Journal of Health Economics 90 (2023) 102782

Contents lists available at ScienceDirect

Journal of Health Economics

journal homepage: www.elsevier.com/locate/jhe

Minimum generosity levels in a competitive health insurance market☆

Minke Remmerswaal [a,b,*], Jan Boone [a,b,c], Rudy Douven [a,d]

[a] CPB Netherlands Bureau for Economic Policy Analysis, The Hague, The Netherlands
[b] Department of Economics, Tilec, Tilburg University, The Netherlands
[c] CEPR, London, United Kingdom
[d] Erasmus School of Health Policy & Management, Erasmus University, Rotterdam, The Netherlands

ARTICLE INFO

Keywords:
Deductible
Healthcare expenditure
Moral hazard
Adverse selection
Panel data
Cost-sharing

ABSTRACT

An important condition for optimal health insurance is that the level of health care coverage is inversely related to the elasticity of demand. We show that this condition is not satisfied for voluntary deductibles in the Netherlands, which are optional deductibles on top of the mandatory deductible introduced by the Dutch government. We find that low-risk types, that mainly choose voluntary deductibles, have a lower elasticity of demand than high-risk types. Moreover, we show that voluntary deductibles introduce equity problems as it results in non-trivial cross subsidies from high-risk to low-risk types. Capping the level of voluntary deductibles (imposing minimum generosity) is likely to be welfare enhancing in the Netherlands.

1. Introduction

Since Arrow (1963) it is well-known that full coverage is optimal in competitive health insurance markets in the case of no moral hazard. Pauly (1968) and Zeckhauser (1970) show that with moral hazard, coverage levels should be set lower to prevent excessive health care consumption by consumers. They argue that the level of coverage should be inversely related to the elasticity of demand, i.e. the higher the elasticity of demand the lower coverage levels should be set. This implies that if a government wants to differentiate coverage levels across consumers, consumers with a high elasticity of demand should face lower levels of coverage. However, differentiation in coverage levels can also introduce consumer choice, and thus selection problems. For example, insurers may try to cherry pick low-risk-types by offering coverage contracts that are not attractive for high-risk types (Rothschild and Stiglitz, 1976). Such risk selection may reduce welfare as well as lead to a cross subsidy from high-risk types – which tend to have low incomes – to low-risk types that are healthier and tend to have higher earnings. This may be seen as undesirable from

☆ We gratefully acknowledge comments and suggestions by two referees of this journal, Judith de Jong, Tobias Klein, Owen O'Donnell, Mieke Reuser, Wynand van de Ven, seminar participants at AHES in Melbourne, 2019, ASHEcon in Washington DC, 2019, RAN conference in Den Haag, 2017, iHEA in Boston, 2017, euHEA in Maastricht, 2018, lolaHESG in Hoenderloo, 2018, and colleagues at CPB Netherlands Bureau for Economic Policy Analysis. Jan Boone thanks NWO for financial support through a Vici grant. An earlier version of this paper circulated under the title "Selection and Moral Hazard Effects of a Voluntary Deductible in a Competitive Health Insurance Market". Minke Remmerswaal worked at the CPB Netherlands Bureau for Economic Policy Analysis while undertaking this research.

* Corresponding author at: Department of Economics, Tilec, Tilburg University, The Netherlands.
E-mail address: minkeremmerswaal@gmail.com (M. Remmerswaal).

https://doi.org/10.1016/j.jhealeco.2023.102782
Received 28 July 2022; Received in revised form 31 May 2023; Accepted 5 June 2023
Available online 14 June 2023

a distribution or solidarity point of view. In many countries policymakers have to make this trade-off between combatting moral hazard by offering more choice in coverage versus selection and distribution problems.[1]

We analyze different insurance coverage levels in the context of Dutch health insurance, where private insurers compete in a regulated market. One of the interesting features of this regulation is that it specifies a minimum and maximum generosity level for health insurance. In particular, there is a mandatory deductible that applies to everyone above age 18 and people can opt for an additional voluntary deductible which is at maximum 500 euros.[2] In the Netherlands, this is a fixed deductible level in euros (i.e. the co-insurance rate equals 1). This paper focuses on the specification of the maximum voluntary deductible (or minimum generosity) for health insurance.

We introduce a theoretical framework to show that specifying a minimum generosity in a health insurance market is in general welfare enhancing. The only exception is a competitive insurance market where there are no (adverse) selection issues, for example due to a perfect risk adjustment system. In such a case the market will offer optimal generosity levels. However, we are not aware of any country with a competitive health insurance market that does not suffer from selection problems. The intuition why a minimum generosity level increases welfare is that the market inefficiently reduces generosity in order to cherry pick low-risk types. From this theoretical framework we derive conditions under which offering a voluntary deductible (on top of a mandatory deductible) increases welfare. To illustrate how countries can check these conditions, we use Dutch data on individual healthcare expenditures. We show that for the Netherlands it is actually better to abolish the voluntary deductible; that is, equate the minimum and maximum generosity level in health insurance.[3]

To assess whether allowing for a voluntary deductible is welfare enhancing in the Netherlands, we need to separate moral hazard from selection effects. This is empirically challenging (Bajari et al., 2014; Trottmann et al., 2012; Geruso and Layton, 2017). Randomized controlled experiments, such as the RAND and Oregon experiments (Newhouse and the Insurance Experiment Group, 1993; Finkelstein et al., 2012; Chiappori et al., 1998), are less suited as the idea of randomization is to remove endogenous choice, and thus selection problems. That is, in such an experimental set-up there is no selection effect to start with. The majority of papers analyzing non-experimental data remove selection by using econometric techniques such as instrumental variables (examples are Eichner, 1998; Trottmann et al., 2012; Van Vliet, 2004; Alessie et al., 2020) or by using a structural model (Bajari et al., 2014; Gardiol et al., 2008). These approaches often require restrictive assumptions. Bajari et al. (2014) and Einav et al. (2013) measure and visualize selection by showing the distribution of health status and expected health risks parameters across multiple health insurance plans.

Our identification strategy relies on the start of cost-sharing in the Netherlands at age 18 and the yearly increase of the mandatory deductible size by the government during 2008–2013. This identification, in the spirit of a regression discontinuity design, allows us to separate selection from moral hazard effects in a clean way. In many adverse selection papers, such as Bajari et al. (2014), Einav et al. (2013) and Cutler and Reber (1998), endogenous contracts complicate the analysis as contracts differ in a number of dimensions.[4] In the Netherlands, endogenous contract dimensions are limited, because the coverage of the mandatory (basic) healthcare plans across the Dutch population only differs by the level of the deductible. The treatments covered by basic insurance are set by the government, not by the insurers. Insurance contracts can differ in their provider networks, but these differences are small in the period that we analyze.

We find that average healthcare expenditures for persons who never choose a voluntary deductible after age 18 (in our data) are almost twice as high just before age 18 than average healthcare expenditures of persons who have chosen a voluntary deductible after 18 in our data. This difference is mainly the result of selection, not a reduction of moral hazard. Indeed, expenditure decisions before 18 are not affected by cost sharing after 18.[5]

Moreover, we find that 18 year olds who have never chosen a voluntary deductible reduce their healthcare spending by 26 euros (on average) in response to a 100 euro increase in the (mandatory) deductible. However, for 18 year olds who have chosen a voluntary deductible we find on average no response to an increase in the (mandatory plus voluntary) deductible. Other potential causes for these findings – for example the difference in price responsiveness between the group who chooses and not chooses for a voluntary deductible could be explained by whether individuals are at the margin – are not confirmed in our data. The lower moral hazard effect for the group with a voluntary deductible (compared to the group with only a mandatory deductible) is one of the main reasons why abolishing the voluntary deductible for the Netherlands will increase welfare.

Next we perform a policy analysis at the population level to determine the moral hazard and selection effects and the size of cross subsidies from high-risk to low-risk individuals due to voluntary deductibles in the Netherlands. To do this, we extend our analysis for 18 year olds to the full population. We use a panel regression design that exploits annual variation in the size of the mandatory and voluntary deductibles. This allows us to quantify the cross subsidies: the part of the reduction (discount) in the insurance premium that is not due to increased out-of-pocket expenditure and reduced healthcare use because of the higher deductible. This

[1] This holds for countries with an insurance market based on managed competition with risk adjustment and community rating, such as for example in Australia, Germany, Israel, the Netherlands, Switzerland and in the US (McGuire and Van Kleef, 2018).

[2] Everyone under 18 years old does not pay a premium nor any form of cost-sharing for health insurance in the Netherlands.

[3] We do not measure the welfare gains from offering choice of a voluntary deductible, but simulations by Handel et al. (2022) indicate that these gains are very small. We come back to this aspect in Section 7.

[4] In the United States for example, selection may also occur from having "no insurance". The majority of the American body of literature was conducted before the introduction of the Patient Protection and Affordable Care Act, and thus before trying to introduce mandatory health insurance.

[5] Note that an anticipation effect would work in the opposite direction. If I know I will choose a high voluntary deductible after 18, I should increase my expenditure before 18 when it is still free. But we find the opposite: a voluntary deductible goes hand-in-hand with low expenditure before 18.

reduction in the premium is due to selection of people with lower expected costs (due to imperfect risk adjustment). We estimate that allowing for a voluntary deductible of up to 500 euros leads to a cross subsidy from high-risk to low-risk individuals of between 38–80 euros per low-risk person per year.

Our research provides lessons for other countries that implement choice in demand-side cost-sharing in a competitive health insurance markets with community rating and (imperfect) risk adjustment. First, in general it is a good idea to impose a minimum generosity level in the health insurance market to prevent selection activities on low-risks and cross subsidies from high-risk to low-risk types. Second, the lower moral hazard effect, i.e. the lower responsiveness of expenditure with respect to the deductible, for people with a voluntary deductible implies that the minimum generosity level should be equal to the maximum generosity level in the Netherlands. This may also be the case in other countries. Einav et al. (2013) find a similar effect and coin this "selection on moral hazard". Moreover, policy makers have to consider another problem of introducing voluntary cost-sharing schemes in competitive markets. If risk adjustment is imperfect then this may lead to selection effects which ultimately can result in unraveling markets (Cutler and Reber, 1998), and unintended cross subsidies between low-risk and high-risk individuals (Marzilli Ericson and Sydnor, 2017). For the Netherlands we find no evidence for a death spiral. As health insurance is mandatory in the Netherlands, a death spiral would not lead to uninsured citizens. Also, there is no evidence of a spiral in terms of insurers reducing coverage or provider networks. Apparently, the combination of mandatory health insurance together with a mandatory and maximum deductible level is sufficient to prevent any spiral developing. But we do find selection effects and fairly substantial cross subsidies.

Therefore, we conclude that policymakers should be reluctant to implement a voluntary cost-sharing scheme if there is already a mandatory cost-sharing scheme in place. If, unlike the Netherlands, generosity choice is seen as important by policy makers due to clear heterogeneity in risk aversion, implementing a voluntary deductible scheme can be optimal. But capping the voluntary deductible is likely to be welfare enhancing in this case as well.

In Section 2 we explain how we think about moral hazard, adverse selection and cross subsidies between low-risk and high-risk individuals in our paper and we derive a hypothesis we can test empirically. Section 3 describes the institutional setting of the Dutch healthcare sector. Section 4 explains our administrative data set and provides several descriptive statistics. We describe our empirical strategy in Section 5 and present our results in Section 6. Section 7 extends the sample to a larger age range for policy analyses. We conclude in Section 8.

2. Moral hazard, adverse selection and welfare

This section presents a model showing that imposing a minimum generosity on a health insurance contract is, in general, welfare enhancing. In the Netherlands this minimum generosity level takes the form of a maximum voluntary deductible of 500 euros (on top of the mandatory deductible). Further, we derive conditions under which it is optimal to have the same minimum and maximum generosity level. In the Dutch context this implies abolishing the voluntary deductible (or equivalently set the maximum voluntary deductible equal to zero).

We start with an overview of the health insurance context (see Section 3 for a detailed discussion of the Dutch institutional setting). Next, we present a formal model (technical details and proofs can be found in the appendix) and derive three policy implications.

2.1. Background

Health insurance leads to moral hazard: the increased use of healthcare because health insurance lowers a person's marginal costs of healthcare use (Zweifel, 1987; Einav and Finkelstein, 2018). Most health insurance plans have cost-sharing schemes in place to prevent moral hazard.

The effect of cost-sharing in health insurance on healthcare demand and expenditures is well established and for excellent overviews of the literature, we refer to Baicker and Goldman (2011), McGuire (2012), Geruso and Layton (2017) and Einav and Finkelstein (2018). See also Appendix 9.A for an overview of the literature on moral hazard and adverse selection in Dutch health insurance.

Moral hazard was one of the reasons for Dutch policy makers to introduce a voluntary deductible: it helps to contain healthcare expenditure. In addition, a voluntary deductible offers choice in the level of cost-sharing. Offering choice in health insurance plans is beneficial if people have (sufficiently) different risk preferences (Cutler and Zeckhauser, 2000; Handel et al., 2022). But a health insurance plan with higher cost-sharing (i.e. a voluntary deductible in Dutch health insurance) is likely to be profitable (only) for people who need little healthcare and have low healthcare expenditures. As a result, these low-risk persons are more likely to choose the voluntary deductible. The result is selection. In this paper, we estimate selection in terms of the difference in healthcare expenditure of persons in plans with and without a voluntary deductible, in absence of cost-sharing.

The last effect of the voluntary deductible that we discuss in this paper are cross subsidies between persons with and without a voluntary deductible. Whether such cross subsidies take place is important for Dutch politicians and policy makers who value risk solidarity among high-risk and low-risk individuals. If due to higher cost-sharing, people pay more for healthcare out-of-pocket, it is seen as fair that they also pay a lower premium. However, if the premium discount for a voluntary deductible is too high, this may constitute a cross subsidy from high-risk (with only a mandatory deductible) to low-risk individuals (who are more likely to choose a voluntary deductible and have no or low costs anyway). We formalize this discussion in the following model.

2.2. A model of health insurance

Consider two types of consumers denoted l, h with l for a low-risk type and h for a high-risk type. The expected healthcare needs of the two types satisfy $\theta_l < \theta_h$ and the variance of these healthcare expenditures is denoted by σ^2 for both types. Furthermore, the fraction of low-risk type persons is given by $\zeta \in [0, 1]$.

When a person of type $i \in \{l, h\}$ buys health insurance with coverage $q \in [0, 1]$, her utility is given by

$$u_i = y - (1-q)\theta_i - p - \tfrac{1}{2} r\sigma^2(1-q)^2 + \varepsilon(q) \tag{1}$$

where y denotes the person's income or wealth, p the price (premium) of health insurance and r captures the degree of risk aversion. To keep the model tractable we do not model differences in risk preferences and assume $r_l = r_h = r$. We come back to this below and in Section 7. The fraction of healthcare expenditure covered by health insurance is denoted by q. The remaining variance in out-of-pocket expenditure for the consumer is given by $\sigma^2(1-q)^2$.[6]

Finally, $\varepsilon(q)$ is part of the way in which we capture moral hazard. Without insurance coverage, this moral hazard term disappears ($\varepsilon(0) = 0$). Furthermore, it is increasing in q: $\varepsilon'(q) \geq 0, \varepsilon''(q) \leq 0$. The term $\varepsilon(q)$ captures the utility an agent derives from (cost-inefficient) treatments. It can also be seen as including the part q that the agent pays out-of-pocket for these treatments. The fact that ε is increasing in q then captures that the value of treatment exceeds the out-of-pocket price. But it is inefficient because treatment value falls short of the total (or social) cost. This will be captured in the insurer's profit function below where the total costs of q exceed the benefit. One can think here of doing additional tests or scans when it is not medically necessary to do these. The effect of moral hazard on utility is symmetric between types as we think of the utility effect as being close to zero. The effect on healthcare costs and hence on insurers' profits does vary by type. The elasticity of demand with respect to coverage q is captured by this effect on total expenditure.

The idea of Eq. (1) is that a person is hit by a health shock and spends money on treatments to restore her health. The utility function captures the financial consequences of paying for the necessary treatments,[7] and adds a (reduced form) moral hazard term.

To ease notation, we do not let income y vary by type. But our results are the same if we allow high risk types (low health status) to have lower income than low risk types: $y_h < y_l$. Such a correlation between health status and income further motivates the idea that a government wants to reduce transfers (cross subsidies) from low health agents with low income to healthy agents with high incomes.

We consider a competitive health insurance market with oligopoly insurers where the government imposes community rating, as is the case in the Netherlands. Although this means that health insurers cannot explicitly separate insured persons through risk-rating, they can implicitly separate the types by offering two insurance contracts with features (q_l, p_l) and (q_h, p_h) where high-risk and low-risk types are supposed to choose their own contract.

We follow the literature and use the binding incentive compatibility (IC) constraint that the h type would like to mimic the l type.[8] That is, insurers have to make sure that h contract is sufficiently attractive for the high-risk type h that she does not buy the cheaper l contract.

$$u_h \geq y - (1-q_l)\theta_h - p_l - \tfrac{1}{2} r\sigma^2(1-q_l)^2 + \varepsilon(q_l) \tag{2}$$

When the h type buys the h contract, her utility is given by u_h. If she would buy the l contract, her utility is given by the right hand side of Eq. (2). This contract features $q = q_l, p = p_l$. If the inequality holds, the h type (weakly) prefers to buy the h contract and does not deviate to the l contract.

Using the expression for u_l in Eq. (1), we can write the IC_h constraint as

$$u_h \geq u_l - (1-q_l)\Delta\theta \tag{IC_h}$$

where $\Delta\theta = \theta_h - \theta_l$. Eq. ($IC_h$) again illustrates that making sure that high-risk type h chooses her own contract implies that her utility u_h should be sufficiently high compared to u_l. Eq. (IC_h) also illustrates that the lower the coverage of the low-risk plan q_l is, the lower is the right-hand side of the IC constraint becomes.

Selection and moral hazard effects

A (risk neutral) insurer's profits of contract (q_i, p_i) bought by type i are given by:

$$\pi_i = p_i - q_i\theta_i + \rho_i - \eta_i(q_i) \tag{3}$$

where the firm earns the premium p_i, has expected costs $q_i\theta_i$ and receives risk adjustment contribution ρ_i for contracting type i. Moral hazard is captured by $\eta_i(q_i)$ in the following way: $\eta_i(0) = 0, \varepsilon'(q_i) - \eta_i'(q_i) < 0$. If there is no insurance coverage, the consumer behaves efficiently and there are no extra costs. As q_i increases, the consumer becomes less cost sensitive in a way that is inefficient:

[6] Note that we model demand-side cost-sharing here as a co-insurance rate, while the empirical application below considers a deductible. The reason is that this framework (based on Rothschild and Stiglitz, 1976) is better known and more tractable than a theoretical set-up with deductibles. The main intuition is the same for co-insurance rates and deductibles.

[7] A health shock can cause a loss in utility. This would be modeled by adding a (negative) constant to the utility function above. As this does not affect the analysis, we leave out this constant to ease notation.

[8] Afterwards we check whether the other IC constraint is satisfied as well: the low-risk type does not want to buy the (expensive) high-risk contract.

the gain for the consumer ($\varepsilon'(q_i)$) is strictly lower than the cost for the insurer $\eta_i'(q_i)$. This is because due to cost-sharing a consumer pays a fraction of a treatment's costs and therefore accepts treatments with a gain larger than his or her out-of-pocket payment but lower than the treatments' costs for the insurer.

We define the part of type i's total healthcare expenditure that does not depend on q_i as $C_i = \theta_i + \eta_i(1)$; total expenditure if there is full coverage $q_i = 1$. The moral hazard effect is defined as $MH_i = \eta_i(q_i) - \eta_i(1) < 0$: the fall in expenditure due to less than full coverage, $q_i < 1$.

Optimal health insurance coverage

We define total welfare for type i as $w_i = u_i + \pi_i$; hence, we find

$$w_i = y - \theta_i - \tfrac{1}{2} r\sigma^2(1-q_i)^2 + \rho^i + \varepsilon(q_i) - \eta_i(q_i) \tag{4}$$

Taking the derivative with respect to q_i, we find first best insurance coverage for type i. In the appendix we work with general functions $\varepsilon(q), \eta_i(q)$, here with simple linear functions: $\varepsilon * q, \eta_i * q$.

$$q_i^* = 1 - \frac{\eta_i - \varepsilon}{r\sigma^2} \tag{5}$$

If there is no moral hazard ($\eta_i - \varepsilon = 0$), it is optimal to have full insurance coverage: $q_i = 1$. Because the consumer is risk averse ($r > 0$) and the insurer is risk neutral, it is optimal to shift the expenditure risk completely to the insurer. If there is moral hazard, $\eta_i - \varepsilon > 0$, it is optimal to reduce the coverage and leave some expenditure risk for the insured consumer to reduce inefficient care consumption. The more risk averse the consumer is (higher denominator $r\sigma^2$), the smaller the effect of moral hazard on the optimal coverage. Types have different optimal coverage levels because the moral hazard over risk aversion term ($(\eta_i - \varepsilon)/r\sigma^2$) differs between the types. This is line with Pauly (1968) and Zeckhauser (1970). Furthermore, if $\eta_l < \eta_h$, the moral hazard effect is smaller for the low-risk type, then the optimal generosity level is higher for the low-risk type than for the high-risk type.

To characterize the market outcome with imperfect competition, we assume that there are two health insurers, denoted a, b, competing in a Hotelling model. In Appendix 9.B, we show that in market equilibrium $\hat{q}_h = q_h^*$ and $\hat{q}_l < q_l^*$. This means that coverage for the h type is set efficiently in the market (at its first best level in Eq. (5)), but for the l type it is set too low. The stronger the selection effects (cherry picking incentives for insurers) are, the bigger the difference $q_l^* - \hat{q}_l > 0$. Perfect risk adjustment removes selection incentives and we find $\hat{q}_l = q_l^*$. With perfect risk adjustment, the government can leave the coverage choice to the market and there is no reason to intervene from an efficiency perspective.[9]

In Appendix 9.B we derive the following expression for the premium difference $\Delta p = p_h - p_l$ between high-risk and low-risk types. We use this expression in Section 7 to decompose the difference in observed prices paid by people with and without a voluntary deductible.

$$\Delta p = \Delta\text{mark up} + \Delta C - \Delta\rho - \Delta oop + \Delta MH \tag{6}$$

where the insurer's mark up covers loading fees and generates profits, the difference in exogenous costs and the risk adjustment contribution is given by $\Delta C = C_h - C_l, \Delta\rho = \rho_h - \rho_l$, difference in out-of-pocket expenditure $\Delta oop = (1-q_h)\theta_h - (1-q_l)\theta_l$ and the difference in the moral hazard effect $\Delta MH = (\eta_h(q_h) - \eta_h(1)) - (\eta_l(q_l) - \eta_l(1))$.

When considering total welfare ($W = \zeta w_l + (1-\zeta) w_h$), transfers between the parties are irrelevant. The only relevant efficiency effect here is that $\hat{q}_l < q_l^*$ implies a reduction of welfare. We consider the effect of the government setting a lower bound on generosity $\bar{q}_l$ such that insurers have to set $q_l \geq \bar{q}_l$ (this corresponds to an upper bound on the voluntary deductible in the Dutch case). We consider $\bar{q}_l > \hat{q}_l$.[10]

The following proposition focuses on the Dutch case that we analyze below. We assume that the government has set a maximum generosity level (this corresponds to a mandatory deductible) equal to or less than q_h^*. If this maximum generosity level would exceed q_h^*, we know from Rothschild and Stiglitz (1976) that the market would implement the optimal level q_h^* (also see the proof of Proposition 1 in the appendix). Hence, we would observe in our data a lot of people choosing a voluntary deductible on top of the (rather low) mandatory deductible. However, this we do not see. Further, if the maximum generosity is set too low by the government (mandatory deductible set too high), it is less likely that q_l^* is below this maximum generosity level. Consequently, assuming that the maximum generosity level corresponds to q_h^*, makes it most likely that there is a role for $\bar{q}_l$ below the maximum generosity level (most likely that there is a role for voluntary deductibles). From now on we assume that the Dutch government set the mandatory deductible equal to the level that is optimal for the high risk types.

The question is whether a voluntary reduction in generosity $q_l < q_h^*$ should be allowed and, if so, whether it should be capped: $q_l \geq \bar{q}_l$. Setting $\bar{q}_l = q_h^*$ abolishes a voluntary reduction in generosity.

Proposition 1.

- *If there are no selection incentives in the market, the minimum generosity level can be set to zero ($\bar{q}_l = 0$) as the market implements the efficient outcome $q_h = q_h^*, q_l = q_l^*$.*

[9] Note that theoretically we cannot rank q_l^* and q_h^*; but we can do so empirically for our Dutch data as we show below.

[10] Indeed, if $\bar{q}_l < \hat{q}_l$, the government imposed bound will be ignored by the market. This is not relevant in the Dutch case where most insured bunch on the maximum voluntary deductible, suggesting that the minimal generosity constraint is indeed binding for the market.

- *If there are selection incentives in the market, imposing a lower bound $\bar{q}_l$ on coverage increases welfare W.*
- *If there are selection incentives and $q_h^* \leq q_l^*$, it is optimal to set $\bar{q}_l = q_h^*$.*
- *If there are selection incentives, an increase in $\bar{q}_l < q_h^*$ reduces cross subsidies from high-risk to low-risk types by reducing Δmark up in Eq. (6).*

If the insurance market outcome does not feature (adverse) selection problems, there is no need – from a total welfare point of view – for the government to intervene. This can happen if risk adjustment ρ_i is perfect. Alternatively, if the differences in risk preference between the l and h types are such that offering each type its first best coverage q_i^* leads to an outcome that satisfies all incentive compatibility constraints. In either case, there is no need for the government to impose a lower bound $\bar{q}_l$.

In all countries that we are aware of, there are selection issues in competitive health insurance markets. In this case, it is always optimal for the government to impose a lower bound $\bar{q}_l$ on insurance generosity. The market reduces q_l too much to allow insurers to cherry pick low-risk types. This is inefficient and allows a transfer of rents from high-risk to low-risk types.

If first best coverage of high-risk types is lower than that for low-risk types, the mandatory coverage should be set equal to q_h^* and no voluntary reductions in coverage should be allowed.

Finally, increasing the lower bound $\bar{q}_l$ reduces the transfer of rents from high-risk to low-risk types. This reduction in transfers does not affect total welfare (the increase in $\hat{q}_l$ itself does), but is beneficial from a distributive or solidarity point of view. A government may not want to shift money from people in low health to healthy citizens that tend to earn higher incomes.

Below we use Eq. (7) to determine where cross subsidies are present in the Dutch case. The cross subsidy is determined by Δmark up. Because ΔC and $\Delta\rho$ are exogenous with respect to a change in q_l, we use the following decomposition in Section 7.[11]

Corollary 1.

$$\textit{cross subsidy}(q_l) = \Delta p - (oop_l(q_l) - oop_l(q_h^*)) + (MH_l(q_l) - MH_l(q_h^*)) \tag{7}$$

The price difference $\Delta p = p_h - p_l > 0$ between the contract with only a mandatory deductible and the one with a voluntary deductible is observed in our data. This price difference can be partly explained by (i) the higher out-of-pocket payments of the less generous contract compared to the contract with the mandatory deductible $oop_l(q_l) - oop_l(q_h^*) > 0$ and (ii) the reduced moral hazard costs of the less generous contract $MH_l(q_l) - MH_l(q_h^*) < 0$. The part of Δp not explained by these two effects, we denote the cross subsidy to people with the voluntary deductible.

Policy implications

The analysis above allows us to formulate multiple policy implications. The first policy implication is that the determination of insurance coverage can be left to the market if there are no selection problems. This is the case if, for instance, risk adjustment is perfect. Although the Dutch risk adjustment system is quite good, it is certainly not perfect. It has been documented that it is possible for health insurers to identify groups that are either profitable or loss making in expected terms (Withagen-Koster et al., 2018). To the best of our knowledge, this is true for all countries with competitive health insurance markets.

The second implication is that, in this case, it is optimal to mandate a lower bound on insurance generosity. To facilitate selection of low-risk insured, the market offers coverage that is too low. Increasing this coverage by imposing a lower bound raises welfare and reduces the transfer of money from high-risk to low-risk insured. This is already implemented in the Netherlands because there is a set maximum voluntary deductible of 500 euro.

The third implication is: if first best insurance coverage for the high-risk types is below the optimal coverage for low-risk types, these two coverage levels should be equalized. Eq. (5) shows that $q_h^* \leq q_l^*$ if the moral hazard effect for high types exceeds this effect for low types: $\eta_h \geq \eta_l$. If this turns out to be the case for the Netherlands, it is optimal to set the maximum voluntary deductible equal to 0; that is, abolish the voluntary deductible.

In the remainder of this paper we show how one can determine whether this condition on the moral hazard effect is satisfied.

3. Institutional setting

The Dutch curative healthcare sector is characterized by regulated competition, which is written down in the Health Care Act (Zorgverzekeringswet). There is competition among private health insurers and among healthcare providers (Van de Ven and Schut, 2008). To safeguard solidarity and access to care, the government has set up a mandatory basic benefit package for all Dutch citizens.[12]

All inhabitants of the Netherlands, except children up to the age of 18, pay for healthcare costs in three ways. The first part is an health insurance premium that individuals pay directly to their health insurer. Annual premiums are between 1000 to 1250 euros (see Appendix 9.C). A person with a low income can receive an income dependent subsidy to pay for his or her health insurance premium. The second part is an income dependent fee, which the tax collector levies on an individual basis. These income dependent fees must cover exactly 50% of total health expenditures in the Netherlands. The last part is cost-sharing.

[11] The Corollary implicitly assumes that the cross subsidy equals 0 when $q_l = q_h^*$. Without this assumption, the left hand side of the equation becomes cross subsidy(q_l) – cross subsidy(q_h^*).

[12] The Health Care Act has been in place since 2006. Before 2006, there was no regulated competition in Dutch curative healthcare. Insurance was only mandatory for persons with a low or middle income, but not for persons with a high income. The latter had to buy health insurance from a private insurer.

Table 1
Deductibles of health insurance plans in the Netherlands for 2008 up to 2013.

Year	2008	2009	2010	2011	2012	2013
Mandatory (€)	150	155	165	170	220	350
Voluntary (€)	100, 200, 300, 400, or 500					

In 2008, the government introduced a mandatory deductible of 150 euros for all health insurance plans of persons aged 18 and over in the Netherlands.[13] Since 2008, the government increased the level of the deductible each year; see Table 1.[14]

In addition to the mandatory deductible, an individual can opt for a voluntary deductible of maximally 500 euros on top of the mandatory deductible. To illustrate, if a person chooses a voluntary deductible of 500 euros in 2013, he or she faces a *total* deductible of 850 euros. In return for choosing the higher deductible, individuals pay a lower health insurance premium. Insurers are free to determine the size of this premium discount, but on average the discount for a 100 euro voluntary deductible was 45 euros and for a 500 euro voluntary deductible 230 euros in 2013. Aside from the difference in the level of cost-sharing, health insurance plans with and without a voluntary deductible are exactly the same. Health insurers are however free to set the size of the premium discount, so it is possible that some plans offer bigger premium discounts for a voluntary deductible, making them more attractive. Therefore, there is some variation in the size of the premium discount.

In 2013, about 10% of the Dutch population chose a voluntary deductible (see Appendix 9.F). About 80% of the people who chose a voluntary deductible in the previous year, also chooses a voluntary deductible the next year. The uptake of the voluntary deductible in the Netherlands is relatively low considering that Van Winssen et al. (2015) and Handel et al. (2022) show that the voluntary deductible is financially profitable for about half of the population. The latter paper argues that default effects are the most likely explanation for the low uptake.

The deductible applies to nearly all health services in the basic benefit package, such as hospital care, physiotherapy and pharmaceutical care. Only primary care, maternal care, obstetric care, and GP care are exempted.[15] The basic package is determined each year by the government and its coverage is the same for all citizens.[16] There have been changes in coverage over time. We will discuss the changes relevant for this study in this paper. Remmerswaal et al. (2019) provide an extensive summary table of all the changes in the basic benefit package and other policy changes during the period of our study.

Individuals can choose to buy supplementary insurance to cover healthcare that is not part of the basic benefit package. Supplementary insurance is always offered in addition to the mandatory insurance of the basic benefit package, as the coverage of care is very different.[17] Examples of care covered by supplementary insurance are alternative medicine, glasses, contact lenses and cosmetic surgery. Supplementary insurance is therefore an addition to regular insurance, not a substitute. It is offered independently from the basic package, which means that persons are not required to buy regular and supplementary insurance from the same insurer.[18] We focus on the basic insurance market and do not consider supplementary insurance. Re-insuring the mandatory deductible is allowed in very special circumstances, for example for seasonal workers or people with a very low income. However, less than 1.5% of the population has re-insured his or her deductible (Dutch Healthcare Authority, 2014).

4. Data and descriptive statistics

4.1. Data

Proprietary healthcare claims data from Vektis are used for this paper.[19] The data cover all, roughly 17 million, insured inhabitants in the Netherlands between 2006 and 2013. After cleaning and excluding the years 2006 and 2007 (when a rebate, not a deductible, was in place), we retain 97 million observations.[20]

[13] Remmerswaal et al. (2019) compare the effect of the rebate (which was in place in 2006 and 2007) and deductible in Dutch healthcare. They show that a deductible causes a larger reduction in healthcare expenditure than a rebate of similar magnitude.

[14] The mandatory deductible was increased further in 2014 to 365 euros, 375 euros in 2015, 385 euros in 2016, and 385 euros in 2017, 2018, 2019, and 2020 as well. As we only have data for 2008 up to 2013, Table 1 shows the deductibles for these years.

[15] These cost categories comprise less than 10% of healthcare expenditure.

[16] There exist small differences in basic benefit packages, not in terms of coverage but provider networks can differ between insurers (Dutch Healthcare Authority, 2014).

[17] One exception is physiotherapy which is often part of supplementary insurance but also partly covered in the basic benefit package. It is however only covered for persons who need physiotherapy because of a chronic condition and they still have to pay the first few visits themselves. In the robustness analyses, we will come back to this.

[18] Over 85% of the Dutch population bought supplementary health insurance in our data period (Dutch Healthcare Authority, 2014).

[19] Vektis is a private organization that gathers and manages data for all Dutch health insurers. The data are pseudonymized and not publicly available.

[20] The data cleaning steps are described in Appendix 9.D. The same data and a similar data preparation are used in Remmerswaal et al. (2019).

The data include total annual healthcare expenditure under the Health Insurance Act for each individual in the Netherlands.[21] The timing of spending within the year is unknown to us, as only annual expenditure data are available. In addition, we know the composition of total healthcare expenditure by healthcare category, such as hospital care, dental care, physiotherapy, et cetera.[22]

We can follow individuals over time and observe their individual's annual choice of a voluntary deductible. Based on these choices, we construct a binary variable which is 1 if an individual chooses a voluntary deductible at least once in our data, and 0 if this is not the case.[23] Note that this binary variable, unlike the voluntary deductible choice itself, can also be 1 for a person under 18 years old, if he or she has chosen a voluntary deductible (in our data set) after turning 18. Based on this binary variable, we create two groups: group 'at least once a voluntary deductible' and group 'never a voluntary deductible'. In our analysis we exploit the exogenous variation in healthcare expenditures of these two groups just before and after they turn 18 (when the deductibles kick in).

Our data also include person characteristics such as sex, age, indicators of chronic use of care and medication, and a four digit postal code. Age is reported in years and for December 31st in every year.[24] DCG stands for diagnosis cost group ('diagnosekostengroep') and is a binary variable that indicates whether a person is chronically ill and had high healthcare costs in the previous years.[25] PCG, an abbreviation of pharmaceutical cost group ('farmaciekostengroep'), indicates whether a person uses medication chronically. Using the four digit postal code, we can link to each observation the average standardized disposable household income in a postal code area from Statistics Netherlands.[26]

All persons with mental healthcare expenditure between 2008 and 2013 are excluded, because the mental healthcare sector faced additional changes in cost-sharing in 2012 (see Remmerswaal et al., 2019 for a list of policy changes). Between 2008 and 2011, dental care coverage also changed in a different way for persons above and below 18. We do not delete all individuals who use dental care, because almost all inhabitants use dental care and dental costs are low. Therefore, we do not include dental expenditure in our dependent variable: healthcare expenditure under the deductible. We verify in the robustness analyses that these choices do not affect our results. The main dependent variable in this study is total healthcare expenditure for healthcare services for which the deductible applies, but without dental healthcare costs. From hereon, we will refer to this variable as healthcare expenditure under the deductible. See Appendix 9.G for details on the cost categories that are included in this variable.

In our main analyses we follow Remmerswaal et al. (2019) and select all persons aged 15 to 21, but 18 year olds are excluded. The argument for excluding 18 year olds is that the exact date of birth is not available in the data. Therefore, within a year we cannot differentiate a person who turns 18 on January 1st from a person who turns 18 on December 31st (the former person has the deductible for almost the entire year while the latter does not face a deductible at all). By removing 18 year olds from the sample, we reduce the possibility of anticipation and substitution effects. Note that for our additional policy analysis we extend the sample to the full population (see Section 7).

To sum up, for our main analysis we study young adults between 15 and 21 years old (18 year olds are excluded), who did not have any mental healthcare expenditures in the period 2008 up to 2013. We compare two groups, people who did not choose a voluntary deductible at all and people who did choose a voluntary deductible at least once in our dataset, and follow them before and after they turn 18 years old.

4.2. Descriptive statistics

Table 2 describes the baseline sample of our main analysis, divided into 15 to 17 year olds (who do not face a deductible) and 19 to 21 year olds (who do face a deductible), and persons who never chose a voluntary deductible (group *never*) and persons who chose a voluntary deductible at least once (group *at least once*).

Our main variable of interest is healthcare expenditure under the deductible. This variable equals expenditures on the cost categories for which an individual (above 18) pays out-of-pocket until the deductible is exhausted. Appendix 9.G specifies these categories. The categories also apply to people below age 18, although they do not pay anything out-of-pocket.

Individuals in group *never* have, as expected, higher healthcare expenditure (under the deductible) than those in group *at least once*: healthcare expenditure of 19 to 21 year olds in the former group is on average 589 euros, whereas for the latter it is 322 euros which is 267 euros less. Looking at 15 to 17 year olds, we see that this difference in healthcare expenditure already manifests before the deductibles kick in: healthcare expenditure of 15 to 17 year olds in group *never* is on average 557 euros, 233 euros higher than

[21] The data do not suffer from underreporting of healthcare claims. It is a common problem with claims data that people with little healthcare expenditure do not bother to claim their bill to their insurer, because they do not expect to be compensated as they will not exceed their deductible. In our data however, healthcare providers are motivated to report *all* costs directly to patients' health insurers: the providers are only reimbursed if they report the costs to their patient's health insurer. Healthcare providers send their bills to the insurer electronically, who subsequently will bill the patient (if the deductible is not exhausted).

[22] Appendix 9.G includes a list of all categories.

[23] To construct these groups we used additional information from 2006 and 2007 and coded an individual with '1' if he or she chose for a voluntary deductible in 2006 or 2007 as well.

[24] A person who becomes 18 on December 1st in 2013 is classified as being 18 years old in 2013, even though he or she was 17 years old for 11 months that year.

[25] DCG and PCG are variables from the Dutch risk adjustment system, which aim to identify chronic disorders that will lead to high healthcare expenditures.

[26] Average standardized disposable household income is gross household income from which taxes and premiums for public insurance policies have been deducted. It has been standardized for differences in size and households compositions. In our data set, there are on average 3130 persons per four digit postal code.

Table 2
Summary statistics of the baseline sample.

	Never a voluntary deductible		At least once a voluntary deductible	
	15–17 years	19–21 years	15–17 years	19–21 years
Persons with any healthcare expenditures (frac.)	0.75 (0.43)	0.78 (0.41)	0.71 (0.46)	0.70 (0.46)
Persons who exhaust mandatory deductible (frac.)	–	0.35 (0.48)	–	0.26 (0.46)
Healthcare expenditure under deductible (€)	557 (3,719)	589 (3,294)	324 (1,315)	322 (1,672)
Of which:				
Hospital care (€)	373 (3,211)	436 (2,866)	219 (1,098)	247 (1,546)
Physiotherapy (€)	40 (156)	8 (110)	28 (115)	4 (69)
Pharmaceutical care (€)	89 (1,289)	97 (1,101)	51 (497)	49 (313)
Other care (€)	55 (679)	48 (506)	25 (245)	21 (217)
Age (years)	16 (0.82)	20 (0.82)	16 (0.79)	20 (0.81)
Male (frac.)	0.51 (0.50)	0.52 (0.50)	0.53 (0.50)	0.54 (0.50)
Diagnosis cost-related group (frac.)	0.01 (0.09)	0.02 (0.13)	0.00 (0.06)	0.01 (0.08)
Pharmaceutical cost-related group (frac.)	0.02 (0.15)	0.03 (0.16)	0.01 (0.10)	0.01 (0.10)
Household income quintile	3.21 (1.37)	2.92 (1.43)	3.39 (1.30)	3.00 (1.45)
Average deductible level (€)	0 (0)	201 (70)	0 (0)	398 (247)
Number of observations	2,521,889	2,369,686	319,827	700,922

Notes: Standard deviations are reported between parentheses. The share of persons who exhaust their mandatory deductible is not reported for 15 to 17 year olds as they do not have a deductible. Healthcare services category 'other care' combines paramedical care, medical aids, transportation costs of persons lying down and for seated persons, care that is provided across the Dutch borders, geriatric revalidation and other healthcare costs which are not part of the cost categories in Appendix 9.G. Household income quintiles range from 1 to 5, where Quintile 1 refers to the households with the lowest incomes and household Quintile 5 to households with the highest incomes. The differences between all characteristics of 15–17 year olds and 19–21 year olds of group 'Never a voluntary deductible' are all significant at a 1% significance level. These extremely small p-values are a result of the large sample size: even very small differences are highly significant.

healthcare expenditure of 15 to 17 year olds in group *at least once*. The difference of 233 euros suggests that there is a selection effect, as this difference cannot be caused by deductibles. The fraction of individuals with any healthcare expenditures is slightly – 4 percentage points for 15 to 17 year olds – higher for group *never* than group *at least once*, which suggests that the difference in healthcare expenditure between the two groups is mostly at the intensive margin. As one would expect, a higher fraction of people in group *never* exhaust the mandatory deductible than in the group *at least once*.

This table presents already, in a 'crude form', the main result of the paper. The 267 euros difference in expenditures for 19 to 21 year olds is made up of a selection and a moral hazard effect. The difference of 233 euros for 15 to 17 year olds is a pure selection effect as none of them faces a deductible. Hence, the biggest part of the difference in 19 to 21 year old expenditures between people without and with a voluntary deductible is due to a selection effect.

Table 2 however does not present first evidence of a reduction of healthcare expenditures due to the deductible kicking in at age 18, as average healthcare expenditures are roughly the same for 15 to 17 year olds (who do not face a deductible) and 19 to 21 year olds (who do face a deductible) in both groups.[27]

Hospital care accounts for most of healthcare expenditure, and physiotherapy for a small part. Average costs of physiotherapy are substantially lower for persons aged 19 to 21 compared to persons aged 15 to 17. This could indicate an effect of the mandatory deductible kicking in. All four groups are relatively healthy (because they are young): only between zero and 2% is a chronic user of healthcare (i.e. classified with a DCG) and between 1 and 3% is a chronic user of medication (i.e. classified with a PCG). Persons who chose a voluntary deductible exhaust their mandatory deductible less often than persons who never chose a voluntary deductible. Persons who chose a voluntary deductible are more often male. The average household income quintile is slightly higher for 15 to 17 year olds compared to 19 to 21 year olds, and lowest for 19 to 21 year olds who never chose a voluntary deductible.

The average deductible level for 19 to 21 year olds in group *never* between 2008 and 2013 is 201 euros and 398 euros for persons in group *at least once*.[28] Group *never* is considerably larger than group *at least once*, because only a tenth of the Dutch population has chosen a voluntary deductible.[29] Furthermore, we have a limited time period of our data. It is possible that someone will choose a voluntary deductible for the first time in 2014. In our analysis, we would incorrectly classify this person as never having chosen

[27] To see the effect of the deductible on healthcare expenditures, we must look at average healthcare expenditure for each age, as we do below in Fig. 1.

[28] Appendix 9.E shows the voluntary deductible choice of 19 to 21 year olds in group *at least once* between 2008 and 2013. A voluntary deductible of 0 and 500 euros is most common, with on average 50 and 29% respectively. The majority of the group did not choose a voluntary deductible each year when they turn 18. However, the 500 euro deductible gains popularity over time. The 201 euros average mandatory deductible is close to the unweighted average (200 euros) of 150, 155, 165, 170, 220 and 350 euros.

[29] The table shows that the number of observations aged 19 to 21 year old among those who chose a voluntary deductible at least once, is more than twice the number of observations aged 15 to 17 year old. This is a result of the available data and the definition of the group *at least once*: to include a person in the group *at least once*, he or she must choose a voluntary deductible when he or she is 18 or over. If a person turns 18 years old in 2013 and decides not to choose a voluntary deductible in that year, then he or she is not included in the group *at least once*, even though he or she may decide to choose a voluntary deductible the next year. As such, the probability of being included in the group *at least once* is lower for younger persons in the later years of the dataset.

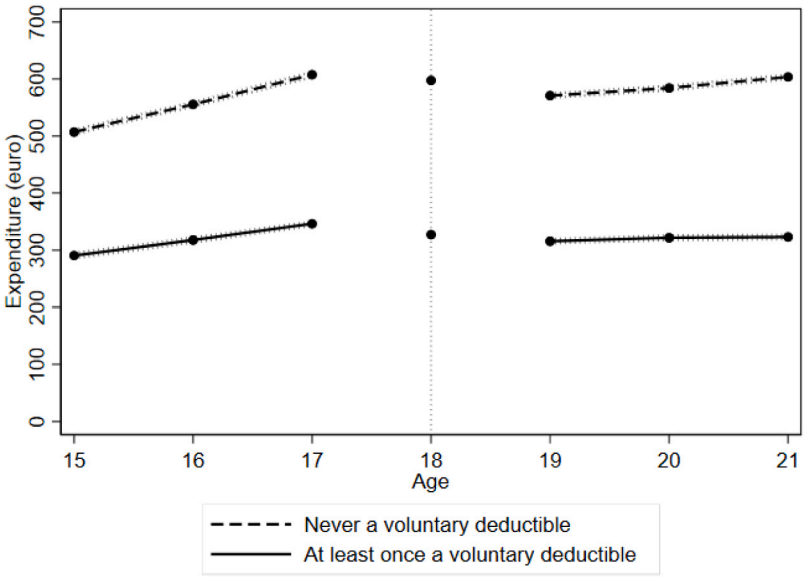

Fig. 1. Mean healthcare expenditure for persons who have never chosen a voluntary deductible or have chosen a voluntary deductible at least once.

a voluntary deductible. This data limitation implies an underestimation of the selection effect: some healthy individuals with low healthcare expenditure are included in group *never*.[30]

Fig. 1 demonstrates further descriptive evidence of selection and moral hazard effects between the two groups.[31,32] The line of average healthcare expenditure of persons in group *at least once* is well below the line of those in group *never*, before and after the deductibles kick in. Furthermore, when extrapolating the lines of 15 to 17 year old and of 19 to 21 year old to age 18, we see that healthcare spending drops at age 18. This is evidence of the effect of the deductible kicking in at 18. In the next section, we formalize this further in our empirical approach.

5. Empirical strategy

The tables and graph in the previous section show first descriptive evidence of selection and moral hazard effects. Here we describe our empirical strategy.[33] Fig. 2 is an adaptation of Fig. 1 and explains the idea of our identification strategy which aims to disentangle selection and moral hazard effects for 18 year olds. First, we extrapolate the healthcare expenditure trend of 15 to 17 year olds up to age 18 for persons in group *never*. This extrapolation is illustrated with a dotted line in Fig. 2. This extrapolated line crosses the vertical line at age 18. This point is denoted by y_{never0}. Hence, y_{never0} shows the average healthcare expenditure for an 18 year old who will never choose a voluntary deductible *in absence of any deductible*. In terms of the model in Section 2, we write $y_{never0} = \theta_h + \eta_h(1)$.

Next, we extrapolate the healthcare expenditure trend of 19 to 21 year olds in group *never* to age 18. y_{never1} denotes healthcare expenditure of an 18 year old who never chose a voluntary deductible *but who faced a mandatory deductible*; $y_{never1} = \theta_h + \eta_h(q_h)$. Therefore, $y_{never1} - y_{never0} = \eta_h(q_h) - \eta_h(1)$ is the effect of the deductible kicking in on healthcare expenditure for persons who never chose a voluntary deductible.[34] We repeat the same steps for persons in group *at least once*. $y_{at\ least\ once0}$ denotes healthcare expenditure for an 18 year old who will choose a voluntary deductible at one point in time, but in absence of a deductible and $y_{at\ least\ once1}$ is the same but when facing a deductible. Lastly, $y_{at\ least\ once1} - y_{at\ least\ once0} = \eta_l(q_l) - \eta_l(1)$ is the moral hazard effect of the mandatory plus voluntary deductible.

$y_{never0} - y_{at\ least\ once0} = \Delta C$ is the selection effect for an 18 year old, as it is the difference in healthcare expenditure between the two groups at age 18, but in absence of any deductible. $y_{never1} - y_{at\ least\ once1} = \Delta C + \Delta MH$ is the selection effect plus the difference in the moral hazard effects of the deductible between the two groups.[35] Note that in most studies, one only observes

[30] We also estimate the selection effect for the more restrictive group of individuals who *always* chose a voluntary deductible in our time period; see Section 6.

[31] Standard errors of expenditures are shown in Fig. 1, but as these standard errors are very small, they are barely visible. They differ from standard deviations of individual expenditures as reported in, for example Table 2.

[32] See Appendix 9.H and 9.I for figures with wider age bandwidths.

[33] Note that in this section we describe our empirical strategy for 18 year olds. Our extension to the entire population for policy analysis can be found in Section 7.

[34] Note that other factors may change as well at 18 which can also lead to a drop in healthcare expenditure. We will address these shortly (Eq. (8) below). To simplify the exposition, first we assume the only change at 18 is the moral hazard effect of the deductible.

[35] We label the moral hazard effect of the deductible as MH_{never} and $MH_{at\ least\ once}$, where the former is $y_{never1} - y_{never0}$ and the latter is $y_{at\ least\ once1} - y_{at\ least\ once0}$. $y_{never1} - y_{at\ least\ once1}$ can be rewritten as $(y_{never0} + MH_{never}) - (y_{at\ least\ once0} + MH_{at\ least\ once})$. The selection effect is $y_{never0} - y_{at\ least\ once0}$.

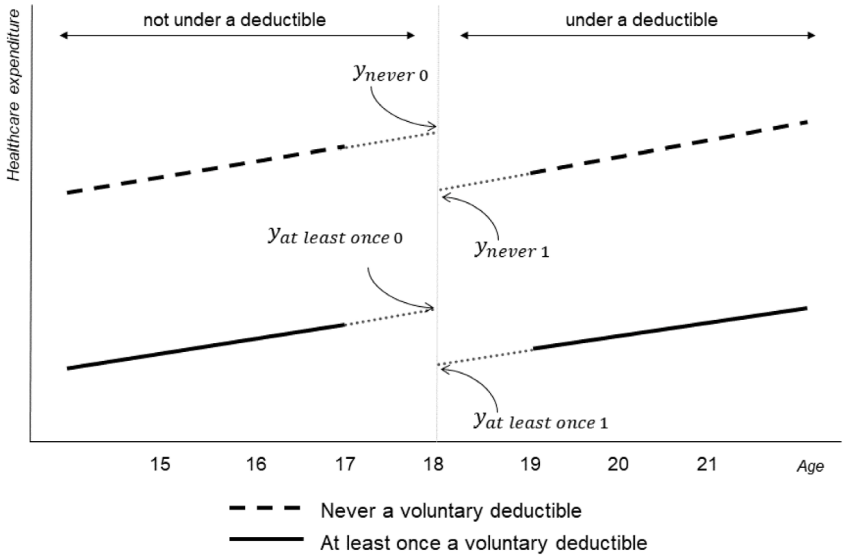

Fig. 2. Graphical illustration of identification strategy.

$y_{never1} - y_{at\ least\ once1}$, which is a mixture of the selection and moral hazard effects. The Dutch setting however offers an opportunity to directly measure a selection effect: $y_{never0} - y_{at\ least\ once0}$.

This identification strategy can be formalized in a design in the spirit of a panel regression discontinuity design (Thistlethwaite and Campbell, 1960).[36] With this design, we can estimate the causal effects of the deductibles on healthcare expenditure.

Regression discontinuity designs exploit discontinuities in assignment variables. In this case, age is the running variable: when a person's age is known, then it is also known whether he or she receives the treatment:

$$T = \begin{cases} 1 & \text{if } age \geq 18 \\ 0 & \text{if } age < 18 \end{cases}$$

where T is a binary variable which indicates whether an individual receives the treatment or not.

A necessary assumption for the regression discontinuity design is that persons cannot influence the running variable (Lee and Lemieux, 2010). As persons cannot manipulate their age in administrative data, this assumption holds. Furthermore, all factors that potentially affect healthcare expenditure must evolve smoothly with age, i.e. the only shock that may occur at 18 is the introduction of the deductible. This assumption will not hold here: at 18, many students in the Netherlands graduate from high school, go to university and move out of their parental house. This is likely to affect healthcare spending at 18, and is therefore a threat to the validity of a regression discontinuity design. Remmerswaal et al. (2019) show that the share of young adults moving and going to university indeed increases at 18. However, they also show that the sizes of these shocks (and several others which may invalidate the design) are constant over time. Aside from the deductible kicking in at age 18, our data offer another useful source of exogenous variation: the Dutch government has increased the size of the mandatory deductible level over the years (see Table 1). In our design we therefore estimate the effect of a *change* in the deductible size. Hence, by exploiting this variation over time, we take out all constant changes at age 18, including moving out of the parental house. As a result, we can identify the marginal effect of an *increase* in the deductible, but not the mere effect of the deductible kicking in at 18.[37]

Eq. (8) formulates the design for estimating the effects for group *never* and group *at least once* simultaneously:

$$y_{it} = \alpha_{gt} + \alpha_i + \beta_g a\tilde{g}e_{it} + \gamma_g T_{it} a\tilde{g}e_{it} + \tau_g T_{it} + \nu n_t T_{it} I_N + \lambda a_{it} T_{it} I_{ALO} + \epsilon_{it} \tag{8}$$

[36] The idea of the identification of the effects – extrapolating the relationship between age and healthcare expenditure to the discontinuity under both regimes – is very much like a regression discontinuity design, but our design also differs from a 'traditional' regression discontinuity design, in the sense that we do not have many values of the running variable (age) and add a panel dimension with time and individual fixed effects. That is why we prefer to call the design in the 'spirit' of a regression discontinuity design. In the paper we go through the assumptions necessary for a regression discontinuity design.

[37] Fig. 2 suggests that we do identify a level moral hazard effect of the deductible kicking in at 18. However, this is not the case as we can only identify marginal effects of changes in the deductible. Appendix 9.K is an adaptation of Fig. 2 and illustrates how we exploit the annual variation in the deductible to identify the marginal effect of the deductible. The drop in healthcare expenditures at age 18, i.e. the reduction of the moral hazard effect due to the deductible, is larger in year t_1 than in year t_0 because deductible level $D_2 > D_1$. The difference between the moral hazard effects in both years can be attributed to the difference between deductible levels D_1 and D_2. Note that we *do* estimate a level effect of becoming 18 years old as well. However, the issue here is that we cannot separate the level effect of the deductible kicking in – in which we are interested – from the other level effect when turning 18 years old, such as moving out of the parental house. Parameter τ_g from model (8) is a mix of both level effects. Therefore we only focus on the marginal effects of the deductible, which we can identify properly.

We include individuals i in periods $t \in \{2008, 2009, 2010, 2011, 2012, 2013\}$ where $age_{it} \in \{15, 16, 17, 19, 20, 21\}$. y_{it} denotes healthcare expenditure under the deductible of individual i in period t. α_{gt} are time fixed effects. The index g indicates that we allow the time fixed effects to vary for group *never* and *at least once*. α_i are individual fixed effects. Age is centered to 18 and denoted as $a\tilde{g}e_{it} = age_{it} - 18$. It captures the linear trend between age and healthcare expenditure (see Fig. 1).[38] $T_{it} = 1$ if $age_{it} \in \{19, 20, 21\}$, and zero otherwise. Hence, γ_g allows this linear trend to be different for persons before and after they turn 18. The coefficient of age can vary for groups *never* and *at least once*, as is denoted by the index g in β_g, γ_g. Parameter τ_g captures the effect of becoming 18 on healthcare spending, which can differ for the two groups.

We assume a linear relationship between the moral hazard effect and the size of the deductible. That is, the range of deductibles in our data is relatively small (individuals have no deductible or a deductible between 150 and 850 euros) and thus a linear approximation is not unreasonable.[39] n_t denotes the size of the mandatory deductible in period t for persons who never chose a voluntary deductible (who are indicated by I_N); ν is one of the main parameters of interest as it captures the effect of a marginal increase in the mandatory deductible for persons who never chose voluntary deductible. For persons in group *at least once*, indicated by I_{ALO}, a_{it} denotes the level of the (mandatory plus voluntary) deductible for a person i in period t who chose a voluntary deductible at least once (in our data period). Because it is the total deductible size (mandatory plus the voluntary), it varies across both person i and year t. This is because people can choose and change their voluntary deductible size.[40] λ captures the effect of a marginal increase in the total deductible size for persons who chose a voluntary deductible at least once.

The panel structure of the data allows us to include individual fixed effects (α_i) and time fixed effects (α_{gt}). Individual fixed effects control for time invariant differences in health status or personality traits (e.g. risk attitude) which might be related to both choosing a voluntary deductible and healthcare expenditures. Time fixed effects control for changes over time that affect persons' behavior above and below 18, for example changes in basic benefit package coverage. The standard errors ϵ_{it} are clustered at the individual level in all models to correct for correlation (Lee and Lemieux, 2010). Clustering the standard errors differently, for example by age or age times year, does not change our results; it leads to (even) smaller standard errors for group *never* and similar standard errors for group *at least once* (see Appendix 9.M).

As discussed above, we identify parameter ν assuming that there are no other changes between 2008 and 2013 when people turn 18 years old and which could affect healthcare spending: ϵ_{it} is independent of n_t.[41] We identify parameter λ assuming that an individual's choice of the voluntary deductible, which is part of variable a_{it}, does not depend on a change in unobserved factors over time. The latter assumption is quite restrictive: it is possible that a person changes the level of his or her voluntary deductible when he or she suffers from a health shock. However, we argue this is not an issue here. Few 19 to 21 year old individuals seem to strategically increase and lower their choice of the voluntary deductible in response to a health shock. In fact, most of the individuals in group *at least once* choose a voluntary again, once they have chosen it (see Section 3). Moreover, strategic behavior along these lines tends to increase λ while we basically find $\lambda = 0$ for this limited age group.

6. Results

This section presents the results of the baseline model and discusses a number of robustness analyses. Technical details can be found in the Appendix.

6.1. Baseline model

The results of model (8) are presented in Table 3.[42] For persons who never chose a voluntary deductible, we find a ν coefficient of −0.26 which is statistically significant at a 1% significance level. To illustrate the economic significance of this estimate, a coefficient of −0.26 means that for 18 year olds who have never chosen a voluntary deductible, a 100 euro increase in the mandatory deductible reduces healthcare expenditure on average by 26 euros per person. This corresponds to a deductible elasticity of −0.09.[43]

[38] Remmerswaal et al. (2019) show that the hypothesis of a linear relationship between healthcare spending and age cannot be rejected for this age range and for persons who never chose a voluntary deductible.

[39] We could also add an interaction of $a\tilde{g}e_{it}$ with year t or with n_t and a_it, to allow the slope of age to vary over time or to vary with the size of the deductible. It is possible that healthcare expenditure increases less steeply with age in years in which the deductible size was higher. However, adding such interactions would be quite demanding for the model and data, considering that we only look at three age years before and after becoming 18 years old. We have estimated model (8) with such interactions and find similar results to the results of the baseline models. However, the coefficients have larger standard errors because they are estimated less efficiently.

[40] This in contrast to the mandatory deductible which – for people above 18 – does not vary across individuals in a given year.

[41] We believe this assumption is credible, it would be violated if for example new technology, changes in benefit packages or policy rules were implemented that hits healthcare expenditure for particular ages in the range 15 to 21 years old in different ways. However, we are not aware of such changes.

[42] In Appendix 9.N we present and discuss the results of estimating model (8) across men and women and income quintiles.

[43] We calculate the elasticity as follows:

$$\epsilon_n^y = \frac{\partial y}{\partial n}\frac{\bar{n}}{\bar{y}_{never}} = -0.26\frac{201}{589} = -0.09 \tag{9}$$

where $\bar{y}_{never}$ is the average healthcare expenditure of 19 to 21 year olds who never chose a voluntary deductible (Table 2), $\bar{n}$ is the average (mandatory) deductible size. Note that this elasticity is not a price elasticity. Price in "price elasticities" in health insurance tends to refer to co-payments or co-insurance rates. To avoid confusion, we use the term deductible elasticity.

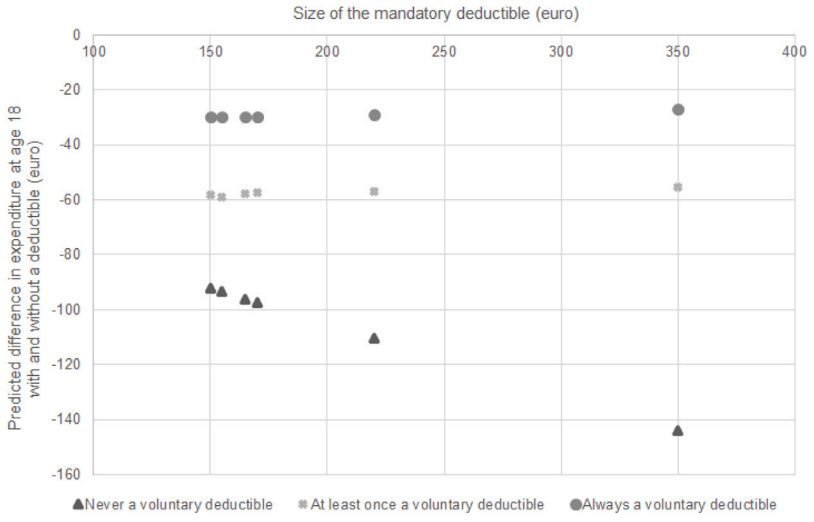

Fig. 3. Difference in predicted healthcare expenditure at age 18.

Table 3
Deductible and selection effects for 18 year olds.

	Never a voluntary voluntary deductible	At least once a voluntary deductible
Marginal effect of €1 increase in total deductible size (ν, λ)	−0.26*** (0.07)	0.00 (0.02)
Deductible elasticity	−0.09	0.00
Selection effect (€)	–	339.60 (0.64)

Notes: Standard errors are reported between parentheses and clustered at the individual level. *, **, and *** indicate significance based on a two sided test at the .10, .05, and .01 levels, respectively. Coefficients ν and λ are estimated with model (8). The full results of the estimation are presented in Table 9.L.

The size of the deductible elasticity is roughly two times smaller than the price elasticity from the RAND experiment (Newhouse and the Insurance Experiment Group, 1993). This may be a result of institutional differences, Dutch healthcare has GPs as gatekeepers for example, or differences in demographics and healthcare use between the United States and the Netherlands. The level of the deductible is also much lower in the Netherlands compared to the United States. Lastly, our deductible elasticity applies to a very specific age category: young adults around 18 years old.

The estimated coefficient λ for 18 year olds who have chosen a voluntary deductible at least once is 0.00 and not statistically significant at a 10% significance level. But λ is significantly different from −0.26, the value we find for ν. After controlling for age effects, individual and time fixed effects, a marginal increase in the level of the deductible does not reduce healthcare expenditure of this group significantly. Two intuitive reasons for this are the following. First, healthcare expenditure for many people in this group is well below their deductible level. As a result, they are not marginally affected by an increase in the total deductible size. Second, a zero deductible elasticity also suggests that the healthcare they *do* consume is both necessary and valuable, such as an appendectomy, as they accept these treatments regardless of their deductible size.

Note that the coefficient of zero does not imply that persons who chose a voluntary deductible at least once do not respond to the deductible at all. In fact, in Fig. 1, we observe a drop in healthcare expenditure at 18. This reduction in healthcare expenditure can be the effect of the deductible kicking in, as well as other changes at 18. But this deductible effect would not be related to the size of the deductible as that is captured by coefficient λ (which happens to be zero).

Fig. 3 shows the difference in predicted healthcare expenditure for 18 year olds under a deductible and not under a deductible for each year, that is, it shows the difference for example between $\hat{y}_{never1}$ and $\hat{y}_{never0}$ from Fig. 2 for persons who never chose a voluntary deductible. As expected from finding a negative coefficient for ν, the difference becomes larger as the deductible size increases. This is not the case for persons who chose a voluntary deductible at least once: the difference in healthcare expenditure at age 18 (i.e. $\hat{y}_{at\ least\ once1} - \hat{y}_{at\ least\ once0}$) is negative, but it does not become bigger over time.

If we estimate model (8) without individual fixed effects (i.e. ordinary least squares) then we get a value for λ equal to −0.21 (see Table 9.L in the appendix). This negative coefficient reveals the existence of selection effects: within the *at least once* group, people with a higher deductible tend to have lower costs, but without controlling for selection effects (by including individual fixed effects α_i) this is picked up as a deductible effect.

Table 4
Other comparison groups.

	Always a voluntary deductible	Always a voluntary deductible of €500
Marginal effect of €1 increase in total deductible size (λ)	0.01 (0.08)	0.31 (0.62)
Selection effect (€)	462.90 (1.26)	534.30 (1.56)
Observations	209,249	140,569

Notes: Standard errors are reported between parentheses and clustered at the individual level. *, **, and *** indicate significance based on a two sided test at the .10, .05, and .01 levels, respectively. Coefficients λ is estimated with model (8). The selection effect is determined compared to individuals who never chose a voluntary deductible. The full results of the estimations are available upon request.

The average selection effect is 339.60 euros.[44] This result indicates that mainly selection, not moral hazard, explains the difference in healthcare expenditure between persons in plans with and without a voluntary deductible. A large selection effect is intuitive as only a small, highly selective group of individuals chooses the voluntary deductible. A selection effect of 340 euros is more than the observed 267 euros difference in healthcare expenditure between 19 to 21 without and with a voluntary deductible (reported in Table 2). Therefore, it must be the reduction in healthcare spending due to the mandatory deductible that makes the difference in healthcare spending between the two groups smaller. That is, persons in group *never* (with higher expected expenditures) reduce their expenditure more in response to the deductible than persons in group *at least once*. It is also more than the 233 euros average difference in healthcare expenditure between 15 to 17 year olds, which seems in line with the diverging trends in healthcare expenditure for persons in the two groups below 18 years old in Fig. 1.

The main analysis compares individuals who never chose a voluntary deductible to persons who chose one *at least once*. If we narrow the latter group to individuals who *always*, i.e. each year, chose a voluntary deductible we find a similar coefficient of λ: an increase in the total deductible size does not affect healthcare spending significantly (see Table 4).[45] The selection effect has become larger with 462.90 euros.[46] When selecting only persons who always chose the maximum voluntary deductible of 500 euros, the selection effect becomes even larger: 534.30 euros. The estimated λ coefficient, is large, but insignificant.[47]

We interpret these findings as follows. When moving from the group who chose a voluntary deductible at least once to the group who always chose a voluntary deductible and finally to the group who always chose the highest voluntary deductible, we move to people who are (expected to be) progressively more healthy. Hence, one would expect the selection effect to increase moving from the first group to the last. Indeed, this monotone increase is what we find in the data.

6.2. Robustness analyses

To test whether our results are driven by our specification choices, we have performed several additional analyses. The results and in some cases the estimation methodology are presented in the appendices. Here we describe the main findings.

First, we test if the difference between coefficients λ and ν is driven by the fact that persons in group *never* and group *at least once* are not at the margin in the same way. It is possible that the level of the mandatory plus voluntary deductible is so high and the expenditure levels of persons in group *at least once* so low, that there are few individuals affected by the deductible at the margin. If that is the case, the zero response to the deductible change would be the result of the fact that they are simply not affected by the change in the deductible size, whereas individuals in group *never* do respond because they are affected at the margin by a change in the deductible size. It could also be that persons with a voluntary deductible are simply so healthy, they do not have any healthcare consumption they can reduce when their deductible increases. In Appendix 9.O however, we perform two different analyses to control for these issues and to make sure the coefficients λ and ν are comparable: in the first analysis we use the end-of-year price instead of the deductible size as the main regressor and the second analysis we exclude the healthiest individuals who we expect will not respond to a change in the deductible size. The results of both analyses are similar to our main results.

With the second robustness analysis (see Appendix 9.P) we show that we find the same results when we divide the sample differently and compare persons who are *high risk* to persons who are *low risk* based on spending patterns when aged under 18. A potential disadvantage of our approach of comparing persons who chose a voluntary deductible at least once to persons who never chose a voluntary deductible is that persons are assigned to groups based on an individual endogenous choice (i.e. the choice of voluntary deductible). To check whether this endogeneity affects our results, we discard group *never* and group *at least once* and divide the sample into two other groups: group *high risk* and group *low risk*. A person is considered low risk (and consequently assigned to group *low risk*) when he or she had low healthcare expenditures (less than 200 euros) each year before turning 18 years old. Similarly a person is considered high risk (and is assigned to group *high risk*) when he or she had high healthcare expenditures (more than 500 euros) each year before turning 18 years old. As shown in Table 8, we find a much larger coefficient for *high risk*

[44] See Section 5 and Footnote 35 for the definition of the selection effect.

[45] λ is the same, 0.01, if we select persons who sometimes choose a voluntary deductible, but not each year (i.e. not always).

[46] Compared to persons who never chose a voluntary deductible.

[47] We can also estimate model (8) for persons who sometimes choose a voluntary deductible, but not always. These persons are in group *at least once*, but not in group *always*. λ is 0.00 as well and as expected the selection effect is smaller with 296.94 euros.

persons than for *low risk* persons. This result is the same when we only keep persons who never chose a voluntary deductible within groups *high risk* and *low risk*, thus the difference in coefficients is not driven by the level of the deductible chosen in the two groups (see Table 9).[48] Finally, we ran the regression with the mandatory deductible only (and we ignore the voluntary deductible) for each group. Again we find (not separately reported) that the coefficient for the *high risk* group is bigger (in absolute value) than for the *low risk* group. The result of the latter regression is now a combination of the *low risk* group reacting less to a change in deductible and relatively many low risk agents taking a voluntary deductible and hence not being on the margin of the mandatory deductible.

Our interpretation is that low risk people (with low healthcare use in the past) can expect low costs in the future (after turning 18). When costs are low, there is little opportunity to change healthcare use in response to a change in the deductible. The healthcare that they use, has high value to them (e.g. in response to breaking a leg). The high risk group uses more healthcare and has more opportunities to adjust this use in response to a change in the deductible. To illustrate, a patient receiving physiotherapy treatment can try to extend twenty sessions by one or two additional sessions that have lower (extra) value. But it is harder for someone who does not receive any physiotherapy at all, to arrange one or two sessions out of the blue. This is different from saying that someone with low expenditure is not at the margin and does therefore not respond to a change in the deductible.

Next we show that the age bandwidth choice of three years before and after the threshold does not affect the results. In Appendix 9.Q we show that for a two- or four-year age bandwidth, we find results very similar to a three-year age bandwidth.

We also check the effects of a group of chronically ill persons with annual high healthcare expenditures, who are unlikely to respond to an increase in the deductible, because they will exceed the deductible anyway. Excluding the chronically ill is therefore expected to increase the average moral hazard effect. We exclude persons with, before turning 18, healthcare expenditures in the highest decile.[49] As shown in Table 9.R, we find a much larger coefficient ν of −0.70 for persons in group *never*. That translates into a deductible elasticity of −0.26. We find a coefficient λ of −0.03, insignificant at a 10% significance level, for persons in group *at least once*. Furthermore, we find that excluding the persons with high healthcare expenditures lowers the selection effect.

A potential problem for our estimation is an anticipation effect for 17 year olds: they consume more care at age 17, while it is still 'free' because the deductible will kick in when they turn 18. Such anticipation effects can result in a bigger drop of healthcare expenditure at 18, and thus inflate our results. Therefore, we omit 17 year olds from the analysis. This does not affect our findings: we find a ν coefficient of −0.23 for group *never* and a λ coefficient of 0.00 for group *at least once*. The implied selection effect is 341.07 euros, as before.

We also test how sensitive our results are to excluding persons who at one point in time use mental healthcare. When we include them (while not including their mental healthcare expenditure), we again get similar results. Hence, our choice to exclude people with mental care expenditures does not affect our results.

Appendix 9.S presents the results of estimating model (8) for three different categories of healthcare expenditure separately: hospital care, physiotherapy, and pharmaceutical care. The results for hospital care are similar to the main results: the estimated coefficients ν and λ equal −0.26 and 0.01 respectively. This is not surprising as most of the expenditures are hospital expenditures. Furthermore, we find that the coefficients on physiotherapy are significant but small, ν is −0.03 and λ is −0.01. Again the effect is smaller for the group that chooses a deductible at least once. Because average expenditures on physiotherapy are very small, these small coefficients do translate into quite large elasticities. The coefficient for pharmaceutical care is small and insignificant. This means that pharmaceutical expenditures at 18 do not change with the total deductible size. This in contrast to other studies that find a negative elasticity for pharmaceutical care. Explanations for this include the following. As shown in Table 2, for our age categories' mean pharmaceutical expenditures are relatively low; below 100 euros and clearly below the mandatory deductible. Hence, they tend not to be marginally affected by a change in the deductible. Second, in the Netherlands pharmaceuticals have to be prescribed by a physician and hence tend to be used when they are actually effective. Finally, there could be an offset effect: as the deductible increases, people substitute toward drugs which tend to be cheaper than other treatments. Table 11 in appendix 9.V shows that pharmaceutical expenditure is elastic when considering a wider age range (ages 13 to 65). This suggests that the first explanation is the more convincing one.

Lastly, we reestimate model (8) with a different dependent variable: healthcare services for which the deductible applies, but without dental healthcare costs and without physiotherapy costs. In 2012 and 2013 the coverage of physiotherapy in the basic benefit package was limited: before 2012 patients had to pay the first 8 visits themselves, but in 2012 they had to pay the first 12 and in 2013 the first 20 visits themselves (see Remmerswaal et al., 2019). Such a change in the basic benefit package could be –incorrectly– picked up by our parameters ν and λ. Appendix 9.T shows that the results only become slightly smaller when we reestimate the model. This makes sense as physiotherapy is only a small fraction of total healthcare expenditure for 15 to 21 year olds (see Table 2).

In each of these robustness analyses we find that there is a significant reduction in expenditures in response to an increase in the mandatory deductible (for people who never chose a voluntary deductible). For people with a voluntary deductible, there is no significant response in healthcare expenditure to changes in the total (mandatory plus voluntary) deductible. We find a selection effect that varies between 201.62 and 377.72 euros.

[48] A similar analysis for the group choosing always a voluntary deductible has too few observations to be meaningful.

[49] We only construct deciles based on healthcare expenditure below 18, to make sure we do not select on the deductible effect. The deciles are calculated per year and age category and we only exclude individuals each year in the highest decile (before they turn 18).

7. Policy analysis

The results in the previous section show that the voluntary deductible leads to a substantial selection effect but not to an additional reduction of healthcare expenditure. In terms of our model in Section 2, the moral hazard effect is higher for the high-risk types than for the low-risk types. Hence, $q_h^* \leq q_l^*$ for 15–21 year olds. These results apply to people around 18 year old but may be different at the full population level. To draw conclusions for policy makers and politicians, we extend our analysis to persons of all ages in the Dutch population. We show how much healthcare spending will increase if the voluntary deductible is abolished and we show whether cross subsidies take place from high-risk individuals, for whom a voluntary deductible would be unprofitable, to low-risk individuals, who choose a voluntary deductible.

Δp, oop_l and MH_l

Estimating and separating selection and moral hazard effects for the full population is less clean from an econometric perspective than the analysis for 18 year olds from the previous sections. For policy analysis however, we argue the results are still informative because we only need an upper and lower bound on the moral hazard effect to derive our policy implications (see also our discussion in Appendix 9.V).

In Section 2 we explain how and when the voluntary deductible can lead to cross subsidies. Eq. (7) shows how the cross subsidy can be derived from the difference in insurance premium (between contracts without and with a voluntary deductible), the change in out-of-pocket expenditure and moral hazard due to the voluntary deductible.

To estimate the impact of the voluntary deductible on cross subsidies, we first compute the mean premium discount (Δp) for persons who chose a voluntary deductible. The exact premium discount each person received is not available in our data (see also Section 3). Instead, we assign the average premium discount by year and voluntary deductible level as published by the Dutch Healthcare Authority to each person in our data according to their choice of the voluntary deductible and the year (see Appendix 9.C). We compute their weighted average premium discounts to account for, for example, that there are more persons with a 500 euro voluntary deductible than a 400 euro deductible (see Appendix 9.F). Since we are interested in the aggregate effect, working with averages here is fine.

Next, we calculate the mean extra out-of-pocket payment ($oop_l(q_l) - oop_l(q_h^*)$) due to the voluntary deductible for each individual in our data given his or her choice of the voluntary deductible and healthcare expenditures. Note that this measure will be relatively low as it is zero for persons who do not exceed their mandatory deductible.

The last component we need is the mean reduction in moral hazard due to the voluntary deductible being in place ($MH_l(q_l) - MH_l(q_h^*)$). The results in the previous section show that 18 year olds who chose a voluntary deductible at least once do not respond to an increase in the deductible ($\lambda = 0$). This implies that the moral hazard term is zero. We will use this result as a lower bound because MH_l may be higher at the full population level. To determine an upper bound for MH_l, we estimate a similar model for the extended age range of 13 to 65 year olds.[50] For 13 to 65 year olds, we cannot use the same design as for 15 to 21 year olds. Therefore, we measure an average effect over all age categories from 13 to 65 with a panel regression design. Unlike model (8) here we do not include group *never* and include age dummies (α_{it}^{age}) which capture age specific effects which are not necessarily linear in age over this wider age range. We estimate:

$$y_{it} = \alpha_t + \alpha_i + \alpha_{it}^{age} + \lambda a_{it} T_{it} + \epsilon_{it} \tag{10}$$

where y_{it} is identical to y_{it} in model (8) and α_t and α_i are time and person fixed effects respectively.

In our empirical strategy in Section 5 we assume that time fixed effects α_t control for the changes over time that affect *all* ages in the bandwidth. Think of the government making changes to the set of treatments covered by the mandatory insurance. This assumption is less plausible for wide age bandwidths. For example, if new treatments are added to the basic benefit package which are only relevant to persons over 50, the time fixed effects will no longer adequately control for this as these are identified on people younger than 18. With apologies for the stereotype, adding a walker or Viagra to the basic package may not be picked up correctly by the time fixed effects. We can estimate Eq. (8) – in the form of Eq. (10) – because for the group who chose a voluntary deductible at least once, we have variation in deductible within a year.

The estimated λ coefficient is no longer (around) zero for persons who chose a voluntary deductible at least once: λ is −0.23, which implies that a marginal increase in the total deductible size leads to a reduction of healthcare expenditure.[51] Although our design becomes less powerful when higher age ranges are included in the model, we use the estimated coefficient λ of −0.23 as an upper bound of the reduction of moral hazard due to the voluntary deductible as it is likely an overestimation of the moral hazard effect for 13 to 65 year olds with a voluntary deductible. To illustrate, we get an overestimation of this moral hazard effect if people reduce their deductible for the next period because they expect higher expenditure in that period. Appendix 9.V discusses further why an overestimation is likely. Even with this upper bound for MH_l, we find that there is a cross subsidy from high-risk to low-risk agents. To compute the mean moral hazard effect with regard to the voluntary deductible we multiply λ with the average size of the voluntary deductible.

[50] This age bandwidth was chosen because for persons below age 13 we cannot determine whether they have chosen a voluntary deductible or not and because there are almost no persons above 65 who chose a voluntary deductible (see Appendix 9.U).

[51] The full results of estimating model (10) are presented in Table 12 in Appendix 9.V.

Table 5
Impact of the voluntary deductible on cross subsidies.

	18 year olds	Full population	
		Upper bound	Lower bound
Mean premium discount (Δp, €)	171	147	147
Mean OOP payment w.r.t. voluntary deductible (oop_i, €)	76	68	68
Mean moral hazard effect w.r.t. voluntary deductible (MH_i, €)	0	0	−41
Mean cross subsidy (($\Delta p - oop_i + MH_i$), €)	95	80	38
Total moral hazard effect w.r.t. voluntary deductible (× €1,000)	0	0	−37,148
Total cross subsidies (× €1,000)	1,124	72,139	34,991
Mean premium increase for persons without a voluntary deductible (€)	6	6	3

Notes: Due to rounding, the numbers may not add up.

Cross subsidies between persons with and without a voluntary deductible

Table 5 shows the impact of the voluntary deductible on cross subsidies for 18 year olds and the full population, the latter with an upper and lower bound of the impact on cross subsidies. For the upper bound, we use $\lambda = 0$, estimated for 18 year olds, and for the lower bound impact we use $\lambda = -0.23$, as described above.[52] Note that 18 year olds have a higher premium discount and larger OOP payments than the full population because they chose higher voluntary deductible levels.

For 18 year olds, we find substantial cross subsidies for the voluntary deductible: 95 euros per person with a voluntary deductible. The average out-of-pocket payment due to the voluntary deductible and the average moral hazard effect do not add up to the average premium discount. In other words, in this age bracket there is a cross subsidy from high-risk people (who never choose a voluntary deductible) to low-risk people (who tend to choose a voluntary deductible at least once). Abolishing the voluntary deductible and pooling everyone into one contract (with one mandatory deductible) would lower the premium for high-risk individuals.

For the lower bound estimate of the full population, we find slightly smaller effects because the premium discount is lower. The cross subsidies due to the voluntary deductible are between 38 and 80 euros per person. Abolishing the voluntary deductible will most likely increase healthcare spending as the moral hazard effects listed in Table 5 (in total roughly 37 million euros) would disappear. But recall from the analysis above that we view this as an upper bound on the moral hazard effect (and a lower bound on cross subsidies). Even with this upper bound on moral hazard, we find that the reduction in the premium is 'excessive' compared to reduced expenditure and increased out-of-pocket payments: voluntary deductibles increase the cross subsidy from high-risk to low-risks.

If we assume 'all else equal' after abolishing the voluntary deductible then the premium for people who did not choose a voluntary deductible would decrease with 3 to 6 euros whereas for persons with previously a voluntary deductible the premium would increase with 38 to 80 euros.[53] We observe only modest benefits for the high-risk types, which is due to the fact that in our data a large majority of the population does not opt for a voluntary deductible. But more recently the fraction of people with a voluntary deductible has almost doubled from 7% in 2012 to 13%.[54] Hence, the benefits for high-risk types of abolishing the voluntary deductible are higher now and will continue to grow with the share of people with a voluntary deductible.

Summarizing, we find that abolishing the voluntary deductible in the Netherlands will lead at most to a modest increase in healthcare spending. Although there is a substantial difference between the mean expenditure of people with and without a voluntary deductible, most of this difference is due to selection and only a small part is due to moral hazard.

The large selection effect could suggest that the market will unravel. This has not happened so far however. Potential explanations are that the Dutch are risk averse and do not easily opt for less generous coverage (Handel et al., 2022; Van Winssen et al., 2015) or that health insurers do not fully translate the selection effects into premium discounts (Douven et al., 2016). This is related to the observation that the share of individuals choosing a voluntary deductible is relatively small. If there are plenty of individuals with good health who do not choose a voluntary deductible, this mitigates the large selection effects and prevents a downward spiral (Marzilli Ericson and Sydnor, 2017).

For measuring total welfare effects also the benefits of offering choice must be taken into account. We follow here (Handel et al., 2022) who study heterogeneity in choice quality of the Dutch voluntary deductible (but they do not measure adverse selection and related cross subsidies). They estimate that abolishing the voluntary deductible results in a rather small average welfare loss of offering choice between 3 to 8 euros per person, depending on the choice of risk and inequality aversion of the population.[55]

Hence, abolishing the voluntary deductible has four favorable effects: there is a relatively small increase in healthcare costs, the cross subsidies from high-risk to low-risk types due to the voluntary deductible are removed, abolishing the voluntary deductible also

[52] That is, the upper bound for the moral hazard effect corresponds to the lower bound on the cross subsidy; and the other way around.

[53] Note that it is not obvious that 'all else will be equal': for example, if people with and without a voluntary deductible are pooled together their pooled elasticity and hence the mark-up can be different.

[54] See https://www.vektis.nl/uploads/Publicaties/Zorgthermometer/Zorgthermometer%20Verzekerden%20in%20Beeld%202021.pdf.

[55] Handel et al. (2022) argue that more effective choices by consumers would increase the uptake of the voluntary deductible and thus the benefits of offering choice. However, an increase in uptake might also increase the cross subsidies from the high-risk to low-risk types.

removes the risk of market unraveling and, as shown in Section 2, it improves welfare because the market sets insurance generosity for low-risks too low ($\hat{q}_l < q_l^*$). These benefits need to be weighed against the small benefits of offering choice for people who differ in their degree of risk aversion.

8. Concluding remarks

We show that – unless there are no selection issues in the health insurance market – it is optimal for the government to impose a minimum generosity level (maximum deductible) on health insurance contracts offered in the market. If, in addition, the moral hazard effect for high-risk types is at least as strong as for low-risk types, it is optimal to equate the minimum and maximum generosity. We use Dutch data to illustrate how this condition can be tested empirically. We conclude that for the Netherlands it is optimal to abolish the voluntary deductible.

We find significantly lower moral hazard effects for people who choose a voluntary deductible compared to people who choose the mandatory deductible only. Our results are cleanest for 18 year olds as we can use a quasi-experimental approach in the spirit of regression discontinuity design. Persons who do not choose a voluntary deductible reduce their healthcare spending at age 18 by 26 euros on average, when faced with a 100 euro increase in the deductible. However, persons at the same age who have chosen a voluntary deductible do not respond to an increase of the deductible. Also after extending the age range to 13 to 65 year olds we find modest moral hazard effects. We use a panel regression and find that a 100 euro increase in the deductible leads to an average reduction in healthcare spending of 23 euros per person for persons who chose a voluntary deductible, which we interpret as an upper bound on the moral hazard effect.

For the population as a whole these results translate into modest moral hazard effects: abolishing the voluntary deductible would increase healthcare expenditure for the Netherlands but only by a small amount. This finding that people choosing a higher deductible have lower moral hazard effects is not confined to the Netherlands. It is in line with the findings of Einav et al. (2013). We also document cross subsidies: the voluntary deductible increases the premium for people who do not choose a voluntary deductible, as the sum of the out-of-pocket payment and reduction of moral hazard due to the voluntary deductible is lower than the premium discount that they receive.

Politicians and policy makers should weigh the pros and cons of the voluntary deductible and decide whether the extra utility gain from offering choice outweigh the small reduction of moral hazard, the unintended cross subsidies (a reduction of risk solidarity) and the risk of market unraveling.

Our result that a minimum generosity level in health insurance is welfare enhancing is policy relevant for any country with a market based community-rated health insurance system. We analyzed the optimal level of the minimal generosity in the Dutch context. The advantage in the Netherlands is that the voluntary deductible is orthogonal to the coverage of basic insurance. For a (mandatory) health insurance contract there is the choice to increase the deductible by 100, 200, 300, 400 or 500 euros per year. This increase in the deductible is independent of the coverage (in terms of treatments and provider network) of the insurance contract. This facilitates our analysis. But a voluntary deductible is not a necessary feature for countries to benefit from the observation that a minimum generosity is welfare enhancing in a health insurance market suffering from selection problems.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jhealeco.2023.102782.

References

Alessie, R.J.M., Angelini, V., Mierau, J.O., Viluma, L., 2020. Moral hazard and selection for voluntary deductibles. Health Econ. 29 (10), 1251–1269.
Arrow, K.J., 1963. Uncertainty and the welfare economics of medical care. Amer. Econ. Rev. 53, 941–973.
Baicker, K., Goldman, D., 2011. Patient cost-sharing and healthcare spending growth. J. Econ. Perspect. 25 (2), 47–68.
Bajari, P., Dalton, C., Hong, H., Khwaja, A., 2014. Moral hazard, adverse selection, and health expenditures: A semiparametric analysis. Rand J. Econ. 45 (4), 747–763.
Chiappori, P.A., Durand, F., Geoffard, P.Y., 1998. Moral hazard and the demand for physician services: First lessons from a French natural experiment. Eur. Econ. Rev. 42, 499–511.
Cutler, D.M., Reber, S.J., 1998. Paying for health insurance: The trade-off between competition and adverse selection. Q. J. Econ. 113 (2), 433–466.
Cutler, D.M., Zeckhauser, R.J., 2000. Chapter The anatomy of health insurance. In: Culyer, A.J., Newhouse, J.P (Eds.), In: Handbook of Health Economics, vol. 1A, Elsevier, Amsterdam.
Douven, R., Remmerswaal, M.C., Kranendonk, L., 2016. Keuzegedrag Verzekerden En Risicosolidariteit Bij Vrijwillig Eigen Risico. Tech. Rep., Centraal Planbureau.
Dutch Healthcare Authority, 2014. Marktscan En Beleidsbrief Zorgverzekeringsmarkt: Weergave Van de Markt 2010–2014. Tech. Rep., Dutch Healthcare Authority.
Eichner, M.J., 1998. The demand for medical care: What people pay does matter. Am. Econ. Rev. Pap. Proc. 88 (2), 117–121.
Einav, L., Finkelstein, A., 2018. Moral hazard in health insurance: What we know and how we know it. J. Eur. Econom. Assoc. 16 (4), 959–982.
Einav, L., Finkelstein, A., Ryan, S.P., Schrimpf, P., Cullen, M.R., 2013. Selection on moral hazard in health insurance. Amer. Econ. Rev. 103 (1), 178–219.
Finkelstein, A., Taubman, S., Wright, B., Bernstein, M., Gruber, J., Newhouse, J.P., Allen, H., Baicker, K., 2012. The Oregon health insurance experiment: Evidence from the first year. Quat. J. Econ. 127 (3), 1057–1106.

Gardiol, L., Geoffard, P.Y., Grandchamp, C., 2008. Separating Selection and Incentive Effects in Health Insurance. Tech. Rep., CEPR Discussion Paper No. 5380.
Geruso, M., Layton, T.J., 2017. Selection in health insurance and its policy remedies. J. Econ. Perspect. 31 (4), 23–50.
Handel, B., Kolstad, J., Minten, T., Spinnewijn, J., 2022. The Social Determinants of Choice Quality: Evidence from Health Insurance in the Netherlands. Tech. Rep., NBER Working Paper No. 27785.
Lee, D.S., Lemieux, T., 2010. Regression discontinuity designs in economics. J. Econ. Lit. 48, 281–355.
Marzilli Ericson, E., Sydnor, J., 2017. The questionable value of having a choice of levels of health insurance coverage. J. Econ. Perspect. 31 (64), 51–75.
McGuire, T.G., 2012. Chapter Demand for health insurance. In: Pauly, M.V., McGuire, T.G., Barros, P.P. (Eds.), In: Handbook of Health Economics, vol. 2, Elsevier, Amsterdam.
McGuire, T.G., Van Kleef, R.C., 2018. Risk Adjustment, Risk Sharing and Premium Regulation in Health Insurance Markets: Theory and Practice. Elsevier, Amsterdam.
Newhouse and the Insurance Experiment Group, 1993. Free for All? Lessons from the RAND Health Insurance Experiment. Harvard University Press, Cambridge, Massachusetts.
Pauly, M.V., 1968. The economics of moral hazard: Comment. Amer. Econ. Rev. 58 (3), 531–537.
Remmerswaal, M.C., Boone, J., Bijlsma, M., Douven, R., 2019. Cost-sharing design matters: A comparison of the rebate and deductible in healthcare. J. Public Econ. 170, 83–97.
Rothschild, M., Stiglitz, J., 1976. Equilibrium in competitive insurance markets: An essay on the economics of imperfect information. Q. J. Econ. 90 (4), 629–649.
Thistlethwaite, D.L., Campbell, D.T., 1960. Regression-discontinuity analysis: An alternative to the ex post facto experiment. J. Educ. Psychol. 6 (51), 309–317.
Trottmann, M., Zweifel, P., Beck, K., 2012. Supply-side and demand-side cost sharing in deregulated social health insurance: Which is more effective? J. Health Econ. 31, 231–242.
Van Vliet, R., 2004. Deductibles and health care expenditures: Empirical estimates of price sensitivity based on administrative data. Int. J. Healthc. Finance Econ. 4, 283–305.
Van Winssen, K.P.M., van Kleef, R.C., van de Ven, W.P.M.M., 2015. How profitable is a voluntary deductible in health insurance for the consumer? Health Policy 119 (5), 688–695.
Van de Ven, W.P.M.M., Schut, F.T., 2008. Universal mandatory health insurance in the Netherlands: A model for the United States? Health Aff. 27 (3), 771–781.
Withagen-Koster, A., Van Kleef, R.C., F., E., 2018. Examining unpriced risk heterogeneity in the Dutch health insurance market. Eur. J. Health Econ. 19, 1351–1363.
Zeckhauser, R.J., 1970. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. J. Econ. Lit. 2 (1), 10–26.
Zweifel, P., 1987. Bonus systems in health insurance: A micro-economic analysis. Health Policy 7, 273–288.

Further reading

Baicker, K., Mullainathan, S., Schwartzstein, J., 2015. Behavioral hazard in health insurance. Q. J. Econ. 130 (4), 1623–1667.
Cockx, B., Brasseur, C., 2003. The demand for physician services: Evidence from a natural experiment. J. Health Econ. 22 (6), 881–913.
Croes, R., Katona, K., Mikkers, M., Shestalova, V., 2018. Evidence of Selection in a Mandatory Health Insurance Market with Risk Adjustment. Tech. Rep., TILEC Discussion Paper 13.
Dutch Healthcare Authority, 2011. Marktscan En Beleidsbrief Zorgverzekeringsmarkt: Weergave Van de Markt 2007–2011. Tech. Rep., Dutch Healthcare Authority.
Dutch Healthcare Authority, 2013. Marktscan En Beleidsbrief Zorgverzekeringsmarkt: Weergave Van de Markt 2008–2013. Tech. Rep., Dutch Healthcare Authority.
Ellis, R.P., 1986. Rational behavior in the presence of coverage ceilings and deductibles. Rand J. Econ. 17 (2), 158–175.
Johansson, N., Jakobsson, N., Svensson, M., 2019. Effects of primary care cost-sharing among young adults: Varying impact across income groups and gender. Eur. J. Health Econ. 20, 1271–1280.
Keeler, E.B., Rolph, J.E., 1977. The demand for episodes of treatment in the health insurance experiment. J. Health Econ. 7 (4), 337–367.
Klein, T.J., Salm, M., Hayen, A., 2018. Does the Framing of Patient Cost-Sharing Incentives Matter? the Effects of Deductibles Vs. No-Claim Refunds. Tech. Rep., CEPR Discussion Paper No. 12908.
Klein, T.J., Salm, M., Upadhyay, S., 2020. The Response to Dynamic Incentives in Insurance Contracts with a Deductible: Evidence from a Differences-in-Regression-Discontinuities Design. Tech. Rep., IZA Discussion Paper Series No 13108.