. Einav et al. (2015) estimate a discount rate of around 7 percent per week, which means that the yearly discount rate is 98 percent. Their preferred interpretation of this model parameter is that it is "a behavioral parameter that also reflects individuals' understanding of the insurance coverage contract, in particular the salience of the (future) nonlinearities of the contract." Dalton et al. (ming) relate this to current prices being more salient.

[11] In fact, the coefficient on the current price will capture the effect of both the evolution of the current price *and* the evolution of the future price. Qualitatively, we will see a similar reaction in either case, as prices are highly positively correlated. Because of the correlation, our estimates will put the more weight on the current price the more individuals discount. To see this, consider the two polar cases. Suppose individuals are fully myopic. Then, they will react only to the current price, which is included in the regression. Instead assume that they are fully forward-looking and don't discount. Then, they will react only to the future price. As this price is highly positively correlated with the current price that is included in the regression, we will qualitatively capture the effect of the future price even if we include the current price into the regressions.

Figure 2: Experienced effects



Notes: This figure shows the experienced effects of a foregone refund and of a deductible payment. The amount for both is the same.

3.2 Framing as gains and losses

Loss aversion, one of the elements of prospect theory, postulates that decision-makers value losses more than corresponding gains (Kahneman and Tversky, 1979). This concept can also be applied to patient cost-sharing for health insurance where cost-sharing incentives can be perceived either as a loss or a gain. Whether patients perceive payments as a loss or a gain depends on their reference point.

In a theoretical contribution, Johnson et al. (1993) consider the case when patients perceive their financial situation after paying insurance premiums as reference point. This is the situation we have depicted in Figure 1. Compared to this reference point they will view a deductible payment as a loss. In contrast, if they receive a refund at the end of the year under a no-claim refund policy, then they will view this payment as a gain relative to their reference point. Thus, patients should react stronger to cost-sharing under a deductible policy than under a no-claim refund policy.

This is illustrated in Figure 2, which shows a utility curve ("value") under loss aversion when the cost-sharing limit has not been reached. The reference point is the financial situation after paying health insurance premiums. The utility curve is steeper to the left than to the right of the cutoff point, and in both directions the utility curve becomes flatter with increasing distance to the reference point. The figure shows that the experienced effect of a deductible payment is higher than the effect of an equivalent reduction in the refund under a no-claim refund policy. When patients make decisions about medical care use they will compare the value of the treatment with the value they assign to their own payment, either a deductible payment or a foregone refund under a no-claim refund policy. Thus, patients will use less medical care under a deductible policy than under a no-claim refund policy.[12]

Above, we have assumed that patients perceive the financial situation after paying health insurance premiums as their reference point. However, van Winssen et al. (2016) point out that individuals may also have other reference points such as their financial situation before paying insurance premiums. In that case, they will feel to be in the loss domain under both schemes. Then, there will be no stronger response to the price under a deductible than under a no-claim refund.

Our central hypothesis is therefore that individuals are loss-averse and form their reference point in a way so that they react more strongly to incentives when they are framed in terms of losses instead of gains. This hypothesis is testable, as individuals would *not* react to the difference in framing if *either* it is the case that individuals are not loss-averse *or* if they use the same reference point under a no-claim refund and a deductible.

4 Data

Our data come from a large Dutch health insurer. They were obtained under a pilot project in which a new payment model for GPs was designed and implemented (see Hayen et al., 2015, for details). This pilot ran from July 2014 to the end of 2015, but the data also include pre-intervention years. The data cover the years 2006 to 2015.

From a set of approximately 200 GPs that participated either in the intervention or were part of the control group, we first selected all insurance enrollees who were registered with those GPs. For our main analysis, we use observations for individuals who were at least 19 years old (such that they were subject to cost-sharing for the entire year—information on the month of birth is not available to us), were insured for the entire year, and did not opt for a voluntary deductible.[13] Excluding individuals with a voluntary deductible from our sample does not cause serious concerns, as we exclude only about 2 percent of the observations for that reason.[14] Our estimation sample consists of an unbalanced panel with about 85,000 insured individuals in any

[12]In our empirical analysis, for reasons we have given in Section 3.1, we study the reaction to current prices. The analysis is carried out on a monthly level. The reasoning in the paragraph above also applies to the monthly level: Figure 2 then shows utility defined over gains and losses at the monthly level, when the reference point is wealth after the monthly premium has been paid.

[13]Individuals who opt for a voluntary deductible are likely different in terms of loss aversion, risk attitude, ambiguity aversion, debt aversion, omission bias, and liquidity constraints (van Winssen et al., 2016).

[14]The fraction of individuals with a voluntary deductible varies over the years. In the beginning, it was about 1 percent and then grew to 4 percent over the course of 10 years.

given year.[15]

To assess the representativeness of our sample, we compare the average age and the proportion of women, both among individuals who are at least 19 years old, to the respective values for the Dutch population. Differences are small. Individuals in our sample are about 1 year older on average and the percentage of women is by one percentage point higher.

Our data include information on each person's insurance policy, length of enrollment, and medical care utilization at the claims level. These claims-level data are very detailed. Among others, they include the billing code of the particular service, a full description of the service, the full DTC for medical specialist care, the start date of the treatment, the billing date (by the provider to the insurance company, not the billing date to the patient), the amount paid by the insurance, and the amount paid out-of-pocket by the patient. For prescription drugs, we know the exact quantity delivered (e.g. 5 ml of Tobramycin) and their classification according to the Anatomical Therapeutic Chemical Classification System with defined daily doses (World Health Organization, 2006).

We inferred information on whether or not an individual has one of 25 chronic conditions (e.g. type 2 diabetes) based on claims for prescription drugs. In case a person has billed more than 180 daily doses of medicines linked to the treatment of a particular chronic condition, she was believed to have this condition.[16]

The data also contain demographic information (age, gender), and a person's postal code at the 6-digit level.[17] We use the latter to add information on socioeconomic status: we add income (as measured in December 2008) and the percentage of non-western immigrants (as measured over 2010) at the 6-digit postal code level using data from the Dutch Central Bureau of Statistics (Statistics Netherlands, 2012). We also add 2006 socioeconomic status scores at the 4-digit postal code level, as calculated by the Netherlands Institute for Social Research (Netherlands Institute for Social Research, 2006).[18] These status scores are calculated based on income, education, and employment status.

Outcome variables in our study are measures of health care expenditures. Our measures of

[15] People who met this requirement were both insured with the health insurer, and registered with one of the GPs in our sample. The panel is unbalanced because individuals may have changed insurer or may have chosen a voluntary deductible in some years (and were therefore excluded in those years). In our analysis below, we control for time effects and thereby also for differences in sample composition over the years. Individuals in our sample were always observed for entire years and our main parameters of interest are estimated from within-year variation.

[16] We obtained additional data from the Royal Dutch Pharmacists Association to translate each prescription drug claim into the number of defined daily doses obtained. For each individual, we then summed the number of defined daily doses obtained per medicine group as defined by their 6-digit Anatomical Therapeutic Chemical Classification (e.g. metformine) and used data from the Dutch Health Care Institute to link a medicine's classification to the treatment of chronic conditions (e.g. metformine is used for the treatment of type 2 diabetes).

[17] A 6-digit postal code consists of a number with 4 digits and two letters. In 2007, there were 456,913 different postal code areas (http://postcodeinfo.nl/files/factsheet-postcodeinfonl.pdf, accessed November 2017) and 17 million inhabitants (http://statline.cbs.nl/StatWeb/publication/?PA=37296ned&D1=a&D2=0,10,20,30,40,50,60,(l-1),l&HDR=G1&STB=T, accessed November 2017). This means that there are on average 37 inhabitants per postal code area.

[18] There are 4,054 4-digit postal codes (https://home.kpn.nl/pagklein/postcodes.html, accessed November 2017). This means that there are on average 4,193 inhabitants per postal code.

health care expenditures include all care that is included in the basic health insurance package and is in principle subject to patient cost-sharing. The latter restriction excludes GP care and maternity care. Our two outcome variables in the main analysis measure monthly health care expenditures at the intensive margin and at the extensive margin, respectively. Our measure of health care expenditures at the intensive margin is given by $y_{it} = \log(spending_{it} + 1)$, where $spending_{it}$ are the health care expenditures of individual i in month t.[19] Our measure of health care expenditures at the extensive margin is a binary indicator that equals 1 in case a person consumed a positive amount of medical care in a given month, and it equals 0 in case she did not consume medical care in this month.

Our main independent variable in both models is the price p_{it} for person i in month t, which we define as the marginal cost-sharing rate for the first unit of care consumed within a month. This price equals 0 in case a person has exhausted her deductible or exceeded the no-claim refund limit at the beginning of a month, and 1 in case she has not. Whether or not medical spending is above the cost-sharing limit is determined based on the treatments that have been started up to that point of the year, independent of whether this care has already been billed to the health insurer (see Section 2.3 and 2.4).

Additional control variables include age, gender, risk score deciles, and the expected end of year price at the beginning of January. The risk score is given by a person's predicted annual expenditures divided by the sample's average annual expenditures. We obtained this prediction by performing a linear regression of annual health care expenditures on several risk characteristics, which are age and gender (gender fully interacted with a third-order polynomial in age), last year's spending decile, indicators for chronic illnesses in the previous year, and quartile splits for income, the percentage of non-western immigrants, and the socioeconomic status score in the neighborhood. With this we aim to mimic the way risk scores are calculated in Dutch risk adjustment scheme between health insurers.[20] Based on individual risk scores we compute a set of risk score decile indicators at the person-year level. We also estimate the expected end-of-year ("future") price from the perspective of the beginning of the year (see also Section 3.1). This price is calculated as the average end-of-year price at the level of a cell formed of the risk score decile and age/gender category (18-40, 40-65, and over 65 years of age).

Table 1 presents summary statistics for three different samples. The first column is for our estimation sample. The second column is for the full sample of individuals in our data who are at least 19 years old. The third column is for the sample of individuals who are in our data in all years. About half of the individuals in our estimation sample are female. They were in their late 40s on average. Average gross income at the postal code level is about 2200 euros per

[19] See Online Appendix A.1 for details on how this was constructed.

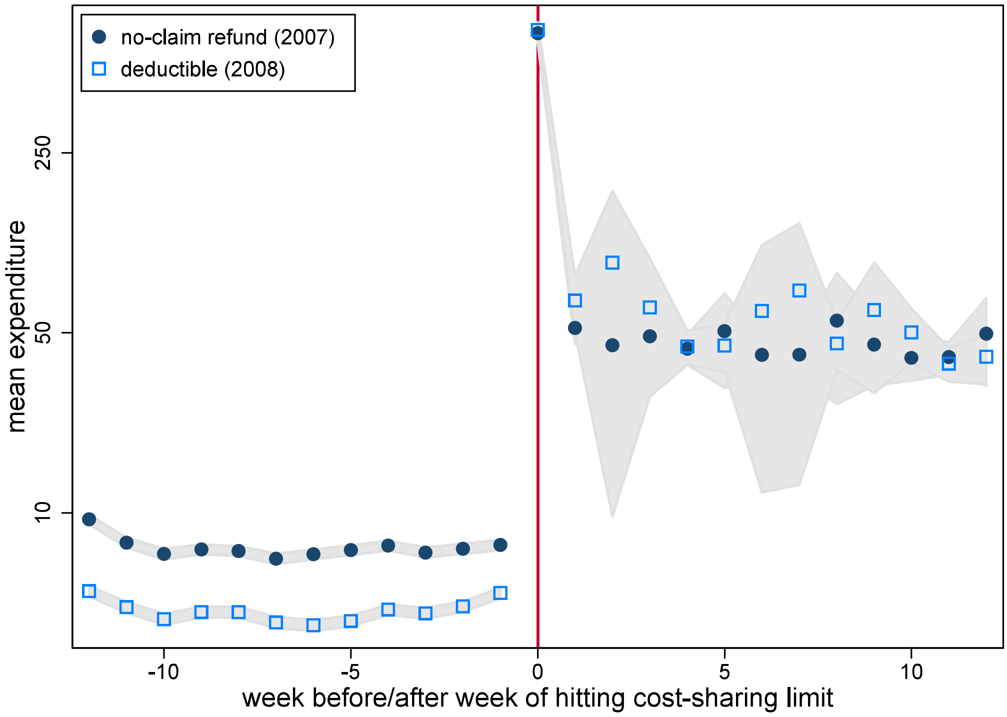
[20] See Online Appendix A.2 for details. In the national risk-equalization scheme, insurance companies compensate one another for differences in their risk pools. This is necessary because risk-rating is not allowed. The actual risk scores in the national risk-equalization scheme use additional information that was not available to us. For instance, they use income at the individual level, while we use income at the 6-digit postal code level. We expect the difference to be small, as a postal code covers only 37 persons on average (footnote 17).

Table 1: Summary statistics for different samples

	estimation sample	full sample	balanced sample
individual and neighborhood characteristics			
female	0.52	0.52	0.52
age	49.80	48.75	53.77
income (if not missing)	2182.37	2188.60	2167.38
indicator income missing	0.04	0.04	0.03
care consumption per month			
expenditure	202.55	194.72	186.02
has claim	0.42	0.41	0.44
number observations	8,771,940	10,965,456	3,854,952

Notes: This table shows summary statistics for individuals who are at least 19 years old, insured for the entire year, and choose the minimum deductible. One observation is a person-month. The first column is for all individuals in our estimation sample who automatically meet these requirements. The second column is for the full sample of individuals meeting these criteria. There are more observations because in our estimation sample we also require that information for the respective previous year is available. This information is used to calculate the risk score. The third column is for the balanced sample of individuals who have been at least 19 years old in 2007 and were in our data until 2015. All presented values are means across individuals and time. Statistics for the estimation and the balanced sample are reported as averages across the years 2007 to 2015. Income is gross income at the 6-digit postal code level in nominal euros. Care consumption is the consumption of care that falls in principle under the no-claim refund policy or the deductible. Has claim is one if care consumption is non-zero in a given month.

Figure 3: Care consumption around week in which cost-sharing ends



Notes: This figure shows average health care expenditures in weeks before and after exceeding the no-claim refund limit or the deductible in 2007 and 2008, respectively. Cost-sharing incentives were framed as a no-claim refund in 2007 and as a deductible in 2008.

month. In a given month, average health care expenditures that are subject to cost-sharing are about 200 euros, and the likelihood that there are positive expenditures in a given month is 42 percent.[21]

Comparing means across columns shows that individuals in the balanced sample, who remained with both the insurer and the sampled GPs throughout our study period, are somewhat older on average. They have a higher probability to have a claim in a given month, but lower average health care expenditures. Still, differences across samples are not big, especially given the fact that the balanced sample is less than half the size of the estimation sample. As a robustness check, we also obtain estimation results for the balanced panel, and we find that results are very similar for our baseline estimation sample and the balanced panel.[22]

5 Descriptive evidence from an event study

The central question in this paper is whether the way patients react to cost-sharing incentives depends on the framing. Figure 3 presents model-free evidence from an event study. The

[21] Table C.3 in the Online Appendix presents average care consumption and the fraction who hit the cost-sharing limit by year, respectively. Figure C.3 plots the consumption of care within the year. It shows that there is no strong seasonal component.

[22] These results are shown in Table C.6.

general empirical approach in this paper is different and explained in detail in Section 6.

For constructing the figure we follow individuals who exceed the no-claim refund or deductible limit in a given year, and we compare their health care spending in the weeks before and after hitting the cost-sharing limit. We do so for a balanced panel of individuals whom we observe at least 12 weeks before hitting the cost-sharing limit and 12 weeks thereafter in the same calendar year. This means that we exclude all observations for individuals who do not hit the cost-sharing limit during the year or who hit the cost-sharing limit very late or very early in the year.

The horizontal axis shows weeks relative to the week of hitting the deductible limit (or the no-claim refund limit). The vertical axis shows average weekly health care expenditures. Each dot shows the average health care expenditures for a different week alternatively in the year 2007, when a no-claim refund was in place, and in the year 2008, when a deductible policy was in place. To the right of the vertical line, after hitting the cost-sharing limit, the price of care is zero. The figure shows that after reaching the cost-sharing limit average expenditures are very similar under the no-claim refund in 2007 and the deductible in 2008. By construction, there is a peak in health care expenditures in the week that the deductible limit (or the no-claim refund limit) is hit and in the weeks thereafter health care expenditures show no clear trend over a 12 week period. To the left, before hitting the cost-sharing limit, the current price is one, and individuals are subject to cost-sharing incentives. Under both schemes, average expenditures are lower to the left than to the right of the red line. However, expenditures are substantially lower in 2008 under the deductible than in 2007 under the no-claim refund. Accordingly, the increase in health care expenditures at the time of exceeding the cost-sharing limit is much larger under the deductible scheme than under the no-claim refund scheme.

Thereby, this figure shows strong suggestive evidence for our main hypothesis that health care expenditures respond stronger to deductibles than to no-claim refunds. We call this evidence suggestive because the increase in health care expenditures upon hitting the deductible cannot be interpreted as causal effect. It could also be explained for example by serial correlation in health care expenditures if high health care expenditures in the week of hitting the cost-sharing limit tend to be followed by other weeks with high health care expenditures. In the following sections we turn to a more formal analysis to address such concerns.

6 Empirical approach

Our goal is to estimate the effect of the price of care on health care utilization, alternatively under a no-claim refund policy and a deductible policy. We specify the estimation equation

$$y_{it} = p_{it} d_t^{\text{no-claim refund}} \beta^{\text{no-claim refund}} + p_{it} d_t^{\text{deductible}} \beta^{\text{deductible}} + x_{it}' \gamma + \varepsilon_{it}, \tag{1}$$

where p_{it} is the price of care for individual i at the beginning of month t, $d_t^{\text{no-claim refund}}$ is a binary indicator for months in which the no-claim refund was in place, $d_t^{\text{deductible}}$ is a binary indicator for months in which the deductible was in place, and x_{it} is a vector of controls that includes a constant term. ε_{it} is an error term, and γ is a vector of parameters. $\beta^{\text{no-claim refund}}$ and $\beta^{\text{deductible}}$ are the main parameters of interest. They measure how health care expenditures respond to prices under a no-claim refund and a deductible, respectively. The difference between $\beta^{\text{no-claim refund}}$ and $\beta^{\text{deductible}}$ can be interpreted as the effect of framing.[23]

In our baseline specifications, the dependent variable y_{it} is either the log of health care costs of individual i in month t plus 1 (as health care costs can be zero), so that coefficients are approximately percentage changes in total spending. Or the dependent variable is a binary indicator for any expenditures in a given month, in which case coefficients measure effects at the extensive margin.

Determinants of health care costs in x_{it} include a full set of month and year dummies in order to control for seasonal variation in health care needs related to, for instance, the flu and for changes across years that are, for instance, driven by health care inflation. x_{it} also includes indicator variables for 5 year age bins and for gender. In addition, x_{it} includes indicators for the deciles of the individual risk score and the expected end-of-year price from the perspective of the beginning of the year (see Section 3.1). The risk score controls for differences in expected health care expenditures across individuals, and the expected end-of-the-year price captures in addition that cost-sharing limits and health care prices change over the course of our study period such that in different years individuals with the same risk score face a different probability of exceeding the cost-sharing limit and experiencing a zero price. Finally, x_{it} includes the log of lagged health care expenditures in the previous three months plus one in order to control for autocorrelation in health care expenditures.

Next, in Section 6.1, we discuss potential threats to identification of $\beta^{\text{no-claim refund}}$ and $\beta^{\text{deductible}}$. Then, in Section 6.2, we explain how we address related concerns using an instrumental variables approach.

6.1 Threats to identification

In (1), p_{it} is individual-month specific. It is a function of past care consumption by the same individual within the same calendar year: p_{it} is 0 when individuals have exceeded the cost-sharing limit and 1 otherwise. This means that a first concern could be that our estimates may suffer from omitted variables bias when we do not sufficiently control for time-invariant individual differences in expected spending that lead to a higher likelihood that p_{it} is 0.

To see this, suppose that the omitted variable was the risk score (for which we actually control in our analysis, among others). For a given p_{it}, individuals with a high risk score will

[23] For this interpretation we have to exclude other potential explanations for stronger reactions to cost-sharing incentives in years in which a deductible was in place. We do so in Section 8.

on average consume more care compared to those with a low risk score. Because prices are a function of past health care consumption, p_{it} will therefore on average be lower for individuals with a higher risk score, which means that a regression of health care consumption on prices will lead to an upward-biased estimate of $\beta^{\text{no-claim refund}}$ and $\beta^{\text{deductible}}$ when we do not control for the risk score. This concern is partly already addressed by including numerous control variables for health care needs in x_{it}, including the risk score. If the bias in both coefficients would be the same, then we could still measure framing effects by comparing both biased coefficients with one another. But it is not obvious why this would be the case.

A second potential concern is also related to the fact that p_{it} is a function of past health care consumption in the same year. The price in t depends on past values of ε_{is}, when $s < t$. If there is autocorrelation in the error term, for example because health shocks are correlated over time, then p_{it} will also be correlated with ε_{it}, because p_{it} is a function of ε_{is}, while ε_{it} is correlated with past ε_{is}. This will result in prices being endogenous.[24]

A third potential concern is that seasonal trends in health care consumption differ between groups of the population. In equation (1) we control for period effects, but not for interactions of period effects with individual characteristics. It could for instance be that younger individuals are more likely than older individuals to have skiing accidents, which tend to occur in the winter months at the beginning of the year. A consequence of this would be that prices are negatively correlated with the error term, as the error term also captures the deviation of the group-level seasonal trend for younger individuals from the general seasonal trend, which is captured by period fixed effects.

6.2 Instrumental variables approach

We overcome the three aforementioned threats to identification using an instrumental variables approach. Our instrument for the price is an average simulated price a particular group faces in a given month.

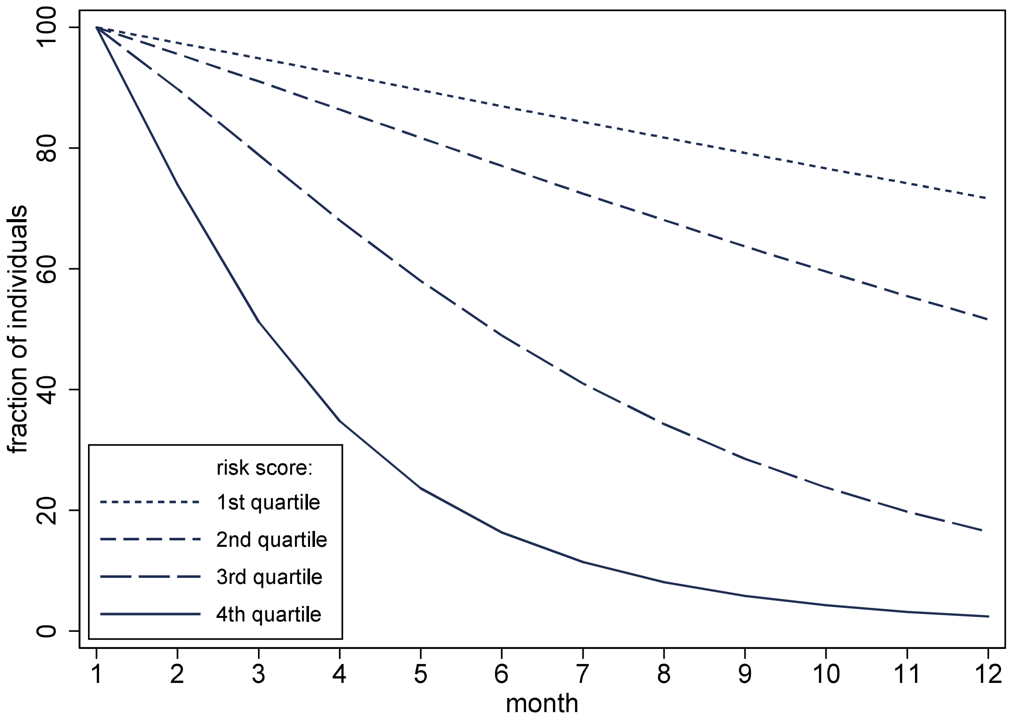
For this, we first define groups as a specific combination of the risk score decile, the age group, and gender.[25] Then we nonparametrically estimate the empirical distribution function of monthly health care consumption for each group. Here we pool across all months and all years.[26] Then, we take draws of monthly expenditures for a set of simulated individuals and compute whether they would have exceeded the cost-sharing limit at the beginning of a given month. From this we compute the simulated fraction of individuals in a group who have not hit

[24] Note that this means that we cannot use a fixed effects estimator to control for level differences, because a fixed effects estimator would only yield consistent estimates if strict exogeneity would hold. This does not hold here because the regressor p_{it} is correlated with some ε_{is}, $s < t$, for the same individual.

[25] Age groups are defined as age 0-18, 19-39, 40-64, and 65+. For the first four risk score deciles, we combine the two last age groups into one to ensure that there are enough observations per group. Therefore, there are 468 groups (10 deciles times 3 age categories times 2 genders minus 8 merged groups) in our baseline sample.

[26] Whether or not we correct for price changes between the years before estimating the empirical distribution function does not affect the validity of our instruments. In our analysis below we have not corrected for price changes. However, we have re-run the analysis taking price changes into account and found very similar results.

Figure 4: Simulated fraction of individuals below cost-sharing limit by risk score



Notes: This figure shows the simulated fraction of individuals who have not hit the cost-sharing limit. We present the average by quartile of the risk score, across years. Figure C.6 in the Online Appendix shows the actual (non-simulated) fractions by risk score quartile.

the cost-sharing limit for each month in each year. We do so separately for each year, because the cost sharing limit changed across years. This fraction is denoted by z_{it} and it is used as instrument for p_{it}.

Figure 4 shows the average value of the instrument by month and by risk score quartiles (i.e., here we average over years and aggregate groups to the 4 risk score quartiles for readability). It illustrates the variation in average prices between groups that we use to identify price effects.[27] For all patients, prices are 1 in January. For patients with high risk scores, average prices quickly decline over the following months and then approach 0. For patients with low risk scores, this decline is much slower and average prices are still close to 1 at the end of the year. In our empirical strategy we control for level differences in average prices between groups by including characteristics that were used to form groups in x_{it}.

Our instrumental variables approach is built on institutional knowledge, in particular knowledge on the mechanical nature in which prices depend on past health care consumption and cost-sharing limits. There is variation in the instrument across groups because for different groups we draw from different empirical distribution functions. There is variation over time within

[27]Figure C.6 in the Online Appendix shows the actual fractions by risk score for 2007 and 2008. One can see that overall patterns are very similar. Figure C.7 also shows variation in the instrument across years for a given month.

the year because as the year progresses more and more simulated individuals will exceed the cost-sharing limit. Variation in the instrument across years is driven by variation in institutional rules.

When estimating the empirical distribution functions we deliberately pool across months and years. By pooling health care expenditures across months, the instrument is purged of differences in seasonal patterns across groups in the population. By pooling expenditures across years, we make sure that the variation in the instrumental variable across years is only driven by changes in institutional rules, and not by how individuals respond to these changes. That is, simulated average prices for the same group and month differ across years because of changing cost-sharing limits, but not because of patients' responses to changing cost-sharing limits or because of different responses to prices in years with a deductible and in years with a no-claim refund.

Formally, for the instrument z_{it} to be valid, we need it to be relevant and an exclusion restriction needs to hold. As for relevance, z_{it} is related to p_{it} by design. In Section 9.1 below, we document that the first stage F-statistic is very high. The exclusion restriction is that

$$\mathbb{E}\left[\varepsilon_{it}|z_{it},x_{it}\right] = \mathbb{E}\left[\varepsilon_{it}|x_{it}\right],$$

which means that conditional on year and month dummies, age, gender, risk score, future price in January, and expenditures in the previous 3 months, the simulated fraction of individuals without cost-sharing is mean independent of ε_{it}. This restriction is likely to hold because of the way we construct our instrument.[28]

This can be seen by re-visiting the three threats to identification we have discussed in Section 6.1: omitted variable bias, autocorrelation in error terms, and seasonal trends that differ between groups in the population. In contrast to individual prices p_{it}, our instrumental variable z_{it} only varies at the group level. It is therefore unrelated to omitted variables capturing individual differences within the group (recall that we control for differences across groups by controlling for x_{it}) and past values of the error term at the individual level.. Furthermore, by drawing from the empirical distribution function that was estimated from the pooled data our instrumental variables are unrelated to seasonal trends at the group level.

Throughout, reported standard errors are heteroskedasticity-robust and clustered at the level of groups defined by the risk score decile, three age categories, gender, and year. In our defini-

[28] Strictly speaking, a possible violation of the exclusion restriction comes from the fact that when estimating the empirical distribution function of monthly health care consumption, then we also use observations for individual i. Formally, the instrument is therefore a function of the error terms of individual i. However, the impact thereof will vanish in the limit, when the number of observations goes to infinity. We are not concerned about this because of the large number of observations. On average, there are about 170,000 observations for each group in our estimation sample and the smallest group has 44,388 observations. In principle, one could construct the instrument at the individual level by estimating the empirical distribution function using data from all other individuals, but this would substantially increase the computational burden (by a factor N, where N is the number of individuals). We have experimented with this for a small number of individuals and have found that it did not noticeably affect the estimates of the empirical distribution function. Therefore, we have not implemented this for the full sample.

tion of clusters we follow the level of variation in our instrumental variable.

The empirical approach we take in this paper is new in the context of studying the effects of patient cost-sharing. It combines the idea of exploiting variation within years and across groups, as suggested by Ellis et al. (2017), who use average prices for different health insurance plans as instruments, with the idea of constructing an instrument by combining draws from an empirical distribution function in combination with knowledge about variation in institutional rules, which was used in other settings for example by Currie and Gruber (1996a,b) and Cutler and Gruber (1996).

Our simulated instrumental variables approach can be seen as a generalization of a triple differences-in-differences grouping estimator. The first difference is within groups and year and across months; the second difference is across groups; and the third difference is across policy regimes. For illustration, suppose that there are only two groups and two points in time within the same year, and that for both groups the average price changes between the two points in time, but by a different amount. Then the instrumental variable estimator relates the difference in the change in spending between the two groups to the difference in the change in prices, also controlling for seasonal variation common to the two groups. This is akin to a differences-in-differences estimator within the year. The third difference comes from comparing estimates of price effects for years when the no-claim refund was in place to estimates for years when a deductible was in place.

7 Results

7.1 Main results

We now turn to our results. Our central hypothesis is that individuals are loss-averse and form their reference point in a way so that they react more strongly to financial incentives when they are framed in terms of losses. This means that we expect the coefficient on the current price interacted with an indicator for the deductible regime to be more negative than the coefficient on the current price interacted with an indicator for the no-claim regime.

Our main results are shown in Table 2. The dependent variable in column (1) is the log of expenditures plus one.[29] Therefore, reported coefficients are (approximately) percentage changes.[30] We find the price effect to be a 16.9 percent decrease in expenditures under a no-claim refund and a 38.7 percent decrease under a deductible. This implies that under a no-claim refund individuals spend 16.9 percent less on health care in months before hitting the no-claim refund limit than in months after hitting the no-claim refund limit. Correspondingly, under

[29] In Table C.4 below we show that adding 10 or 0.1, respectively, leads to different point estimates but the same qualitative findings.

[30] Recall that we use the log of expenditures plus 1 as the dependent variable. The exact percentage change of the dependent variable associated with a coefficient of -0.169 on the current price in the no-claim regime is $e^{-0.169} - 1 = -0.155$. In the following, we will abstract from such small differences.

Table 2: Main results

	(1) expenditure	(2) has claim
(current price) x (noclaim regime)	-0.169***	-0.0484***
	(0.0486)	(0.00924)
(current price) x (deductible regime)	-0.387***	-0.0735***
	(0.0245)	(0.00482)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	8,771,940	8,771,940
number of clusters	468	468
p-value equality current price coefficient	0.0000	0.0053

Notes: Instrumental variables estimates. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

a deductible individuals spend 38.7 percent less on health care in months before hitting the deductible limit than in months after hitting the deductible limit. While these estimates might seem large, our estimate for the effect of deductibles is similar and even somewhat smaller compared to the one of Brot-Goldberg et al. (2017), who find a price effect of deductibles on monthly health care expenditures of 42.2 percent for a sample of highly paid employees of a large American firm.

The last row of the table shows the p-value for a test of the null hypothesis that the two price effects are equal. This null hypothesis is rejected, in line with our main hypothesis that individuals react stronger to prices under a deductible policy than under a no-claim refund policy.

In column (2), the dependent variable is an indicator for having a claim. Hence, reported coefficients show price effects at the extensive margin. The price effect is a 4.8 percentage point decrease in the probability to have a claim in a given month under the no-claim regime and a 7.4 percentage point decrease under the deductible. The difference is statistically different from zero at the 5 percent level.

7.2 Effect heterogeneity

Next, we characterize effect heterogeneity. We do so to investigate whether the average effects we estimate are driven by particular groups or whether framing has an effect across all groups. For this we leverage the fact that our sample is for a broad sample of the population and that we observe background variables such as income in the neighborhood.

Our general approach to characterize effect heterogeneity is to split the sample.[31] We start with investigating whether price effects and the impact of framing differ by income. The price effect could differ between income groups if for example individuals with lower income react stronger to cost-sharing incentives than individuals with higher income, and the effect of framing could differ between income groups if they use different reference points. We divide individuals in 4 groups, according to the average income at the 6-digit neighborhood level, where quartile 1 refers to the group with the lowest income. A 6-digit postal code has on average 37 residents (footnote 17). Table 3 shows that results for each of the income groups are similar to the baseline results, and that there are no systematic differences across groups, with the exeption that individuals in the lowest income quartile react less strongly to cost-sharing incentives under the no-claim refund. The last row of the table shows average income by group. We can conclude from this analysis that individuals in poorer neighborhoods do not react more strongly to cost-sharing incentives than individuals in richer neighborhoods. This finding is in line with evidence from the RAND health insurance experiment, where the effect of deductibles on total health care expenditures did also not significantly differ across income groups (Manning

[31] Here and in the following we present results for the effect on expenditures. Results for the extensive margin (having at least one claim in a given month) are presented in Online Appendix C.

Table 3: Results by income

	(1) quartile 1	(2) quartile 2	(3) quartile 3	(4) quartile 4
(current price) x (noclaim regime)	-0.109**	-0.211***	-0.215***	-0.208***
	(0.0546)	(0.0486)	(0.0516)	(0.0569)
(current price) x (deductible regime)	-0.381***	-0.381***	-0.429***	-0.405***
	(0.0276)	(0.0275)	(0.0302)	(0.0333)
future price in January	Yes	Yes	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes	Yes	Yes
5 year age brackets	Yes	Yes	Yes	Yes
dummies decile risk score	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes
month dummies	Yes	Yes	Yes	Yes
number of observations	2,052,060	2,206,692	2,083,392	2,084,640
number of clusters	467	468	468	468
p-value equality current price coefficient	0.0000	0.0001	0.0000	0.0003
average income	1,439	1,847	2,278	3,174

Notes: Instrumental variables estimates for different subsamples defined by income quartile at the 6-digit neighborhood level. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $* \; p < 0.10$, $** \; p < 0.05$, $*** \; p < 0.01$.

Table 4: Results by gender and age

	(1) female	(2) male	(3) age 19-64	(4) age 65+
(current price) x (noclaim regime)	-0.234***	-0.0335	-0.206***	-0.129**
	(0.0521)	(0.0549)	(0.0524)	(0.0520)
(current price) x (deductible regime)	-0.456***	-0.282***	-0.303***	-0.350***
	(0.0356)	(0.0285)	(0.0282)	(0.0321)
future price in January	Yes	Yes	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes	Yes	Yes
5 year age brackets	Yes	Yes	Yes	Yes
dummies decile risk score	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes
month dummies	Yes	Yes	Yes	Yes
number of observations	4,548,720	4,223,220	6,797,688	1,974,252
number of clusters	234	234	360	175
p-value equality current price coefficient	0.0000	0.0000	0.0575	0.0000

Notes: Instrumental variables estimates by age and gender, respectively. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

et al., 1987). Furthermore, within each group the price effect is significantly stronger under a deductible scheme than under a no-claim refund scheme. Also, the difference between the price effects under a deductible and a no-claim refund is similar across income groups.

Table 4 shows results by age and gender. Individuals in all age and gender groups react stronger to deductibles than to no-claim refunds. Moreover, price responses are similar for persons above age 65 and for persons below age 65. This is an interesting result since many previous studies on the effects of patient cost-sharing examine either the working age population or people above age 65, e.g. Medicare beneficiaries. In the Netherlands, individuals buy the same type of health insurance below and above age 65, and insurance coverage does not change discontinuously at age 65, as it does for instance in the U.S. (Card et al., 2008). However, we find differences in price effects on expenditures by gender. Women react stronger to cost-sharing incentives than men, both under a deductible and a no-claim refund.

Next, we split the sample by risk score. The idea behind this is that we expect individuals with a higher risk score to react less strongly to price changes. About 80 percent of the individuals with an above-median risk score will exhaust the deductible or no-claim refund limit by the end of the year, while this is the case for only about 20-30 percent of the individuals with a below-median risk score. Therefore, one might expect below-median risk score individuals to react more strongly to hitting the cost-sharing limit, as it is *a priori* less likely that they do (see

Table 5: Results by risk score

	(1) low risk score	(2) high risk score
(current price) x (noclaim regime)	-0.541***	-0.111
	(0.0879)	(0.0708)
(current price) x (deductible regime)	-0.569***	-0.277***
	(0.0590)	(0.0562)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	4,359,216	4,412,724
number of clusters	198	270
p-value equality current price coefficient	0.6282	0.0002

Notes: Instrumental variables estimates by risk score. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Section 3.1). Table 5 is in line with this prediction. We find that the price effect for individuals with below-median risk scores is generally stronger than for individuals with above-median risk scores.

7.3 The effect of framing on annual health care expenditures

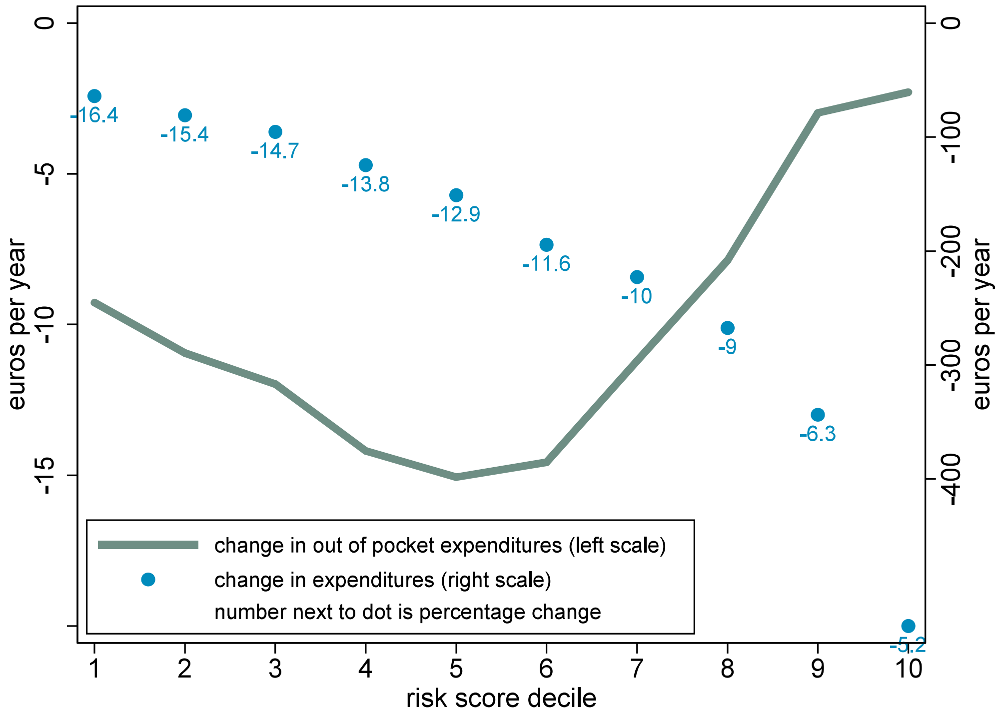
Our main results so far are that no-claim refunds and deductibles reduce monthly health care expenditures by 16.9 percent and 38.7 percent, respectively, in months in which patients are still subject to cost-sharing. Yet, these numbers are not indicative for the effects of no-claim refunds and deductibles on annual health care expenditures since cost-sharing incentives apply only for months in which cost-sharing limits are not yet exceeded. In this section we aim to compute the effect of no-claim refunds and deductibles on annual health care expenditures and on annual out-of-pocket spending.

For this, we conduct a simulation study. In a first step, we calculate the residuals and estimate their empirical distribution function. Then, we take our estimation equation and a current price of 1 at the beginning of January as a starting point and simulate, for each individual in our data set, health care expenditures in January by calculating the predicted mean of the log of health care expenditures plus one, according to equation (1), and adding a draw from the empirical distribution function of the residual. This is then transformed into actual health care expenditures by applying the exponential function and subtracting 1. Next, we calculate simulated prices for all individuals at the beginning of February, again predict the mean of the log of health care spending plus one given those prices, and so on. We do so for 2015 and both the case in which cost-sharing incentives are framed as a deductible, as it was the case in 2015, and for the case in which cost-sharing incentives are framed as a no-claim refund. For the latter, we use $\beta^{\text{no-claim refund}}$ instead of $\beta^{\text{deductible}}$.

Figure 5 shows the result. On average, annual health care expenditures are 8.6 percent lower when cost-sharing incentives are framed as a deductible than when they are framed as a no-claim refund. In absolute terms, annual health care expenditures decrease more for higher risk score deciles. In relative terms, annual care expenditures decrease more for lower risk-score deciles. The underlying mechanism is that low risk-score individuals tend to exceed cost-sharing limits later in the year or not at all, and they are thus subject to cost-sharing incentives for a longer period than high risk-score individuals.

Turning to the effect of the framing of cost-sharing incentives on out-of-pocket spending, we find that it follows a U-shaped pattern. The effect is lowest for individuals with a high risk score. For individuals who exceed the cost-sharing limit under both cost-sharing schemes there is no difference in out-of-pocket spending between the two schemes. The effect is also small for individuals with the lowest risk-scores. These individuals have low health care expenditures and low out-of-pocket spending under both schemes. The effect of framing on out-of-pocket spending is largest for individuals in risk-score deciles 5 and 6.

Figure 5: Effect of framing on yearly health care expenditures and out-of-pocket payments



Notes: This figure shows, for each risk score decile, the simulated effect of framing cost-sharing incentives as a deductible instead of a no-claim refund (dots). Numbers next to the dots are the percentage changes. Absolute values are given on the right axis. The solid line is the effect on out-of-pocket payments. The simulation is done for 2015.

Overall, it is striking that changing the framing of cost-sharing incentives leads to large effects on health care spending—a reduction of 207.97 euros per year on average—, while having very modest effects—a reduction of 10.06 euros on average—on out-of-pocket expenditures.

8 Potential alternative mechanisms

In the previous section we have shown that individuals react more strongly to financial incentives in years in which they are framed as a deductible than in years in which they are framed as a no-claim refund. We have pointed out in Section 6.1 that this is a before-after design. In order to interpret the difference between the two as the pure effect of framing we have to exclude other potential explanations for stronger reactions to cost-sharing incentives in years in which a deductible was in place. Therefore, we now discuss alternative mechanisms that can potentially explain our findings. These are time discounting, liquidity constraints, differences in information, and salience. The discussion will lead us to conclude that none of them is likely to explain our findings, and therefore loss aversion is our preferred explanation.

8.1 Time discounting

Deductibles and no-claim refunds differ not only in the framing of cost-sharing incentives, but also in the timing of payments. In general, deductibles have to be paid several weeks or months after the treatment, whereas the no-claim refund is only paid out at the end of the first quarter of the following year. This raises the question whether this difference in the timing of payments could explain our results.

In this section, we present empirical evidence that strongly speaks against this alternative explanation. In order to examine whether different responses can be explained by differences in the timing of payments, we re-do our analysis and take the delay in payments into account. For this, we assume a high, but still plausible annual discount rate of 10 percent, and we allow for compounding within the year. We further assume a delay of 3.5 months under the deductible regime, corresponding to the deductible payment being due at the end of the third month after the treatment. For the no-claim refund, we assume that the delay is 3.5 months in December, 4.5 months in November, and so on, because the no-claim refund is paid out at the end of the first quarter in the following year. Making these assumptions allows us to re-compute current prices and thus taking the difference in timing into account. In particular, we set the current price to $(1/(1+0.1))^{m/12}$ for payments m months in the future.

Results are presented in Table 6. They are very similar to our main results presented in Table 2. Thus, discounting payments at a 10 percent rate is not enough to explain our findings.

In addition, we use a stylized model to compute which discount rate would be high enough to explain the difference in effects between deductibles and no-claim refunds that we find. Online Appendix B contains details of our calculations. We find that the annual discount rate

Table 6: Specification with discounted prices

	(1) expenditure	(2) has claim
(discounted current price) x (deductible regime)	-0.186***	-0.0535***
	(0.0538)	(0.0102)
(discounted current price) x (noclaim regime)	-0.397***	-0.0753***
	(0.0252)	(0.00495)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	8,771,940	8,771,940
number of clusters	468	468
p-value equality current price coefficient	0.0000	0.0274

Notes: Instrumental variables estimates. We take the difference in timing into account by using the discounted current price instead of the actual current price. We use an annual discount rate of 10 percent. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

would need to be higher than 400% to explain our results. We consider this discount rate to be implausibly large.

8.2 Liquidity constraints

One possible explanation for this very high discount rate could be liquidity constraints. However, in our context, we believe that for institutional reasons liquidity constraints are unlikely to explain our results. The maximal amount an individual had to pay was 375 euros per year in 2015 (see Section 2), which is an amount that most Dutch persons can raise. In case this nevertheless proves difficult, individuals have the possibility to arrange payment in installments with the health insurer. Moreover, as explained in Section 2.3, usually the payment is not immediately due at the time of treatment since the the treatment and the billing dates (to the insurance) do not coincide. Figure C.1 in the Online Appendix shows that the average delay between the treatment and billing date is about 60 days, and there will be an additional delay before the insurance will send an invoice to the individual.

Evidence that liquidity constrains are unlikely to explain the different effect of deductibles and no-claim refunds also comes from our analysis for different income groups at the 6-digit postal code level. We would expect that liquidity constraints are more common for individuals with lower incomes and therefore, if these were the main driver of our results, then we would expect effects to be stronger for individuals with lower incomes. However, in Table 3 we do not find that the price effects differ between income groups.

8.3 The role of information

Another possible explanation why individuals react stronger to deductibles than to no-claim refunds is that they are better informed about economic incentives under a deductible policy than under a no-claim refund policy. This can be the case if they are either better informed about the policy rules or better informed about the price of care at a given point of time under a deductible policy. Yet, as discussed in Section 2.4 knowledge about both no-claim refunds and deductibles was almost universal in the Dutch population throughout our study period with the exception of the year 2008 immediately after the introduction of the deductible. Also, the way people were informed about whether they were above or below the cost sharing limit was similar under the no-claim refund policy and the deductible policy.

8.4 Salience

A further possible explanation for different responses to no-claim refunds and deductibles could be related to the salience of cost-sharing incentives. Salience can be defined as "the phenomenon that when one's attention is differentially directed to one portion of the environment rather than to others, the information contained in that portion will receive disproportionate weighting in subsequent judgements." (Taylor and Thompson, 1982).

In the context of cost-sharing, Dalton et al. (ming) find that strong responses to spot prices, as opposed to expected end-of-year prices, can be attributed to the greater salience of spot prices. In contrast, we compare the response to spot prices under a deductible policy to the response to spot prices under a no-claim refund.

Our results could be explained by salience if more attention is drawn to cost-sharing incentives under a deductible than under a no-claim refund. However, it is not clear why this should be the case in the Dutch institutional setting. Neither the no-claim refund nor the deductible result in immediate payments at the time of treatment. Also, the information individuals receive about cost-sharing incentives is comparable between the two cost-sharing schemes (see Section 2.4). Deductibles could be said to be more salient than no-claim refunds if individuals pay more attention to losses than to gains. But this would redefine the concept of loss aversion as a special case of salience.

In summary, while we cannot completely rule out other mechanisms, we find that loss aversion is the most convincing and plausible explanation for our results.

9 Validity of instruments and sensitivity analysis

In this section, we first assess whether the instruments we use are relevant. Then, we present results from a placebo test in order to assess the credibility of our empirical approach. After this, we present results for a number of robustness checks for alternative specifications of explanatory variables and outcome variables, and we examine whether our results are sensitive to changes in the sample composition and the basic health insurance plan over time. Here, we also ask whether end-of-the-year effects could explain our results. All tables are presented in the Online Appendix.

9.1 Instrument validity

One of the conditions for an instrument to be valid is that the instrument must be related to the endogenous variable conditional on controls. Table C.2 shows that the simulated average price at the group level is highly predictive of the individual price p_{it}. The relevant F-statistics are extremely high. They are 1,216 and 8,046, respectively.

Turning to the exclusion restriction, one way to assess the credibility of our empirical approach is to perform a placebo test. In the Netherlands children below the age of 18 are not subject to patient cost-sharing. For this reason, our main analysis was conducted for a sample of individuals at ages 19 and above.[32] If we apply our empirical approach to a sample of individuals between ages 15 and 17 and code the price as if they were actually subject to cost

[32] We did not include 18 year old individuals because we only observe the birth year and not the month of birth.

sharing, then we should not find any effect. And indeed, Table C.3 shows that price effects are not significantly different from zero for this group and also not significantly different from one another. Thus, our empirical approach performs well in a setting where we know that the true parameters are zero.

9.2 Sensitivity analysis

Next, we conducted a number of additional robustness checks. We first examine whether or not our estimation results are sensitive to different specifications of the covariates in x_{it}. These results are shown in Table C.4. Recall that in our baseline specifications x_{it} includes variables for 5 year age brackets, gender, year and month dummies, indicators for the decile of the risk score, the future price in January, and the log of expenditure in the previous 3 months plus 1. In the first column of Table C.4 we use a specification with a full set of risk score-year interactions and a full set of year-month interactions. The results are similar to the baseline results presented in Table 2, suggesting that a specification that is not fully-interacted in this way is general enough. For the second column we include 6 lags of the log of the respective monthly expenditure plus 1 instead of the log of the expenditure in the previous three months plus 1. Results are qualitatively similar as well. In the last column, we use the baseline specification but do not control for the expected price at the end of the year. This again leads to qualitatively similar results, but now the magnitude of the estimated effects is higher. Thus, we find that our main finding that the price effect is more negative under a deductible than under a no-claim refund is robust to using different specifications of explanatory variables.

Next, we examine whether or not our estimation results are sensitive to the specification of the outcome variable. Recall that up to now, we have used the log of expenditures plus 1. If we instead use the log of expenditures plus 0.1 or plus 10, then coefficients will be rescaled.[33] Table C.5 shows that indeed, adding 0.1 and 10 instead of 1 leads to different estimates—as could be expected. But importantly, the main finding remains unchanged: the price effect is more negative under a deductible than under a no-claim refund. Table C.5 also shows that estimation results remain essentially unchanged when we change the definition of health care expenditure and exclude expenditures on mental health care and geriatric care. These were included in the basic health insurance plan only in later years.

Next, we investigate whether results are different for different definitions of our sample. Results are shown in Table C.6. In column (1) we restrict the sample period to one year before and after the change from a no-claim refund to a deductible, by using data for the years 2007

[33]For a simple example suppose that there is a claim with probability 0.4 under no cost-sharing and that the claim size is 50 conditional on there being a claim. Assume that the price effect of cost-sharing is a reduction in the probability that there is a claim by 10 percentage points and a reduction in the claim size by 50 percent once there is a claim. Then, the difference between the expected log of expenditure plus 1 under no cost-sharing and the expected log of expenditure plus 1 under cost-sharing is -0.33. If we use respectively 0.1 and 10 instead of 1, then we arrive at -0.42 and -0.23. This shows that the constant one adds will matter. The estimated effect will tentatively be smaller in magnitude if we add a bigger number, at least in this example.

and 2009 only. We have excluded data from the year 2008, because in 2008 we would use data from 2007 to construct control variables. By excluding data from 2008, we have a cleaner separation between the period in which the no-claim refund was in place and the period in which there was a deductible.[34] Results are similar to our main results for all years. In column (2) we show results for a balanced sample of individuals who are in our sample in all years from 2006 until 2015. In this way, we assess whether our results are sensitive to changes in the sample composition over time during our study period. Results for the balanced sample are similar to the baseline results for the full sample.

Our final robustness check concerns the shifting of health care expenditures across years (Cabral, 2016). For some types of health care, individuals can influence the timing of treatment. For example, they can choose whether they will receive a cataract surgery now or some months later. Individuals who have exceeded their cost-sharing limit in the current year have an incentive to shift treatments from the next year to the current year, because at the start of the new year patient cost-sharing limits will be reset. For those patients, the price of care in the current year is zero, but the price of care at the beginning of the next year is one. To see whether the shifting of care between years has an impact on our estimation results we estimate a model that in addition to the other covariates also includes an interaction term between the price p_{it} an an indicator for the month of December. The results presented in the first column of Table C.7 show that this does not substantially change the coefficients for the price effects under the two cost-sharing schemes. Interestingly, however, individuals do seem to be more sensitive to the price of care in December than in other months, as one would expect. In the second column we do not use data for December and find results that are very similar to the ones presented in the first column and to our baseline estimates.

10 Conclusions

In this study we compare the effects of patient cost-sharing incentives on health care expenditures under a deductible policy and a no-claim refund policy. We estimate that average annual health care expenditures are about 8.6 percent lower under a deductible policy than under a no-claims refund policy, as patients react more strongly to deductibles than to no-claim refunds.

Our preferred explanation for these results is that individuals are loss-averse and respond differently to both schemes since they perceive a deductible payment as a loss and a no-claim refund as a gain. We also discuss alternative explanations such as the different timing of payments under deductibles and no-claim refunds, liquidity constraints, the role of information and salience. Based on additional empirical evidence and in light of the institutional details we conclude that these alternative explanations are unlikely to explain our findings.

Our finding that the framing of incentives can have a large and economically important effect

[34] Figure 3 above was for 2007 and 2008. Figure C.2 in the Online Appendix is for 2007 and 2009 and corresponds more closely to the specification we use here.

on health care expenditures is directly relevant for designing patient cost-sharing schemes. For example, some Dutch health insurers have recently started to offer insurance contracts that allow insurees to pre-pay their deductibles in monthly installments. For these contracts, individuals receive a refund at the end of the year if they have not exceeded their deductible limit, effectively turning the deductible into a no-claim refund. Within a short period around 800,000 insurees have signed up for this scheme. If patients who pay their deductible in monthly installments behave similar to patients under a no-claim refund then our simulation results predict that total annual health care expenditures are around 8.6 percent higher than under a standard deductible. In their calculation of insurance premiums Dutch health insurers do not take this behavioral effect into account. Thus, offering the seemingly innocent option to pre-pay deductibles in monthly installments might cause substantial losses for insurers and lead to an increase in health care costs.

More generally, our results imply that discussions on the optimal design of cost-sharing incentives should not only be centered on coinsurance rates and cost-sharing limits, but they should also revolve around the question how these are presented to patients.

References

Abaluck, J., J. Gruber, and A. Swanson (2018). Prescription drug use under Medicare Part D: A linear model of nonlinear budget sets. *Journal of Public Economics 164*, 106–138.

Brot-Goldberg, Z. C., A. Chandra, B. R. Handel, and J. T. Kolstad (2017). What does a deductible do? The impact of cost-sharing on health care prices, quantities, and spending dynamics. *Quarterly Journal of Economics 132*(3), 1261–1318.

Cabral, M. (2016). Claim timing and ex post adverse selection. *Review of Economic Studies 84*(1), 1–44.

Card, D., C. Dobkin, and N. Maestas (2008). The impact of nearly universal insurance coverage on health care utilization: Evidence from Medicare. *American Economic Review 98*(5), 2242–2258.

Chandra, A., J. Gruber, and R. McKnight (2010). Patient cost-sharing and hospitalization offsets in the elderly. *American Economic Review 100*(1), 193.

Chandra, A., B. Handel, and J. Schwartzstein (2019). Behavioral economics and health-care markets. In B. D. Bernheim, S. DellaVigna, and D. Laibson (Eds.), *Handbook of Behavioral Economics - Foundations and Applications*, Volume 2. Elsevier.

Currie, J. and J. Gruber (1996a). Health insurance eligibility, utilization of medical care, and child health. *Quarterly Journal of Economics 111*(2), 431–466.

Currie, J. and J. Gruber (1996b). Saving babies: the efficacy and cost of recent changes in the medicaid eligibility of pregnant women. *Journal of Political Economy 104*(6), 1263–1296.

Cutler, D. M. and J. Gruber (1996). Does public insurance crowd out private insurance? *Quarterly Journal of Economics 111*(2), 391–430.

Dalton, C. M., G. Gowrisankaran, and R. Town (forthcoming). Salience, myopia, and complex dynamic incentives: Evidence from Medicare Part D. *Review of Economic Studies*.

Ecorys (2011). Evaluatie naar het verplicht eigen risico. Ecorys, Rotterdam, The Netherlands.

Einav, L., A. Finkelstein, S. P. Ryan, P. Schrimpf, and M. R. Cullen (2013). Selection on moral hazard in health insurance. *American Economic Review 103*(1), 178–219.

Einav, L., A. Finkelstein, and P. Schrimpf (2015). The response of drug expenditure to nonlinear contract design: Evidence from Medicare Part D. *Quarterly Journal of Economics 130*(2), 841–899.

Ellis, R. P. (1986). Rational behavior in the presence of coverage ceilings and deductibles. *RAND Journal of Economics 17*(2), 158–175.

Ellis, R. P., B. Martins, and W. Zhu (2017). Health care demand elasticities by type of service. *Journal of Health Economics 55*, 232–243.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, K. Baicker, Oregon Health Study Group, and others (2012). The Oregon health insurance experiment: Evidence from the first year. *Quarterly Journal of Economics 127*(3), 1057–

1106.

Goudriaan, R., R. Lalloesingh, and P. Vermer (2007). Eindevaluatie van de no-claimteruggaveregeling: Onderzoek in opdracht van het ministerie van VWS. APE Consulting, Den Haag, The Netherlands.

Hasaart, F. (2011). *Incentives in the Diagnosis Treatment Combination payment system for specialist medical care: A study about behavioral responses of medical specialists and hospitals in the Netherlands.* Ph. D. thesis, Universiteit Maastricht, Maastricht, The Netherlands.

Hayen, A. P., M. J. van den Berg, B. R. Meijboom, J. N. Struijs, and G. P. Westert (2015). Incorporating shared savings programs into primary care: From theory to practice. *BMC Health Services Research 15*(580), 1–15.

Johnson, E. J., J. Hershey, J. Meszaros, and H. Kunreuther (1993). Framing, probability distortions, and insurance decisions. *Journal of Risk and Uncertainty 7*(1), 35–51.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*(2), 263–292.

Keeler, E. B., J. P. Newhouse, and C. E. Phelps (1977). Deductibles and the demand for medical care services: The theory of a consumer facing a variable price schedule under uncertainty. *Econometrica 45*(3), 641–655.

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

Krabbe-Alkemade, Y., T. Groot, and M. Lindeboom (2016). Competition in the Dutch hospital sector: An analysis of health care volume and cost. *European Journal of Health Economics 18*(2), 1–15.

Lambregts, T. R. and R. C. van Vliet (2017). The impact of copayments on mental healthcare utilization: A natural experiment. *European Journal of Health Economics 19*(6), 1–10.

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

Netherlands Institute for Social Research (2006). Statusscores postcodegebieden. https://www.scp.nl/Onderzoek/Lopend_onderzoek/A_Z_alle_lopende_onderzoeken/Statusscores. The Hague, The Netherlands.

NZA (2012). Marktscan zorgverzekeringsmarkt 2012. Dutch Healthcare Authority NZA, Utrecht, The Netherlands.

OECD Indicators (2015). Health at a glance 2015. OECD Publishing. Paris, France.

Reitsma-van Rooijen, M. and J. d. Jong (2009). Compensatie eigen risico is nog onbekend. *Zorg en Financiering 8*(6), 34.

Remmerswaal, M., J. Boone, M. Bijlsma, R. Douven, et al. (2017). Cost-sharing design matters: A comparison of the rebate and deductible in healthcare. CPB Discussion Paper 367, CPB

Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands.

Remmerswaal, M., R. Douven, and P. Besseling (2015). Toelichting op de effecten van eigen betalingen in de zorg. CPB Achtergronddocument, CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands.

Schäfer, W., M. Kroneman, W. Boerma, M. Berg, G. Westert, W. Devillé, E. v. Ginneken, et al. (2010, 1). The Netherlands: Health system review. *Health Systems in Transition 12*(1). European Observatory on Health Systems and Policies.

Statistics Netherlands (2012). Inhoud kerncijfers postcodegebieden 2008-2010. https://www.cbs.nl/nl-nl/maatwerk/2012/01/kerncijfers-postcodegebieden-2008-2010, CBS Statistics Netherlands, The Hague, The Netherlands.

Taylor, S. and S. Thompson (1982). Stalking the elusive vividness effect. *Psychological Review 89*, 155–181.

Van de Ven, W. P. and F. T. Schut (2009). Managed competition in the Netherlands: Still work-in-progress. *Health Economics 18*(3), 253–255.

Van der Maat, M. and J. de Jong (2010). Eigen risico in de zorgverzekering: Het verzekerdenperspectief. Nivel, Utrecht, The Netherlands.

van Winssen, K. P. M., R. C. van Kleef, and W. P. van de Ven (2016). Potential determinants of deductible uptake in health insurance: How to increase uptake in the Netherlands? *European Journal of Health Economics 17*(9), 1059–1072.

World Health Organization (2006). The anatomical therapeutic chemical classification system with defined daily doses (atc/ddd). http://www.who.int/classifications/atcddd/en/.

Online Appendix

This Online Appendix contains additional details on the data, the derivation of the implied discount rate, and additional tables and figures.

A Additional details on the data

A.1 Services that are subject to cost-sharing

Constructing our main independent variables and the price requires that we know which services are subject to cost-sharing, as some services are exempted from cost-sharing. Such a list was not available to us. We therefore proxied for this using data on actual deductible payments made and information on the overarching care types that are exempt from the deductible, as specified by law. The types of care that are exempted include obstetric and maternal care, care generally provided by General Practitioners (GPs) and participation under a chronic care program.

Specifically, our approach for determining coverage consisted of the following steps (in chronological order):

1. Every specialist service was defined as counting towards the deductible;
2. Services provided by registered obstetricians were defined as not counting towards the deductible;
3. Specialist services were not defined as counting towards the deductible when they were received only by women (90%, to account for administrative errors) and,
   (a) for which in fewer than 1% of the claims a deductible payment was requested, for services that we observed over a 1000 times in the data or
   (b) for which in fewer than 10% of the claims a deductible payment was requested for services that observed less than a 100 times in the data;
4. For any remaining services, we determined coverage along the lines of 3(a) and 3(b). That is, we let coverage fully depend on the percentage of claims in which actual deductible payments were made;
5. Coverage under the no-claim regime was deduced from 2008 deductible data.

Using the percentage cut-off points in 3(a) and 3(b), we take account of administrative errors. We used a more conservative cut-off points for services that were billed less frequently.

A.2 Risk scores

To calculate the risk score, we regress total health care costs for a given year on a gender dummy, fully interacted with a third order polynomial in age, indicators for the decile of costs in the previous year, indicators for chronic conditions, indicators for characteristics at the 6-digit neighborhood level. For the latter we conduct a median split for income, the fraction non-western immigrants, and the CBS socioeconomic status score, respectively. We also include dummies for missing characteristics at the neighborhood level. Finally, we obtain fitted value and divide them by the average. Figure C.4 shows the risk score deciles.

B Implied discount rate

In our empirical analysis, we have obtained estimates of price effects, separately for the no-claim and deductible regime. Our interpretation of the difference between the estimated effects is that the difference arises due to framing. In principle, an alternative explanation is that individuals react stronger to deductibles because then they have to pay earlier, as the no-claim refund is only paid out at the beginning of the following year. Here, we conduct a simple back-of-the envelope calculation with the goal to determine what the individual discount factor must be in order to generate this effect. We find that it would have to be unreasonable low in order to generate this effect, supporting the interpretation of the effects as being due to framing.

For this, we use a framework with a quadratic utility function as in Einav et al. (2013). Denote health care needs in month t by λ_t, health care consumption by m_t, income by y_t, and the implications of cost-sharing by $\delta^\tau p_t$. δ is the monthly discount factor so that δ^τ is the discount factor associated with payments τ months later, when the "current price" is p_t. The flow utility in period t depends on health care consumption,

$$u(m_t) = (m_t - \lambda_t) - \frac{1}{2\omega}(m_t - \lambda)^2 + y - \delta^\tau p_t m_t.$$

Differentiating with respect to m_t allows us to solve for the optimal consumption in t,

$$m_t^*(p_t) = \lambda_t + \omega(1 - \delta^\tau p_t).$$

When individuals have to fully and immediately pay out-of-pocket, we have $\delta^\tau p_t = 1$. Then, optimal health care consumption is equal to health care needs. In contrast, after hitting the cost-sharing limit, $\delta^\tau p_t = 0$ so that health care consumption is equal to health care needs plus ω. For that reason, ω can be interpreted as *ex post* moral hazard.

The effects reported in this paper are percentage changes in health care consumption, where the base is consumption when the current price is equal to zero. In terms of this model, this is

$$\frac{m_t^*(1) - m_t^*(0)}{m_t^*(0)} = \frac{-\delta^\tau \omega}{\lambda_t + \omega}.$$

Different regimes have different discount factors because they have different delays for the payments. For a rough approximation we assume that the average treatment date is after the first half of the year, payment under the deductible regime occurs on average three months later, and payment under the payback rebate occurs at the end of the first quarter of the following year. This means that the difference in timing is 6 months on average.

Assuming that health care needs and moral hazard effects stay the same, the ratio of coefficients obtained in the main analysis is

$$\frac{-0.169}{-0.387} = \frac{\delta^9}{\delta^3},$$

so that the yearly discount factor would have to be

$$\delta^{12} = 0.191.$$

This corresponds to the arguably unrealistically high discount rate of

$$\frac{1}{0.191} - 1 = 424\%.$$

C Additional Tables and Figures

This appendix contains additional tables and figures. We refer to most of them in the main text.

Table C.1 provides additional information on care consumption and the fraction of individuals reaching the cost-sharing limit per year.

Tables C.2 through C.7 are referred to and discussed in Section 9.

Throughout, we focus on the question how the effects of cost-sharing on total expenditures are influenced by framing. Tables C.8 through C.13 contain results for having a claim in a given month as dependent variable.

Figure C.1 shows the delay between treatment and billing date.

Figure C.2 resembles Figure 3, but is for the years 2007 and 2009, instead of 2007 and 2008.

Figure C.3 shows how care consumption evolves within the years.

Figure C.4 shows risk score deciles.

Figure C.5 shows how the average price of care evolves over the course of the year for the years in our study period. The horizontal axis shows months of the year. The vertical axis shows the fraction of individuals whose care consumption was still below the cost-sharing limits at the beginning of the month. These are the individuals for whom the price of care is one. By construction, the average price of care is monotonically decreasing over the course of the year. Level differences between curves are driven by a combination of differences in medical spending across years and differences in no-claim refund limits and deductible limits across years. One can see that by the end of each year about half of the individuals have exhausted the respective cost-sharing limit.

Figure C.6 shows the fraction of individuals that is below the cost-sharing limit in a given month, by risk score quartile and pooled across years.

Figure C.7 shows that the instrument varies both within and across years for a given month and a given risk score quartile.

Table C.1: Summary statistics over time

	mean
care consumption	
2006	1295.92
2007	1478.37
2008	1901.19
2009	2071.94
2010	2144.54
2011	2320.20
2012	2244.82
2013	2505.05
2014	2650.63
2015	2772.96
hit cost-sharing limit	
2006	0.50
2007	0.54
2008	0.63
2009	0.64
2010	0.65
2011	0.65
2012	0.62
2013	0.57
2014	0.56
2015	0.56
number individuals	35,694

Notes: This table shows summary statistics for all individuals who were in the sample in January 2006. Care consumption is the consumption of care that falls under the no-claim refund policy or the deductible. Hit deductible is one for a person in a given year if that person exceeded the deductible or the payback limit in a given year.

Table C.2: First stage

	(1) current price no claim	(2) current price deductible
simulated fraction group not hit		
interacted with noclaim regime	0.874***	-0.0253**
	(0.0179)	(0.0102)
interacted with deductible regime	-0.00274	0.860***
	(0.00252)	(0.00947)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	8,771,940	8,771,940
number of clusters	468	468
F-statistic excluded instruments	1215.93	8046.24

Notes: First stage estimates. We use data at the monthly level. The dependent variable in the first column is the individual price interacted with an indicator for the no claim regime, and the dependent variable in the second column is the individual price interacted with an indicator for the deductible regime. Standard errors are clustered at the risk score decile-gender-age group-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table C.3: Placebo test: no effect for individuals age 15-17

	(1) expenditure	(2) has claim
(current price) x (noclaim regime)	0.117	-0.0294
	(0.141)	(0.0560)
(current price) x (deductible regime)	0.113	0.0180
	(0.0780)	(0.0245)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	359,724	359,724
number of clusters	126	126
p-value equality current price coefficient	0.9741	0.3620

Notes: Instrumental variables estimates for individuals who are between 15 and 17 years old. We compute risk scores and instruments separately for this group. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table C.4: Alternative specifications explanatory variables

	(1) flexible form	(2) distributed lag	(3) no future price
(current price) x (noclaim regime)	-0.188**	-0.198***	-0.563***
	(0.0799)	(0.0265)	(0.0660)
(current price) x (deductible regime)	-0.407***	-0.311***	-0.726***
	(0.0242)	(0.0204)	(0.0380)
future price in January	Yes	Yes	No
log exp. previous 3 months plus 1	Yes	No	Yes
5 year age brackets	Yes	Yes	Yes
dummies decile risk score	No	Yes	Yes
year dummies	No	Yes	Yes
month dummies	No	Yes	Yes
6 distributed lags log expenditure	No	Yes	No
risk score-year dummies	Yes	No	No
year-month dummies	Yes	No	No
number of observations	8,771,940	8,771,940	8,771,940
number of clusters	468	468	468
p-value equality current price coefficient	0.0128	0.0000	0.0056

Notes: Instrumental variables estimates. This table shows results using alternative specifications of the explanatory variables. In the first column we use a more flexible specification with risk score-year dummies and year-month dummies. The next column uses a distributed lag structure instead of the log of the expenditure in the previous 3 months. In the last column we do not condition on the future price in January. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table C.5: Alternative specifications outcome variable

	(1) log(exp.+0.1)	(2) log(exp.+10)	(3) log(exp.) without mental and geriatric care	(4) has claim without mental and geriatric care	(5) log(exp.) at median price
(current price) x (noclaim regime)	-0.280***	-0.0575*	-0.219***	-0.0501***	-0.182***
	(0.0679)	(0.0307)	(0.0520)	(0.00944)	(0.0527)
(current price) x (deductible regime)	-0.556***	-0.221***	-0.416***	-0.0758***	-0.460***
	(0.0343)	(0.0159)	(0.0259)	(0.00495)	(0.0263)
future price in January	Yes	Yes	Yes	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes	Yes	Yes	Yes
5 year age brackets	Yes	Yes	Yes	Yes	Yes
dummies decile risk score	Yes	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes	Yes
month dummies	Yes	Yes	Yes	Yes	Yes
number of observations	8,771,940	8,771,940	8,771,940	8,771,940	8,771,940
number of clusters	468	468	468	468	468
p-value equality current price coefficient	0.0000	0.0000	0.0001	0.0050	0.0000

Notes: Instrumental variables estimates. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

Table C.6: Alternative samples

	(1) only 2007 and 2009	(2) balanced sample
(current price) x (noclaim regime)	-0.281***	-0.0734
	(0.0577)	(0.0466)
(current price) x (deductible regime)	-0.406***	-0.360***
	(0.0531)	(0.0230)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	1,787,376	3,854,952
number of clusters	104	468
p-value equality current price coefficient	0.0412	0.0000

Notes: Instrumental variables estimates. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C.7: End-of-the-year effects

	(1) December interaction	(2) drop December
(current price) x (noclaim regime)	-0.160***	-0.172***
	(0.0490)	(0.0512)
(current price) x (deductible regime)	-0.378***	-0.374***
	(0.0250)	(0.0252)
(current price) x (December)	-0.0786***	
	(0.0158)	
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	8,771,940	8,040,945
number of clusters	468	468
p-value equality current price coefficient	0.0000	0.0000

Notes: Instrumental variables estimates. We use data at the monthly level. Expenditures are measured as the log of one plus the actual amount. Standard errors are clustered at the risk score decile-gender-age group-year level. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C.8: Results for has claim by income

	(1) quartile 1	(2) quartile 2	(3) quartile 3	(4) quartile 4
(current price) x (noclaim regime)	-0.0426***	-0.0551***	-0.0571***	-0.0494***
	(0.0102)	(0.00992)	(0.0109)	(0.0119)
(current price) x (deductible regime)	-0.0676***	-0.0763***	-0.0866***	-0.0836***
	(0.00574)	(0.00576)	(0.00625)	(0.00657)
future price in January	Yes	Yes	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes	Yes	Yes
5 year age brackets	Yes	Yes	Yes	Yes
dummies decile risk score	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes
month dummies	Yes	Yes	Yes	Yes
number of observations	2,052,060	2,206,692	2,083,392	2,084,640
number of clusters	467	468	468	468
p-value equality current price coefficient	0.0131	0.0234	0.0046	0.0026
average income	1,439	1,847	2,278	3,174

Notes: Instrumental variables estimates for different subsamples defined by income quartile at the 6-digit neighborhood level. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C.9: Results for has claim by gender and age

	(1) female	(2) male	(3) age 19-64	(4) age 65+
(current price) x (noclaim regime)	-0.0479***	-0.0206**	-0.0536***	-0.0462***
	(0.00937)	(0.00963)	(0.0112)	(0.0114)
(current price) x (deductible regime)	-0.0663***	-0.0540***	-0.0715***	-0.0538***
	(0.00668)	(0.00475)	(0.00591)	(0.00626)
future price in January	Yes	Yes	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes	Yes	Yes
5 year age brackets	Yes	Yes	Yes	Yes
dummies decile risk score	Yes	Yes	Yes	Yes
year dummies	Yes	Yes	Yes	Yes
month dummies	Yes	Yes	Yes	Yes
number of observations	4,548,720	4,223,220	6,797,688	1,974,252
number of clusters	234	234	360	175
p-value equality current price coefficient	0.0202	0.0003	0.0943	0.4910

Notes: Instrumental variables estimates by age and gender, respectively. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table C.10: Results for has claim by risk score

	(1) low risk score	(2) high risk score
(current price) x (noclaim regime)	-0.183***	-0.0362***
	(0.0290)	(0.0137)
(current price) x (deductible regime)	-0.159***	-0.0564***
	(0.0183)	(0.0115)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	4,359,216	4,412,724
number of clusters	198	270
p-value equality current price coefficient	0.2118	0.0262

Notes: Instrumental variables estimates by risk score. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C.11: Results for has claim and alternative specifications explanatory variables

	(1) flexible form	(2) distributed lag	(3) no future price
(current price) x (noclaim regime)	-0.0490***	-0.0520***	-0.146***
	(0.0132)	(0.00693)	(0.0148)
(current price) x (deductible regime)	-0.0789***	-0.0577***	-0.157***
	(0.00471)	(0.00460)	(0.00871)
future price in January	Yes	Yes	No
log exp. previous 3 months plus 1	Yes	No	Yes
5 year age brackets	Yes	Yes	Yes
dummies decile risk score	No	Yes	Yes
year dummies	No	Yes	Yes
month dummies	No	Yes	Yes
6 distributed lags log expenditure	No	Yes	No
risk score-year dummies	Yes	No	No
year-month dummies	Yes	No	No
number of observations	8,771,940	8,771,940	8,771,940
number of clusters	468	468	468
p-value equality current price coefficient	0.0375	0.3624	0.3910

Notes: Instrumental variables estimates. This table shows results using alternative specifications of the explanatory variables. In the first column we use a more flexible specification with risk score-year dummies and year-month dummies. The next column uses a distributed lag structure instead of the log of the expenditure in the previous 3 months. In the last column we do not condition on the future price in January. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table C.12: Results for has claim and additional robustness checks

	(1) only 2007 and 2009	(2) balanced sample
(current price) x (noclaim regime)	-0.0597***	-0.0386***
	(0.0107)	(0.00910)
(current price) x (deductible regime)	-0.0875***	-0.0691***
	(0.0102)	(0.00461)
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	1,787,376	3,854,952
number of clusters	104	468
p-value equality current price coefficient	0.0076	0.0006

Notes: Instrumental variables estimates. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

Table C.13: End-of-the-year effects for has claim

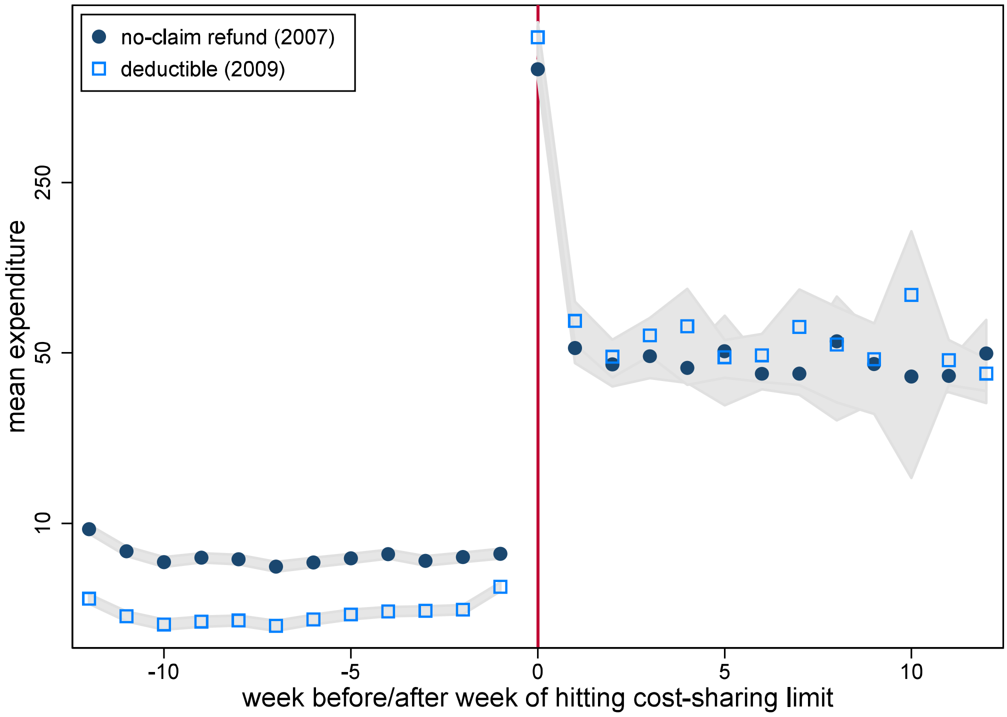
	(1) December interaction	(2) drop December
(current price) x (noclaim regime)	-0.0467***	-0.0459***
	(0.00929)	(0.00967)
(current price) x (deductible regime)	-0.0718***	-0.0685***
	(0.00485)	(0.00473)
(current price) x (December)	-0.0144***	
	(0.00262)	
future price in January	Yes	Yes
log exp. previous 3 months plus 1	Yes	Yes
5 year age brackets	Yes	Yes
dummies decile risk score	Yes	Yes
year dummies	Yes	Yes
month dummies	Yes	Yes
number of observations	8,771,940	8,040,945
number of clusters	468	468
p-value equality current price coefficient	0.0054	0.0179

Notes: Instrumental variables estimates. We use data at the monthly level. Standard errors are clustered at the risk score decile-gender-age group-year level. $*\ p < 0.10$, $**\ p < 0.05$, $***\ p < 0.01$.

Figure C.1: Average delay between treatment and billing date

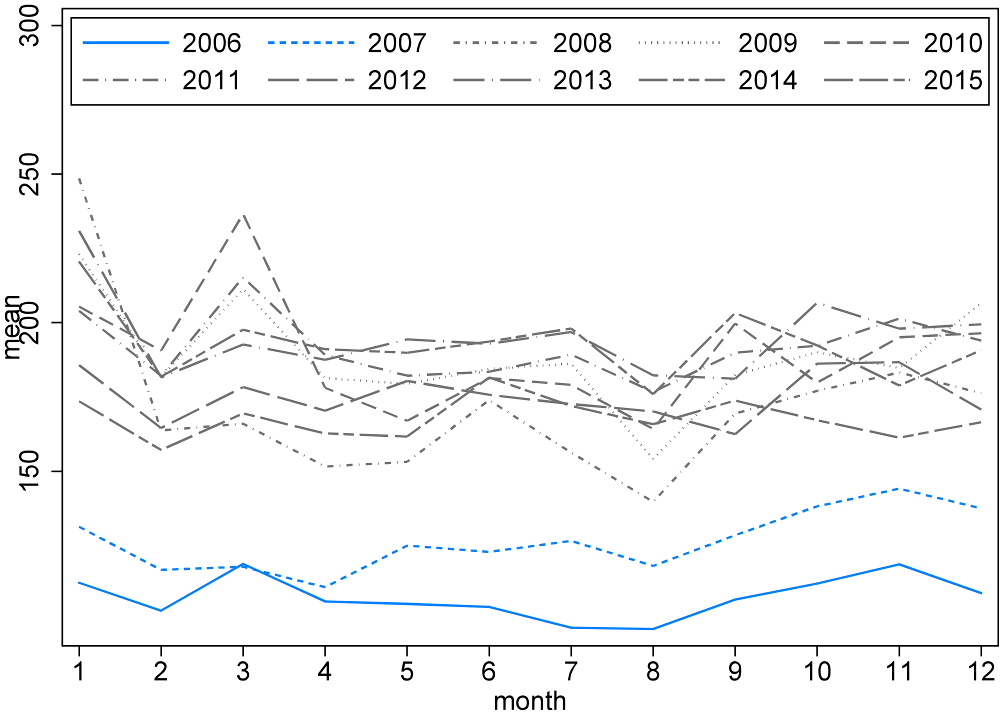
Notes: This figure shows how the average delay between treatment and billing date changed over time. Shown by size of the claim. Calculated using all claims in our data.

Figure C.2: Care consumption around week in which cost-sharing ends



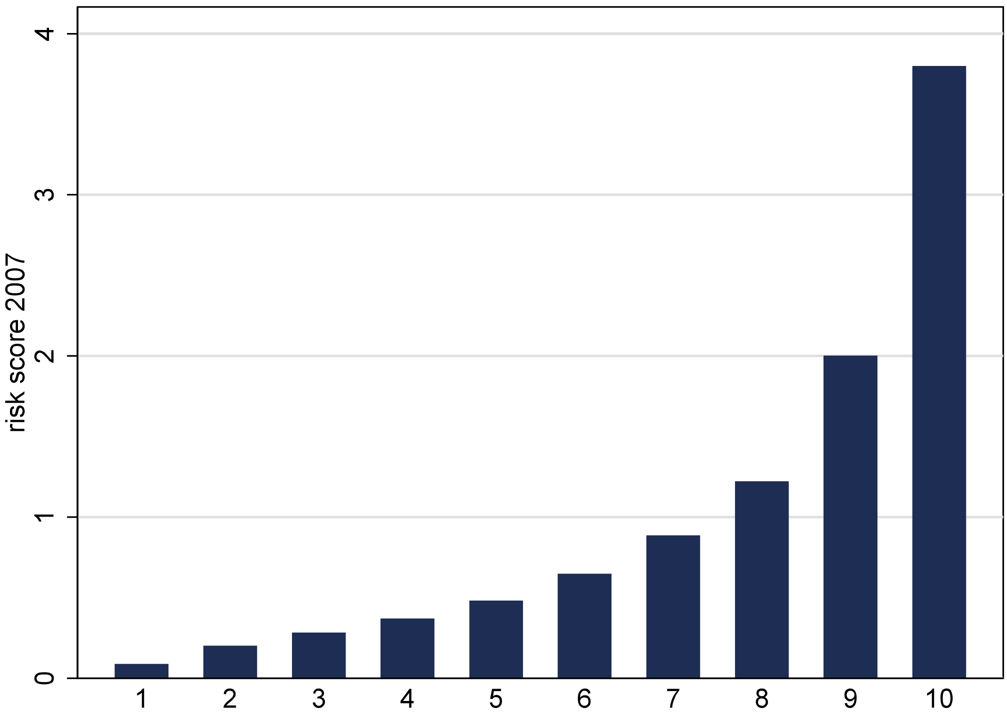
Notes: This figure shows average health care expenditures in weeks before and after exceeding the no-claim refund limit or the deductible in 2007 and 2009, respectively. Cost-sharing incentives were framed as a no-claim refund in 2007 and as a deductible in 2009. Figure 3 in the main text is for 2007 and 2008.

Figure C.3: Care consumption over time



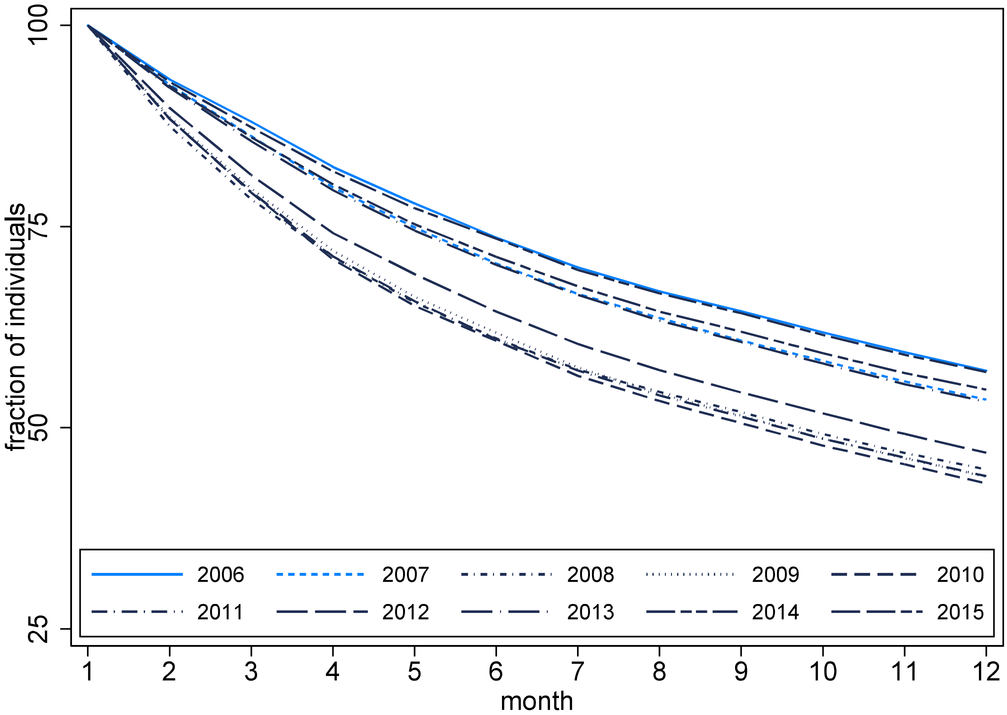
Notes: This figure shows care consumption, by month and for all years between 2006 and 2015. Calculated for the full unbalanced panel.

Figure C.4: Risk score deciles



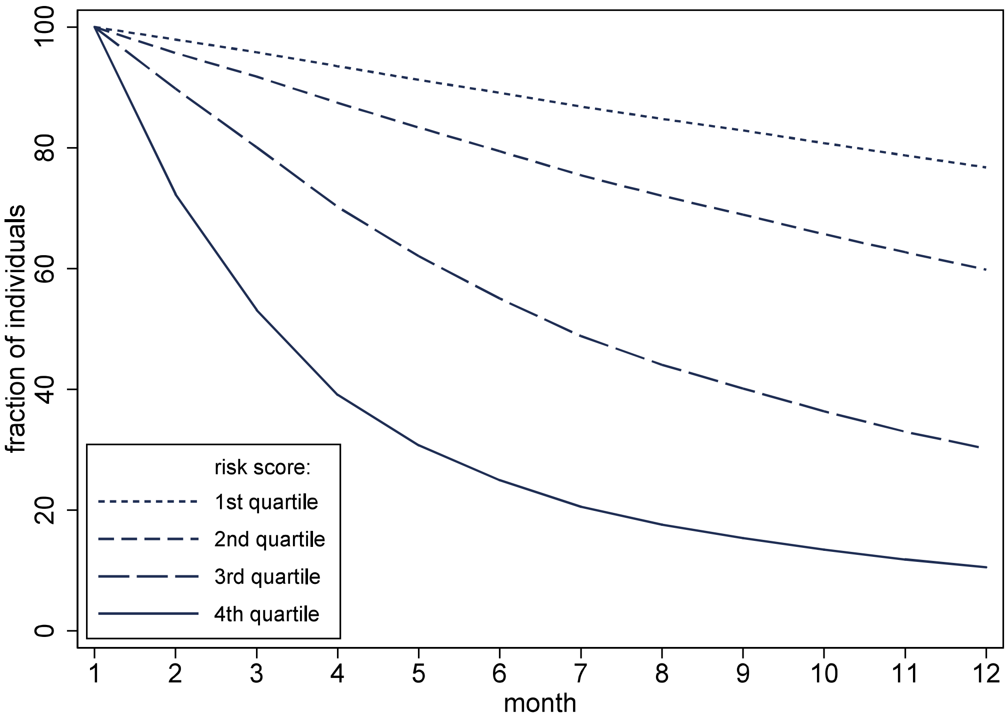
Notes: This figure shows risk score deciles. See Appendix A.2 for details on how the risk score was calculated.

Figure C.5: Fraction of individuals below deductible or no-claim refund limit



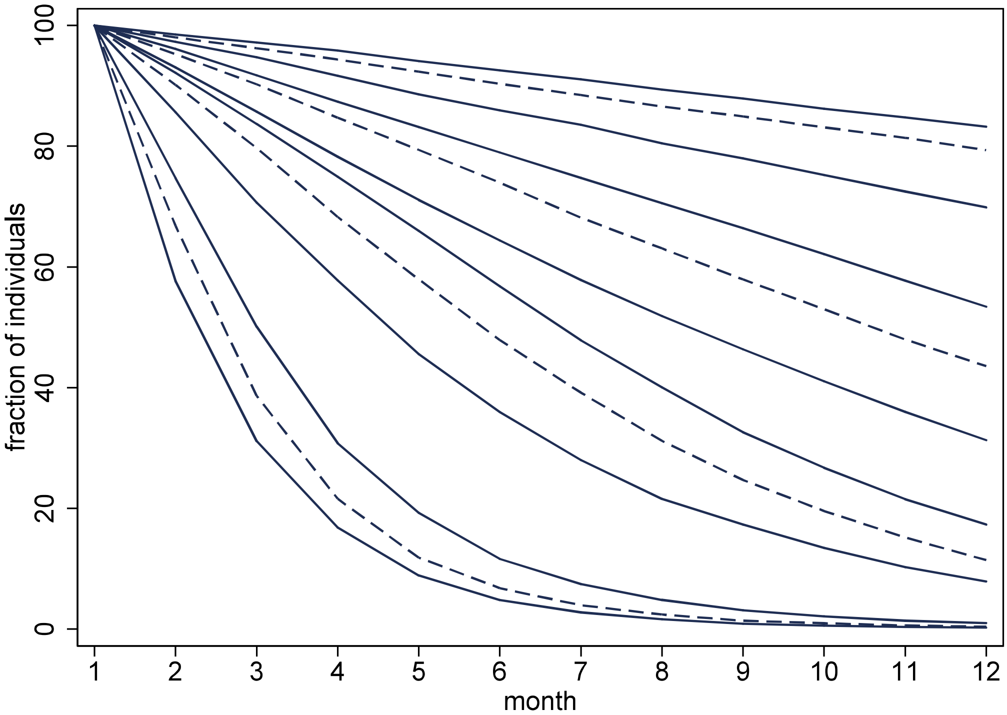
Notes: This figure shows the fraction of individuals who have not hit the deductible or the no-claim refund limit, by month and for all years between 2006 and 2015. Calculated for the full unbalanced panel. The cost-sharing limits were 255, 255, 150, 155, 165, 170, 220, 350, 360, and 375 euros.

Figure C.6: Fraction of individuals below deductible by risk score



Notes: This figure shows the fraction of individuals who have not hit the deductible, by month and by quartile of the risk score, pooled over the years. Figure C.5 shows the fractions by year.

Figure C.7: Variation in instrument



Notes: This figure shows the simulated fractions of individuals by month, for selected risk score deciles and years. The dashed lines are, from left to right, for the 10th, 7th, 5th, and 1st risk score decile in 2007, respectively. The respective solid line to the left of the dashed line is for the same decile and year 2008, and the respective solid line to the right of the dashed line is for the same decile and year 2015.

CINCH working paper series

1 Halla, Martin and Martina Zweimüller**. Parental Responses to Early Human Capital Shocks:** Evidence from the Chernobyl Accident. CINCH 2014.

2 Aparicio, Ainhoa and Libertad González. **Newborn Health and the Business Cycle:** Is it Good to be born in Bad Times? CINCH 2014.

3 Robinson, Joshua J. **Sound Body, Sound Mind?:** Asymmetric and Symmetric Fetal Growth Restriction and Human Capital Development. CINCH 2014.

4 Bhalotra, Sonia, Martin Karlsson and Therese Nilsson. **Life Expectancy and Mother-Baby Interventions:** Evidence from A Historical Trial. CINCH 2014.

5 Goebel, Jan, Christian Krekel, Tim Tiefenbach and Nicolas R. Ziebarth. **Natural Disaster, Environmental Concerns, Well-Being and Policy Action:** The Case of Fukushima. CINCH 2014.

6 Avdic, Daniel, **A matter of life and death? Hospital Distance and Quality of Care:** Evidence from Emergency Hospital Closures and Myocardial Infarctions. CINCH 2015.

7 Costa-Font, Joan, Martin Karlsson and Henning Øien. **Informal Care and the Great Recession.** CINCH 2015.

8 Titus J. Galama and Hans van Kippersluis. **A Theory of Education and Health.** CINCH 2015.

9 Dahmann, Sarah. **How Does Education Improve Cognitive Skills?:** Instructional Time versus Timing of Instruction. CINCH 2015.

10 Dahmann, Sarah and Silke Anger. **The Impact of Education on Personality:** Evidence from a German High School Reform. CINCH 2015.

11 Carbone, Jared C. and Snorre Kverndokk. **Individual Investments in Education and Health.** CINCH 2015.

12 Zilic, Ivan. **Effect of forced displacement on health.** CINCH 2015.

13 De la Mata, Dolores and Carlos Felipe Gaviria. **Losing Health Insurance When Young:** Impacts on Usage of Medical Services and Health. CINCH 2015.

14 Tequame, Miron and Nyasha Tirivayi. **Higher education and fertility:** Evidence from a natural experiment in Ethiopia. CINCH 2015.

15 Aoki, Yu and Lualhati Santiago. **Fertility, Health and Education of UK Immigrants:** The Role of English Language Skills. CINCH 2015.

16 Rawlings, Samantha B., **Parental education and child health:** Evidence from an education reform in China. CINCH 2015.

17 Kamhöfer, Daniel A., Hendrik Schmitz and Matthias Westphal. **Heterogeneity in Marginal Non-monetary Returns to Higher Education.** CINCH 2015.

18 Ardila Brenøe, Anne and Ramona Molitor. **Birth Order and Health of Newborns:** What Can We Learn from Danish Registry Data? CINCH 2015.

19 Rossi, Pauline. **Strategic Choices in Polygamous Households:** Theory and Evidence from Senegal. CINCH 2016.

20 Clarke, Damian and Hanna Mühlrad. **The Impact of Abortion Legalization on Fertility and Maternal Mortality:** New Evidence from Mexico. CINCH 2016.

21 Jones, Lauren E. and Nicolas R. Ziebarth. **US Child Safety Seat Laws:** Are they Effective, and Who Complies? CINCH 2016.

22 Koppensteiner, Martin Foureaux and Jesse Matheson**. Access to Education and Teenage Pregnancy.** CINCH 2016.

23 Hofmann, Sarah M. and Andrea M. Mühlenweg**. Gatekeeping in German Primary Health Care –** Impacts on Coordination of Care, Quality Indicators and Ambulatory Costs. CINCH 2016.

24 Sandner, Malte. **Effects of Early Childhood Intervention on Fertility and Maternal Employment:** Evidence from a Randomized Controlled Trial. CINCH 2016.

25 Baird, Matthew, Lindsay Daugherty, and Krishna Kumar. **Improving Estimation of Labor Market Disequilibrium through Inclusion of Shortage Indicators.** CINCH 2017.

26 Bertoni, Marco, Giorgio Brunello and Gianluca Mazzarella. **Does postponing minimum retirement age improve healthy behaviors**

before retirement? Evidence from middle-aged Italian workers. CINCH 2017.

27 Berniell, Inés and Jan Bietenbeck. **The Effect of Working Hours on Health.** CINCH 2017.

28 Cronin, Christopher, Matthew Forsstrom, and Nicholas Papageorge. **Mental Health, Human Capital and Labor Market Outcomes.** CINCH 2017.

29 Kamhöfer, Daniel and Matthias Westphal. **Fertility Effects of College Education:** Evidence from the German Educationl Expansion. CINCH 2017.

30 Jones, John Bailey and Yue Li. **The Effects of Collecting Income Taxes on Social Security Benefits.** CINCH 2017.

31 Hofmann, Sarah and Andrea Mühlenweg. **Learning Intensity Effects in Students' Mental and Physical Health –** Evidence from a Large Scale Natural Experiment in Germany. CINCH 2017.

32 Vollmer, Sebastian and Juditha Wójcik. **The Long-term Consequences of the Global 1918 Influenza Pandemic:** A Systematic Analysis of 117 IPUMS International Census Data Sets. CINCH 2017.

33 Thamarapani, Dhanushka, Rockmore, Marc, and Willa Friedman. **The Educational and Fertility Effects of Sibling Deaths**. CINCH 2018.

34 Lemmon, Elizabeth. **Utilisation of personal care services in Scotland:** the influence of unpaid carers. CINCH 2018.

35 Avdic, Daniel, Büyükdurmus, Tugba, Moscelli, Giuseppe, Pilny, Adam, and Ieva Sriubaite**. Subjective and objective quality reporting and choice of hospital:** Evidence from maternal care services in Germany. CINCH 2018.

36 Hentschker, Corinna and Ansgar Wübker. **Quasi-experimental evidence on the effectiveness of heart attack treatment in Germany.** CINCH 2018.

37 Pasha, Mochamad, Rockmore, Marc, and Chih Ming Tan. **Early Life Exposure to Above Average Rainfall and Adult Mental Health**. CINCH 2018

38 Elsner, Benjamin and Florian Wozny. **The Human Capital Cost of Radiation:** Long-run Evidence from Exposure outside the Womb. CINCH 2019

39 de la Mata, Dolores and Carlos Felipe Gaviria Garcés. **Exposure to Pollution and Infant Health:** Evidence from Colombia. CINCH 2019

40 Besstremyannaya, Galina and Sergei Golovan. **Physicians' altruism in incentives contracts:** Medicare's quality race. CINCH 2019

41 Hayen, Arthur P., Klein, Tobias J., and Martin Salm. **Does the framing of patient cost-sharing incentives matter?** The effects of deductibles vs. no-claim refunds. CINCH 2019