



Basic versus supplementary health insurance: Access to care and the role of cost effectiveness[☆]



Jan Boone^{a,b}

^a Center, TILEC, Department of Economics, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands

^b CEPR, London, UK

ARTICLE INFO

Article history:

Received 16 October 2017

Received in revised form 5 April 2018

Accepted 22 May 2018

Available online 31 May 2018

JEL classification:

I13

I14

D82

H51

Keywords:

Universal basic health insurance

Voluntary supplementary insurance

Public vs. private insurance

Access to care

Cost effectiveness

Moral hazard

Adverse selection

ABSTRACT

In a model where patients face budget constraints that make some treatments unaffordable without health insurance, we ask which treatments should be covered by universal basic insurance and which by private voluntary insurance. We argue that next to cost effectiveness, prevalence is important if the government wants to maximize the welfare gain that it gets from its health budget. Conditions are derived under which basic insurance should cover treatments that are mainly used by high risk agents with low income.

© 2018 The Author. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

This paper considers a health insurance system where the government offers universal basic health insurance and private parties offer voluntary supplementary insurance. The questions we analyze are: *which* treatments should be covered by health insurance and *how* should they be covered; i.e. by basic or supplementary insurance? Should basic insurance cover treatments with the best cost effectiveness scores, the most expensive ones or treatments used by “needy agents” (low income, high risk)? The motivation for this question is that many countries have such a mixed public/private health care system (Section 2), while the (health) economics literature does not analyze this question.

We introduce a framework where people buy insurance to get access to health care. In particular, agents face a budget constraint for health care expenditures. The agent buys insurance in order to be able to afford treatment if she needs it. A companion paper, Boone (2015), analyzes basic and supplementary insurance in a model with risk averse agents and the standard imperfections adverse selection and moral hazard but without budget constraints. The three main findings in such a model are that (i) basic insurance should cover treatments with the biggest adverse selection problems – this is a market imperfection the government can address, (ii) moral hazard plays no role in deciding whether a treatment should be covered in either basic or supplementary insurance; moral hazard (only) affects the extent to which a treatment is covered – this problem is the same for government and market, and (iii) cost effectiveness (CE) of a treatment does not affect its priority for basic insurance coverage. This last result is surprising given that many governments do take CE into account and are encouraged to do so by academics (Drummond et al., 2005; Gold et al., 1996).

[☆] Comments by Cedric Argenton, Annemieke Meijdam, two anonymous referees and seminar participants in Berlin and Tilburg are much appreciated. Financial support from the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged.

E-mail address: j.boone@uvt.nl

Compared to this standard model, the access to care model yields five new insights. First, CE scores do play a role in prioritizing which treatments should be insured. Because of limited budgets, money should be spent on treatments with a high health return per euro spent. Second, treatments with serious moral hazard issues should not be covered at all by (any) insurance. Third, relatively cheap treatments should not be covered by either basic or supplementary insurance. Fourth, basic insurance should cover treatments that are predominantly used by people who at the margin buy the most valuable treatments. The importance of disease (and treatment) prevalence to determine which treatments should be covered by basic insurance appears to be new. We derive sufficient conditions under which basic insurance should target treatments used especially by vulnerable people with high risk and low income. Finally, the welfare effects of government intervention in health insurance are magnitudes higher in an access to care model. Intuitively, in a standard model, the (adverse selection) inefficiency causes under-insurance for low risk types. Some people have argued that the cost of this is relatively low (Pauly, 2008). In an access to care model, low income people cannot afford high value treatments causing a substantial loss in welfare if there is no government intervention.

Simply covering the most cost effective treatments by basic insurance is likely to waste government resources. To illustrate, consider a highly cost effective treatment k that is only used by high income people who can cover all cost effective treatments themselves. With apologies for the stereotype, say k is a treatment for a golf or polo injury. Then covering k by basic insurance does not yield any health return at all despite k 's high cost effectiveness score. Indeed, the high income types using this treatment will not get access to more treatments if k is subsidized by the government: they have access to all cost effective treatments already.

To facilitate the exposition of the differences between risk aversion and access to care as motivation of health insurance, we assume that agents are risk neutral. Although we work with linear utility and a planner maximizing total welfare, budget constraints introduce endogenously a reason for redistribution.

The access to care model sheds new light on the definition of moral hazard due to health insurance. The health economics literature tends to define moral hazard as follows. Consider a treatment k that an agent uses if it is covered by health insurance but does not use if she would pay the full costs of k herself (Cutler and Zeckhauser, 2000, pp. 576). Then treatment k is often interpreted as over-consumption induced by moral hazard of insurance. The implication is that treatment k should not be covered by insurance. In an access to care model, an agent may not be able to afford treatment k without insurance even though the use of treatment k is socially efficient. Nyman (2003) is one of the first to discuss “efficient moral hazard”. To separate the concepts “efficiency” and “moral hazard”, we define social efficiency of a treatment independently from the question of whether or not it is covered by insurance.

Summarizing, we introduce a model with three imperfections in the insurance market: moral hazard, adverse selection and agents' liquidity constraints. When the government introduces mandatory basic insurance, it needs to address each of these imperfections. It should not cover/subsidize treatments to such an extent that their use becomes inefficient. As basic insurance is mandatory, it can reduce adverse selection inefficiencies in the private market by covering treatments especially used by high risk agents. Finally, the government should cover treatments that free up resources for agents with a high shadow value of health. We derive an index to prioritize treatments taking these three imperfections into account: priority is given by the weighted sum of adverse selection inefficiencies and health shadow values, where the weights are given by cost shares.

The paper is related to the following strands of literature. First, there is the literature on (health) insurance with adverse selection and moral hazard. Seminal papers are Rothschild and Stiglitz (1976), Pauly (1968) and Zeckhauser (1970). This literature is discussed in Boone (2015). No paper in this literature analyzes the optimal combination of basic and supplementary insurance.

Second, Nyman (1999) and Nyman (2003) stress the access to care motive of health insurance. As these papers point out, there is a difference between health insurance and insurance for property, such as a car. The fact that someone bought the car (usually) suggests that she can afford the car and also the repairs that come with it. Then insurance is indeed about reducing consumption risk due to car related expenses. This is different with health insurance. A 50,000 euro cancer treatment may be worth it in terms of utility gained, but a person may not be able to afford it. Nyman does not analyze how the access to care motive affects which treatments should be covered by basic and which by supplementary insurance.

Third, there is a literature on cost effectiveness of treatments. See, for instance, Drummond et al. (2005), Gold et al. (1996), Garber and Sculpher (2011), Meltzer and Smith (2011). The objective of CE analysis is to find the set of treatments that leads to the highest health gain for a given health care budget (e.g. Drummond et al., 2005; Gold et al., 1996). It calculates for each treatment the gain in “health” (usually measured in terms of quality adjusted life years – qaly's) per unit of the budget spent. Let v_k denote the health gain (in utility terms) from using treatment k and δ_k the cost of treatment k . We say that treatment k has a high CE score if v_k/δ_k is high¹: the health return per unit of the budget spent is high.

Roughly speaking, CE analysis proceeds as follows. Rank all treatments in terms of their CE score. Then treatments with the best scores are covered by insurance until the budget is spent. With one exception, the CE literature does not answer the question how treatments should be covered.² Although CE analysis can be applied both by governments and private insurers, it seems fair to say that this literature focuses on public insurance.³ A natural interpretation of this literature is then to say that government sponsored basic insurance should cover the treatments with the highest CE scores. This is intuitive in a model where a treatment is either available through public insurance or not at all.⁴ However, we consider a situation where treatments not included in basic can be covered in supplementary insurance.

As mentioned, the underlying idea of CE analysis is to maximize health given the government's budget constraint. As a consequence “economic evaluations do not usually incorporate the importance of the distribution, of costs and consequences, among different patient or population groups, into the analysis ... Rather, an equitable distribution of costs and consequences across socioeconomic or other

¹ The CE ratio is usually defined as δ_k/v_k : cost per unit of health gain. We work with the CE score such that a higher score and higher cost effectiveness are “better”.

² The only paper in the CE literature – that we are aware of – dealing explicitly with both private and public insurance is (Smith, 2007, pp. 147). He argues that in a setting with both public and voluntary private insurance, “interventions for the statutory package can be selected solely according to their cost-effectiveness”. In his model, there are no market imperfections in the private market. We augment this literature by allowing for market imperfections in the supplementary segment.

³ To illustrate, to seriously advice a (profit maximizing) private insurer which treatments to cover, one would have to take selection effects into account. In books like Drummond et al. (2005) and Gold et al. (1996), (adverse) selection is not discussed.

⁴ See Hoel (2007) for an extension where a treatment that is not insured can be paid for out-of-pocket.

defined groups in society is viewed as a competing dimension upon which decisions are made, in addition to that of efficient deployment of resources” (Drummond et al., 2005, pp. 47).⁵ In an access to care model, maximizing total health, endogenously introduces distributive concerns, as we will see below.

To the best of our knowledge, this is the first paper to combine and analyze the effects of moral hazard, adverse selection and cost effectiveness on optimal health insurance.

Finally, technically speaking, we analyze a two tiered problem similar to DeMarzo et al. (2005) and Faure-Grimaud et al. (1999). In our case, the government sets parameters for basic insurance which then affect the equilibrium interaction between insurers and consumers on the private market. The health insurance context is similar to Bijlsma et al. (2014), but they consider the effects of risk adjustment; not basic versus supplementary insurance.

This paper is organized as follows. The next section discusses health insurance coverage in mixed public-private systems as it is organized around the world. Then we introduce an access to care model. Section 4 introduces the supplementary insurance market. Section 5 characterizes optimal government policy for a government that wants to maximize total health. Here we derive that prevalence is an important factor for basic insurance. Given the government's optimal policy, Section 6 derives which treatments are covered by supplementary insurance. Section 7 derives sufficient conditions under which government uses basic insurance to target “needy” agents. But we also give an example where the government should target resources at low risk types with high income. Section 9 discusses the policy implications of our analysis. The proofs of results are in the appendix.

2. Health insurance coverage in practice

This section gives a brief overview of how public and private insurance are combined in practice. In most OECD countries there exists a combination of public and private health insurance (Colombo and Tapay, 2004, pp. 14). The idea often is that the public system addresses imperfections in the private health insurance market (Blomqvist, 2011; Zweifel, 2011). Roughly speaking, there are three ways in which public and private systems can be combined. First, private insurance may substitute for public insurance. That is, people are either covered privately or publicly. This is the case in Australia, Ireland, Spain and used to be the case in the Netherlands before 2006 (Colombo and Tapay, 2004, pp. 14). Second, private insurance is bought in addition to public insurance to get faster access (i.e. shorter waiting lists), have a broader choice of providers and treatments for a given condition. Examples include Austria, Denmark and Finland (Mossialos and Thomson, 2004, pp. 38/9). Third, private insurance is bought to cover treatment for conditions that are not covered by public insurance. One can think of dental care, physiotherapy and prescription glasses that are not covered by the public insurance system. But also, the public system may not fully cover the costs of a treatment and people can insure the public co-payment on the private market. Examples here include the Netherlands (after 2006), France and Luxembourg (Mossialos and Thomson, 2004, pp. 39–41).

The Patient Protection and Affordable Care Act (“Obamacare”) also combines mandatory public insurance with voluntary private insurance. Ten essential health benefits categories are defined that need (at least) to be covered by health insurance. It is left to the states to define which treatments within these categories will be covered exactly (McDonough, 2011). People are free to buy more coverage if they want, but any health insurance contract needs to cover at least the basics. So how does one decide which treatments belong to the essential health benefits?

Real world systems can have elements from the second and third categories above; e.g. private insurance providing coverage for treatments that are not covered by basic insurance and offering more choice in terms of health care providers. We focus in this paper on the third category (only).⁶

In particular, we consider the case with universal public coverage for a basic insurance package – like the essential health benefits in Obamacare. There is a fixed budget to finance the public system⁷ and not all treatments can be covered. We assume that there is only one treatment for each condition and we do not model providers. The question is: which conditions should be covered by public insurance and which treatments can be left to the private market? We denote insurance offered by the private voluntary market: supplementary insurance.⁸

3. Access to care

This section introduces a simple model that captures that people can experience financial problems to access health care. In particular, we model that there exist treatments that the agent cannot afford if her insurance does not cover these treatments. Further, people may not be able to afford health insurance that covers all available treatments. These access issues have been stressed both in the popular press (Cohn, 2007) and in academic journals (Nyman, 1999; Schoen et al., 2008, 2010). These papers document that without insurance (mainly low income) people forgo treatments that they actually need, because of cost considerations.

Adverse selection models in the tradition of Rothschild and Stiglitz (1976) (RS from here onwards) ignore budget constraints on the side of insured. Nyman (1999) and Nyman (2003) are early papers pointing out that the access motive of insurance is important. We introduce a simpler model than used by Nyman, as we want to characterize which treatments are covered by basic and which by supplementary insurance.

⁵ There are papers trying to address distributive concerns by introducing equity weights into the planner's objective function (see Meltzer and Smith, 2011; Garber and Sculpher, 2011, for a discussion). For instance, Smith (2005) derives how the optimal co-payment in a purely public system falls with a treatment's CE score. If the planner's objective function features equity weights – e.g. in favour of low income people – then treatments used by people with high weights feature lower co-payments as well. As this moves away from the simple idea of maximizing total health, it is harder to implement in practice.

⁶ We ignore private insurance as offering more choice than public insurance, where choice refers to horizontal differentiation. For the same condition, some people prefer treatment *a*, others *b* (some prefer to be treated by provider 1 others by provider 2). Papers modelling this type of choice (and ways in which the insurer can steer patients) include Chernew et al. (2000) and Chernew et al. (2007).

⁷ We assume that basic insurance is financed by general taxation. Hence, there is no reason to make it mandatory. If basic insurance is financed by social security contributions, mandatory insurance does become an issue (see Boadway et al., 2006; Stabile and Thomson, 2014, for a discussion).

⁸ Sometimes this is called complementary insurance (see Mossialos and Thomson, 2004, pp. 16).

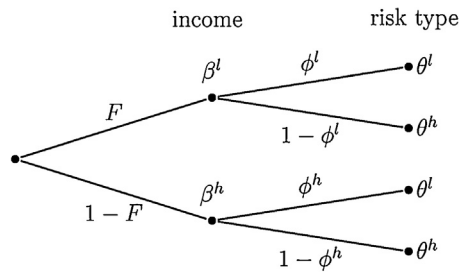


Fig. 1. Fractions of the different income and risk types with $\beta^h > \beta^l$.

3.1. utility

We model access to care in the following way. An agent has a budget β that she can (at max.) spend on health care. That is, this budget is used to spend on insurance premium, co-payments and buying uninsured treatments out-of-pocket. This is modelled in a reduced form way in the sense that we do not derive the budget from a utility function where the consumer decides on allocating income to housing, transport, food, health care etc. We take the budget β for health care expenditure as exogenous.⁹ Hence, we assume that people face a credit market constraint. They cannot borrow money to finance a treatment (investment in health). To illustrate, a bank may be reluctant to lend money for a liver transplant as the operation itself can signal reduced earning capacity.

We assume that agents are risk neutral for two reasons. First, this allows us to focus on access to care, instead of risk aversion, as a motive for insurance. Boone (2015) analyzes the case of risk averse insured. Second, with utility linear in consumption, a planner who maximizes total welfare, focuses on efficiency (maximizing health) and not on redistribution. As discussed in the introduction, this is in line with the CE literature. However, as we will see below, in an access to care model distributional considerations come to the fore endogenously; even if welfare is linear in consumption.

A fraction $F[1-F]$ of the population has a low [high] budget $\beta^l[\beta^h > \beta^l]$ available to spend on health care. We think of people with the high [low] budget $\beta^h[\beta^l]$ as having high [low] income. Budget β captures the *access to care* aspect of the model. Within income class $i \in \{l, h\}$, a fraction $\phi^i[1-\phi^i]$ is low [high] risk type. For each treatment k the high risk type has higher probability θ_k^h of needing the treatment than the low risk type $\theta_k^l \leq \theta_k^h$.¹⁰ Risk type θ captures *adverse selection*. Further, we know that health and income are positively correlated.¹¹ Hence a natural assumption is $\phi^l \leq \phi^h$ (although we will not explicitly use this). Fig. 1 gives an overview of the relevant parameters.¹²

In words, we have four types of agents: they can have either high or low income and they can be high or low risk from the point of view of the insurer. The income heterogeneity speaks to the access to care motive of insurance, while risk types capture the adverse selection inefficiency in insurance markets. We derive conditions under which the government is especially concerned with the low-income, high-risk agents. Intuitively, this agent faces expensive insurance in the supplementary market and may not be able to afford this due to her low income. In contrast, in a standard insurance model without access to care issues, the high risk type gets efficient insurance at an actuarially fair price. The low risk type is faced with an inefficiency.

3.2. Treatments

The set of conditions (“illnesses”) is denoted by $K = \{1, 2, \dots, \kappa\}$. For each condition $k \in K$ there is exactly one treatment; also denoted by k .¹³

We assume that conditions are contractible. That is, a physician can determine whether or not a patient suffers from condition k .¹⁴ An insurer or the planner can verify this information. Hence, there is no moral hazard on what treatment is needed. However, we do consider moral hazard on whether treatment is needed at all.

We model *moral hazard* as follows. There is a probability $\psi_k[1-\psi_k]$ (the same for the h and l type, to ease notation) that the patient is in state $s=0[1]$. State 0 means that the patient has the condition/disease but treatment k is not cost effective in addressing it. For instance, state 0 is a light form of the disease which does not need treatment in the sense that the treatment cost outweighs its benefit. However, if the treatment is fully insured, even an ε -gain will induce the patient to use it. Examples of treatments and the situations in which they are not needed are given here: <http://www.cityam.com/252096/some-40-nhs-treatments-you-could-get-prescribed-dont-help> (accessed on 11.11.16). In state 1 treatment k is cost effective: the utility gain from treatment exceeds the treatment cost.

Formally, the utility of treatment k in state s is denoted by $v_{sk} > 0$ and the cost of the treatment is denoted by δ_k . To ease notation, v_{sk} does not depend on the patient's type and δ_k does neither depend on the patient's type nor on the state s . We normalize the value of no treatment in state $s \in \{0, 1\}$ to 0. Hence, v_{sk} captures the added utility (compared to doing nothing) of treatment. The model is a bit heavy

⁹ One simple way to get such a budget β for health care is to assume a Cobb Douglas utility function over categories of spending like food, transport, housing and health. Then expenditure on each spending category is a fixed fraction of the agent's income.

¹⁰ This implies that types are consistently ranked. See Boone and Schottmüller (2017) for an analysis when this is not the case due to a violation of single crossing.

¹¹ See, for instance, Frijters et al. (2005), Finkelstein and McGarry (2006), Gravelle and Sutton (2009) and Munkin and Trivedi (2010). Possible explanations for this correlation between income and health include the following. High income people are better educated and hence know the importance of healthy food, exercise etc. Healthy food options tend to be more expensive and therefore better affordable to high income people. Or (with causality running in the other direction) healthy people are more productive and therefore earn higher incomes.

¹² The model is reminiscent of models where agents differ on two dimensions with two types on each dimension: Netzer and Scheuer (2010), Smart (2000) and Wambach (2000). However, in these models more than two IC constraints can be binding leading to the possibility that high and low risk types are pooled in equilibrium. Lemma 2 in the appendix shows that in our model at max. two IC constraints are binding. This keeps our model simple and allows us to analyze basic vs. supplementary insurance.

¹³ We focus on “ex post” treatments. For an analysis of ex ante measures (prevention), see Ellis and Manning (2007).

¹⁴ Alternatively, treatment k for a patient suffering from $k' \neq k$ yields zero utility.

Table 1Notation summary of expected value and costs for the government (basic insurance), insurers (suppl. insurance) and out-of-pocket for the patient of treatment k for agent ij .

Treatment k for agent ij	Basic insurance	Suppl. insurance	Out-of-pocket	Total
$c_k^{ij} \geq v_{0k}$				
Expected value				$\theta_k^j(1 - \psi_k)v_{1k}$
Expected costs	$\theta_k^j(1 - \psi_k)(\delta_k - \gamma_k)$	$\theta_k^j(1 - \psi_k)(\gamma_k - c_k^{ij})$	$\theta_k^j(1 - \psi_k)c_k^{ij}$	$\theta_k^j(1 - \psi_k)\delta_k$
$c_k^{ij} < v_{0k}$				
Expected value				$\theta_k^j v_k$
Expected costs	$\theta_k^j(\delta_k - \gamma_k)$	$\theta_k^j(\gamma_k - c_k^{ij})$	$\theta_k^j c_k^{ij}$	$\theta_k^j \delta_k$

on notation with costs and benefits of treatments for four different agents. Table 1 at the end of the section keeps track of the notation introduced here.

In our notation, treatment is not cost effective in state 0 is expressed as: $v_{0k} < \delta_k$. It is cost effective in state 1: $v_{1k} > \delta_k$.¹⁵ In addition to $v_{0k} < \delta_k$, we make the assumption that

$$v_k \equiv \psi_k v_{0k} + (1 - \psi_k) v_{1k} < \delta_k \quad (1)$$

In words, if the treatment is used in both states the use of treatment k destroys social surplus in the sense that the cost exceeds the average value of the health gain. A treatment with high (zero) v_{0k} features serious (no) moral hazard problems – given that (1) holds. If an insurer covers treatment k with a co-payment below v_{0k} , the patient uses the treatment even though this is socially inefficient. In the terminology of (Chandra and Skinner, 2012 July, pp. 663), we consider category II treatments: clear benefits to some patients (group $1 - \psi_k$) but poor average cost effectiveness across all patients.¹⁶

Whereas condition k is contractible, the severity (state) of the condition is not. We assume that the patient can reveal the state 0 or 1 by truthfully reporting her symptoms to her physician. However, she can also exaggerate her symptoms to qualify for treatment in state 0.¹⁷ Hence, a co-payment is used for the patient to truthfully reveal the state. As utility is linear in consumption, a co-payment

$$c_k \geq v_{0k} \quad (2)$$

induces truthful revelation of s by the patient. Hence, the two states 0,1 formalize the idea that a co-payment can reduce moral hazard on the patient side.

Summarizing the points above. We define moral hazard as the use of treatment k in situations where it is socially inefficient. That is, the use of k in state 0. Moral hazard can be avoided by setting a co-payment at least equal to v_{0k} . Although, we focus here on the case where both the government and private insurers induce efficient health care consumption, moral hazard still has an important role to play. It implies that treatments cannot be offered completely free of charge. In an access to care model, agents set aside money that they need to spend on co-payments (and uninsured treatments). This money set aside cannot be spent on the insurance premium to cover more treatments. Hence, we will see a trade off between the number of treatments covered and the set of treatments that can be consumed efficiently. The higher the co-payment, the more treatments become available that can be used efficiently but the lower the insurance premium has to be (and hence fewer treatments are covered).

Our definition of moral hazard as linked to social inefficiency follows from the literature on mechanism design. As mentioned in the introduction, the health economics literature sometimes defines moral hazard in terms of behaviour (Cutler and Zeckhauser, 2000, pp. 576), not efficiency. As discussed at the end of this section, these definitions can conflict in an access to care model. In particular, without insurance low income agents can decide to forgo a treatment k that is actually efficient for them to use: consuming k with insurance is then moral hazard under the traditional definition in the health economics literature, but not under the definition that we use here.

Without loss of generality, an agent has at max. one illness $k \in K$ in the period under consideration: we view an agent needing two treatments k and k' as suffering from one condition $\tilde{k} \in K$ where she needs one treatment $\tilde{k}$ which consists of the combination treatments k and k' . This leads to two advantages. First, we do not impose that the value of treatments can simply be added; we allow for $v_{\tilde{k}} \neq v_k + v_{k'}$. Indeed, co-morbidity can affect the value of a treatment. Second, in terms of contracts, we allow for general co-payments. In the state of the world where an agent needs treatments k and k' , she pays a co-payment equal to $c_{\tilde{k}}$ which does not necessarily equal $c_k + c_{k'}$.

To avoid confusion: when we say that a treatment is covered by insurance, this statement is conditional on the condition. In the notation above, we allow treatment k to be covered in condition k , while it is not necessarily covered in condition $\tilde{k}$. As an illustration, beta blockers can be covered for someone with (only) heart problems but not for a diabetes patient with heart problems (or the other way around).

3.3. Basic insurance

The (total/social) cost of treatment k is denoted δ_k . By covering treatment k by basic insurance, the government subsidizes this treatment for the private sector. After government subsidy of the treatment, the cost of k for the market is denoted by γ_k . We assume that $\gamma_k \leq \delta_k$: the government does not tax treatments. Each time the treatment is used, the government chips in $\delta_k - \gamma_k$. We say that k is not covered by basic insurance if $\gamma_k = \delta_k$. The reduced cost shows up in the insurance premium (if k is also covered by supplementary insurance) or in the price of uninsured treatment (if k is not covered by supplementary insurance).

Given γ_k set by the government, private insurers can offer coverage of k at co-payment c_k . If $c_k = \gamma_k$, treatment k is not covered by supplementary insurance; if $c_k < \gamma_k$ supplementary insurance covers the difference $\gamma_k - c_k$. Summarizing, when treatment k is used, the government pays $\delta_k - \gamma_k$, private health insurance pays $\gamma_k - c_k$ and the patient pays c_k out-of-pocket.

¹⁵ See Boone (2015) for the case where different treatments are optimal in different states.

¹⁶ The analysis below can be extended to deal with treatments $v_k > \delta_k$, but it makes the notation rather cumbersome.

¹⁷ See Schottmüller (2013 July) for a formal analysis of this situation as a cheap talk game.

As there are four different types ij (with income β^i and risk class θ^j , $i, j \in \{l, h\}$), insurers can offer four different contracts indexed by ij . Hence, we write c_k^{ij} for the co-payment on treatment k for type ij . As basic insurance is universal, the cost reduction γ_k applies to all types and γ is not indexed by ij .

Consider an agent ij who sets aside an amount of money C^{ij} (to be determined below) to pay either for co-payments or for uninsured treatments. Given C^{ij} , c_k^{ij} , we can distinguish three sets of treatments to which the agent has access:

$$\begin{aligned} K_e^{ij} &= \{k \in K | v_{0k} \leq c_k^{ij} \leq C^{ij} < \gamma_k\} \\ K_i^{ij} &= \{k \in K | v_{0k} > c_k^{ij} \leq C^{ij} < \gamma_k\} \\ K_n^{ij} &= \{k \in K | \gamma_k \leq C^{ij}\} \end{aligned} \quad (3)$$

The first set of treatments can only be used by the agent if she buys insurance for these treatments as $\gamma_k > C^{ij}$. That is, she cannot pay for it out-of-pocket. If she buys insurance for these treatments, they are used efficiently. Because $c_k^{ij} \geq v_{0k}$, the treatments are only used in state 1. The second set of treatments, if used at all (i.e. if covered by insurance), will be used inefficiently. The co-payment is below v_{0k} , hence the treatment is used in both states 0 and 1. Thus treatments in this set reduce total welfare when used (see Eq. (1)). The final set consists of treatments for which the agent does not need to buy insurance. These treatments can be paid out of C^{ij} . As the agent here is risk neutral (she buys insurance to get access to care) and – by construction – needs at max. one treatment, she will not buy insurance for treatments in the set K_n^{ij} .

To ease the exposition, we assume $\gamma_k \geq v_{0k}$: the government does not induce inefficient health care consumption by making it so cheap that it will be used in both states 0 and 1. To illustrate, allowing for $\gamma_k < v_{0k}$, would lead us to introduce two sets of uninsured treatments (K_{ne}^{ij}, K_{ni}^{ij}); one set where uninsured treatments are used efficiently ($v_{0k} \leq \gamma_k \leq C^{ij}$) and one where treatments are used inefficiently ($\gamma_k < v_{0k}, C^{ij}$). To avoid this cumbersome notation and given that most governments worry nowadays about efficient health care utilization, $\gamma_k \geq v_{0k}$ seems the relevant case to consider.

Treatments with $c_k^{ij} > C^{ij}$ are unavailable for agent ij as she cannot afford the co-payment at the moment that she needs the treatment. This is line with evidence cited by Nyman (2003), (Pauly, 2008, pp. 116) and (Schokkaert and van de Voorde, 2011, pp. 334) where high co-payments cause some people to forgo valuable treatments.

As we discussed above, this is the point where the definition of moral hazard used in the health economics literature differs from ours. A treatment k in the set K_e^{ij} is efficient to use, but without insurance coverage agent ij would not buy k because she cannot afford it. This is interpreted as moral hazard in the health economics literature because it is seen as over-consumption induced by insurance. In a model with budget constraints, this definition is misleading since any $k \in K_e^{ij}$ is actually socially efficient.

4. Insurance market

For given γ_k set by the government, this section describes the supplementary insurance market. In this market, we consider both second and third degree price discrimination. With third degree price discrimination, insurers can risk rate: charge different risk types different prices for the same contract. With second degree price discrimination insurers know that there are different types of insured, but they cannot risk rate. This can be the case either because they do not observe a customer's type at the moment that insurance is bought or because the government does not allow them to risk rate (community rating requirement).

Further, we assume that the supplementary insurance market is perfectly competitive. Hence, we focus on contracts that are priced in an actuarially fair way (as long as this satisfies incentive compatibility, which it does in our model).

4.1. Budget constraints

First, consider a market where insurers can risk rate: they offer a contract for each type ij without worrying about incentive compatibility constraints (as they observe the customer's type). Given γ_k set by the government, if agent ij can buy actuarially fair insurance, her optimal contract is determined by choosing ρ_k^{ij} , c_k^{ij} and C^{ij} which solve:

$$\begin{aligned} V^{ij} = \max_{C^{ij}, c_k^{ij}, \rho_k^{ij}} & C^{ij} + \sum_{k \in K_e^{ij}} \theta_k^j (1 - \psi_k) \rho_k^{ij} (v_{1k} - c_k^{ij}) + \sum_{k \in K_i^{ij}} \theta_k^j \rho_k^{ij} (v_k - c_k^{ij}) + \sum_{k \in K_n^{ij}} \theta_k^j (1 - \psi_k) (v_{1k} - \gamma_k) \\ & - \lambda^{ij} \left(\sum_{k \in K_e^{ij}} \theta_k^j (1 - \psi_k) \rho_k^{ij} (\gamma_k - c_k^{ij}) + \sum_{k \in K_i^{ij}} \theta_k^j \rho_k^{ij} (\gamma_k - c_k^{ij}) - (\beta^i - C^{ij}) \right) \end{aligned} \quad (4)$$

where v_{1k} , v_k denote the additional value/utility of treatment compared to receiving no treatment. Recall that we normalized the value of no treatment to 0. Further, $\rho_k^{ij} \in [0, 1]$ denotes the fraction of supplementary insurance bought by the agent and λ^{ij} denotes the Lagrange

multiplier on agent ij 's budget constraint. As we will see below, for all treatments except one, we have $\rho_k^{ih} \in \{0, 1\}$. The last treatment that the agent covers by insurance can feature $\rho_k \in (0, 1)$ ¹⁸ as the agent hits her budget constraint.¹⁹

The first term of V^{ij} is money set aside C^{ij} : money not spent on co-payments or uninsured treatments can be spent on other goods. We assume that utility is linear in consumption and normalize the marginal utility of income to 1. In Eq. (4), there is no ρ_k^{ij} term for treatments $k \in K_n^{ij}$. These treatments are not insured and bought when needed. Treatments $k \in K_n^{ij}$ are only needed in state 1 and hence the probability that θ^j agent needs such a treatment k equals $\theta_k^j(1 - \psi_k)$. If the treatment is used, it gives utility v_{1k} at cost γ_k for the patient. Finally, treatments in K_n^{ij} do not appear explicitly in the agent's budget constraint as they are paid out of money set aside C^{ij} in case such a treatment in K_n^{ij} is needed.

If $\rho_k^{ij} = 0$ for $k \in K_e^{ij}, K_i^{ij}$, agent ij has no access to the treatment because $\gamma_k > C^{ij}$: treatment k is too expensive to pay out-of-pocket. In this case, the term for this treatment "disappears" both in the value expression in the first line of equation (4) and in the budget constraint in the second line. For $\rho_k^{ij} > 0$, agent buys treatment $k \in K_e^{ij} [k \in K_i^{ij}]$ in state 1 [states 0 and 1]. Hence, the probability that the treatment is needed equals $\theta_k^j(1 - \psi_k) [\theta_k^j]$. Utility of the treatment when used is given by $v_{1k} [v_k]$. The cost of having this treatment covered equals γ_k which is partly paid as co-payment (c_k^{ij}) and partly ($\gamma_k - c_k^{ij}$) as insurance premium. The latter expression appears explicitly in the agent's budget constraint.²⁰

The choice of the amount C^{ij} set aside to finance co-payments affects the budget constraint and the sets K^{ij} in Eq. (3). By increasing C^{ij} less income is available to buy insurance coverage ρ_k , but a larger set of efficient treatments $K_e^{ij} \cup K_n^{ij}$ becomes available. This is the trade off that the agent faces when deciding on C^{ij} (see Section 8).

Once an agent has decided to put C^{ij} aside for co-payments and uninsured treatments, it is optimal for the agent to have co-payment $c_k^{ij} = C^{ij}$ for each treatment k with $\gamma_k \geq C^{ij}$ for which she buys coverage. Given that C^{ij} is available for co-payments, the insurance premium can be reduced by having this co-payment for each treatment. The reduced premium helps to cover more treatments. In a standard insurance model, it is optimal to let the co-payment vary per treatment depending on the extent of moral hazard v_{0k} (Boone, 2015). But in an access to care framework, it is better to determine the highest co-payment that one can afford and then set this co-payment for each treatment for which one buys insurance. As $c_k^{ij} = C^{ij}$ for each k , we now write $c^{ij} = C^{ij}$ for the co-payment.

In a standard insurance model with moral hazard and adverse selection, budget constraints do not play a role. People buy insurance for all treatments where value exceeds costs. Less than full insurance happens only due to moral hazard problems or due to adverse selection where low risk types get under-insured in order to separate types. Under-insurance because people cannot afford health insurance cannot be captured in the standard framework.

In an access to care model, budget constraints are important. Hence, we think of Lagrange multiplier λ^{ij} in Eq. (4) as being bigger than 1. Because of the C^{ij} term in both the objective function and budget constraint, we know that $\lambda^{ij} \geq 1$. Here we assume that

$$\lambda^{ij} > 1 \quad (5)$$

for each type ij . Even if both ij and the government spend their whole budget, there are treatments available that are valuable at the margin. We denote λ the health shadow price (of income). Lemma 1 below derives how λ varies with type ij .

Hence, (5) holds if the set of valuable treatments is big compared to the agents' and government's budget. If agent ij has $\lambda^{ij} = 1$, this agent has access to all valuable treatments. Increasing this agent's income further (or reducing co-payments) will not increase the agent's health. Our analysis applies as long as (5) holds for at least one type ij . It simplifies the exposition to assume that (5) holds for all ij . If no type features $\lambda^{ij} > 1$, the nature of the analysis changes. Budget constraints are then no longer an issue and we are back to the analysis in Boone (2015).

4.2. Incentive compatibility constraints

With second degree price discrimination (community rating), we need to add incentive compatibility (IC) constraints to the analysis in the previous section. Since insurers can no longer identify types, they need to offer ij an insurance contract that she prefers to contracts offered to the other types. If this condition is satisfied, we call ij 's contract incentive compatible.

As above, we consider a perfectly competitive supplementary insurance market. We follow the RS definition of perfect competition equilibrium under community rating: (i) each offered contract makes non-negative profits and (ii) given the equilibrium contracts there is no other contract yielding positive profits.²¹ Lemma 2 in the appendix shows that in this model with budget constraints, two features of the RS model are preserved: (i) θ^h and θ^l types are not pooled into one insurance contract²² and (ii) the relevant IC constraint is θ^h trying to mimic θ^l ; this happens in each income class. A type in one income class does not want (or is not able) to mimic a type in another income class.

¹⁸ In other words, insurance is random in this case for at most one treatment: if the agent needs treatment k , the insurer randomizes (in a verifiable way) to determine whether the treatment can be reimbursed. In the model, the agent is risk neutral. Hence she does not mind this randomization. We do not suggest lottery-insurance as a serious real world possibility. Indeed, we know that people do not like probabilistic insurance (Wakker et al., 1997). Here we use it to resolve an integer problem as it is not optimal for the insurer to raise the co-payment on this treatment above C^{ij} (which would make the treatment unavailable for the agent).

¹⁹ For the θ^l type there are at max. two such treatments with $\rho_k \in (0, 1)$ in case of second degree price discrimination: one when she hits her budget constraint and one for the incentive compatibility constraint.

²⁰ If treatment k is covered by basic insurance but not by supplementary insurance for ij : $\gamma_k = c_k^{ij}$ and $\rho_k^{ij} = 1$ does not increase the supplementary insurance premium in the budget constraint.

²¹ In RS an equilibrium may not exist if the fraction of θ^l -types is large. This paper does not focus on existence of equilibrium. Implicitly, we assume that the fraction of θ^l -types is small enough that an equilibrium exists.

²² Assuming they are different given the treatments they can afford with their budgets: see Lemma 2.

The IC constraint for θ^h with income β^i can be written as²³:

$$V^{ih} \geq \beta^i + \sum_{k \in K_e^{il}} (1 - \psi_k) \rho_k^{il} (\theta_k^h (v_{1k} - c^{il}) - \theta_k^l (\gamma_k - c^{il})) + \sum_{k \in K_i^{il}} \rho_k^{il} (\theta_k^h (\psi_k v_{1k} + (1 - \psi_k) v_{0k} - c^{il}) - \theta_k^l (\gamma_k - c^{il})) + \sum_{k \in K_n^{il}} (1 - \psi_k) \theta_k^h (v_{1k} - \gamma_k) \quad (IC)$$

The utility of type ih choosing her own contract equals V^{ih} . If she mimics the θ^l type with the same income, her utility equals the sum of the following terms. First, her budget β^i as money that is not spent on the insurance premium, co-payment or uninsured treatments can be used for consumption of other goods. When buying insurance for a treatment k in K_e^{il} , K_i^{il} the premium is determined by θ_k^l . But the probability that an insured treatment k is actually used (and yields payoff equal to utility minus co-payment) is given by θ_k^h . When using an uninsured treatment $k \in K_n^{il}$, only the probability θ_k^h (not θ_k^l) is relevant for the mimicking θ^h type.

From RS type of analysis, we know that in equilibrium, the utility of θ^h types is maximized, while the utility of θ^l is maximized subject to constraint (IC). In our model, this happens per income class β^i . That is, given V^{ih} , insurers choose ρ_k^{il} , $C^{il} = c^{il}$ to solve

$$V^{il} = \max_{c^{il}, \rho_k^{il}} c^{il} + \sum_{k \in K_e^{il}} \theta_k^l (1 - \psi_k) \rho_k^{il} (v_{1k} - c^{il}) + \sum_{k \in K_i^{il}} \theta_k^l \rho_k^{il} (v_k - c^{il}) + \sum_{k \in K_n^{il}} \theta_k^l (1 - \psi_k) (v_{1k} - \gamma_k) - \lambda^{il} \left(\sum_{k \in K_e^{il}} \theta_k^l (1 - \psi_k) \rho_k^{il} (\gamma_k - c^{il}) + \sum_{k \in K_i^{il}} \theta_k^l \rho_k^{il} (\gamma_k - c^{il}) - (\beta^i - c^{il}) \right) - \nu^{il} \left(\sum_{k \in K_e^{il}} (1 - \psi_k) \rho_k^{il} (\theta_k^h (v_{1k} - c^{il}) - \theta_k^l (\gamma_k - c^{il})) \right) + \sum_{k \in K_i^{il}} \rho_k^{il} \left(\theta_k^h (\psi_k v_{1k} + (1 - \psi_k) v_{0k} - c^{il}) - \theta_k^l (\gamma_k - c^{il}) \right) + \sum_{k \in K_n^{il}} (1 - \psi_k) \theta_k^h (v_{1k} - \gamma_k) - (V^{ih} - \beta^i) \quad (6)$$

where ν^{il} is the Lagrange multiplier on (IC).

Since no agent wants to mimic the high risk type, $\nu^{ih} = 0$ for each income class i . As low risk agents get a cheaper contract (lower premium) than a high risk type, the latter would like to buy (mimic) the low risk contract. The multiplier ν^{il} measures the extent to which the high risk agent prefers the low risk contract. The higher ν^{il} , the more attractive the l -contract is for the h -type. As shown in Section 6, the higher ν^{il} , the more insurers need to distort the l -contract to make it unattractive for the h -type.

4.3. Timing and information

The government, private insurers and agents know the primitives of the model. Before offering their supplementary insurance contracts, insurers know the treatments covered by basic insurance (γ_k). When evaluating supplementary insurance contracts, agents know what is covered by basic insurance. There is asymmetric information when an agent falls ill: her condition k is verifiable but only she knows the state/severity $s \in \{0, 1\}$ of her condition.

The timing of the game is as follows. In the first stage, the government decides which treatments to cover by basic insurance (i.e. it sets γ_k for each $k \in K$). Then – knowing γ_k – insurers choose contracts. We focus on second degree price discrimination (community rating): private insurers cannot discriminate types θ^l , β^i .²⁴ From the contracts that they can afford, consumers buy the insurance contract which maximizes their expected utility in stage 3. Finally, if an insured falls ill, she gets treatment if either (i) she has insurance coverage for this treatment (and pays the co-payment) or (ii) she can afford to buy it out-of-pocket.

At first sight, one would expect that backward induction is needed to find the government's optimal policy. However, because supplementary insurance is chosen to maximize agents' utility, it turns out that we can use an envelop argument (see Appendix B for details). This allows us to characterize the government's optimal basic insurance package before deriving the contracts offered on the supplementary market. The next section derives the optimal γ_k for $k \in K$. Section 6 derives which treatments are covered by supplementary insurance for each type ij and Section 8 derives the optimal co-payment for each contract.

5. Optimal government policy

The government chooses $\gamma_k \in [v_{0k}, \delta_k]$ to maximize total welfare, defined as

$$W = F(\phi^l V^{ll} + (1 - \phi^l) V^{lh}) + (1 - F)(\phi^h V^{hl} + (1 - \phi^h) V^{hh}) \quad (7)$$

where V^{ih} [V^{il}] is defined in Eq. (4) [(6)]. We follow the CE literature where the government maximizes total health without risk and/or income solidarity considerations.

²³ Substituting β^i from the budget constraint gives a right hand side similar to (4). Although equivalent, it turns out to be convenient to start the right hand side here with β^i instead of C^l .

²⁴ Risk rating is captured by $\nu^{il} = 0$ in Eq. (6): insurers can discriminate between types in terms of premium without worrying about (IC).

When choosing basic insurance, the government has three market imperfections that it needs to address: moral hazard (creating a lower bound for γ_k to prevent inefficient use), adverse selection (covering treatments used by high risk types to reduce distortions for low risk types) and access to care (making treatments available for types with high λ^{ij}). Proposition 1 derives the index balancing these priorities.

These imperfections have the following effects on the government's prioritization of treatments for coverage under basic insurance. Access to care and budget constraints cause $\lambda^{ij} > 1$. Higher λ^{ij} increases priority for treatments used (especially) by ij . Adverse selection leads to $v^{il} > 0$. As the government can reduce the adverse selection inefficiency – in contrast to the market – treatments associated with adverse selection get priority as well. Finally, moral hazard does not affect the priority of treatments directly. However, it does limit the extent to which the government can subsidize treatments covered by basic insurance.

The planner has a budget B – financed by a general tax – that can be spent on basic insurance. The budget constraint can be written as²⁵

$$\begin{aligned} & F\phi^l \left(\sum_{k \in K_e^{ll}} \theta_k^l (1 - \psi_k) \rho_k^{ll} (\delta_k - \gamma_k) + \sum_{k \in K_i^{ll}} \theta_k^l \rho_k^{ll} (\delta_k - \gamma_k) + \sum_{k \in K_n^{ll}} \theta_k^l (1 - \psi_k) (\delta_k - \gamma_k) \right) \\ & + F(1 - \phi^l) \left(\sum_{k \in K_e^{lh}} \theta_k^h (1 - \psi_k) \rho_k^{lh} (\delta_k - \gamma_k) + \sum_{k \in K_i^{lh}} \theta_k^h \rho_k^{lh} (\delta_k - \gamma_k) + \sum_{k \in K_n^{lh}} \theta_k^h (1 - \psi_k) (\delta_k - \gamma_k) \right) \\ & + (1 - F)\phi^h \left(\sum_{k \in K_e^{hl}} \theta_k^l (1 - \psi_k) \rho_k^{hl} (\delta_k - \gamma_k) + \sum_{k \in K_i^{hl}} \theta_k^l \rho_k^{hl} (\delta_k - \gamma_k) + \sum_{k \in K_n^{hl}} \theta_k^l (1 - \psi_k) (\delta_k - \gamma_k) \right) \\ & + (1 - F)(1 - \phi^h) \left(\sum_{k \in K_e^{hh}} \theta_k^h (1 - \psi_k) \rho_k^{hh} (\delta_k - \gamma_k) + \sum_{k \in K_i^{hh}} \theta_k^h \rho_k^{hh} (\delta_k - \gamma_k) + \sum_{k \in K_n^{hh}} \theta_k^h (1 - \psi_k) (\delta_k - \gamma_k) \right) \leq B \end{aligned} \quad (\text{GBC})$$

For each type ij , the equation gives the probability that this type uses treatment k and hence the government needs to contribute $\delta_k - \gamma_k$ for treatment of this type. If either treatment k is not covered by basic insurance ($\gamma_k = \delta_k$) or it is not used by anyone ($k \notin K_n^{ij}$ and $\rho_k^{ij} = 0$ for each ij), the government does not need to contribute anything for this treatment.

In order to characterize the optimal values of γ_k , we define the following indicator function:

$$I_{\gamma_k \leq c^{ij}} = \begin{cases} 1 & \text{if } k \in K_n^{ij} \\ 0 & \text{otherwise} \end{cases} \quad (8)$$

This function indicates that a treatment is bought without insurance.

The following result explains how treatments should be prioritized for coverage by basic insurance. Treatments are ranked on the weighted sum of shadow value and adverse selection terms, where weights are given by cost shares. We explain each of the terms below.

Proposition 1. Rank treatments k on the basis of

$$\sum_{i \in \{l, h\}} \left\{ \sum_{j \in \{l, h\}} CS_k^{ij} [(1 - I_{\gamma_k \leq c^{ij}}) \lambda^{ij} + I_{\gamma_k \leq c^{ij}}] + v^{il} CS_k^{il} AS_k^i \right\} \quad (9)$$

with $\gamma_k \in [v_{0k}, \delta_k]$ taking on one of the values $\{v_{0k}, c^{ll}, c^{lh}, c^{hl}, c^{hh}\}$, where the cost share CS_k^{ij} is defined as ij 's share of the government's cost of changing γ_k and AS_k^i is defined as the adverse selection effect of treatment k for income type i . These expressions are defined in Eqs. (A.3) and (A.4) in the appendix.

By subsidizing treatments through basic insurance, the government frees up agents' resources. These freed up resources can be spent by the agent on buying either additional "health" (insurance coverage for more treatments) or other goods. Further, basic insurance helps to reduce adverse selection by relaxing IC constraints. The proposition says that treatments k should be covered by basic insurance where the returns to agents' freed resources and relaxed IC constraints are highest.

To understand the expression in (9), first consider the case where IC constraints are not binding ($v^{il} = 0$). This is the case with risk rating, but can also happen with community rating (see Example 1 below).

Consider the case with only one type. Then $CS_k^{ij} = 1$ if treatment k is used and zero otherwise. If treatment k is covered by supplementary insurance, reducing γ_k frees up the agent's resources that can be spent on insuring (and hence getting access to) other treatments. The return on these freed resources is given by the health shadow value λ^{ij} . If, however, $\gamma_k < c^{ij}$, i.e. k is bought without insurance (paid out of $C^{ij} = c^{ij}$). Reducing γ_k only increases consumption on outside goods (weight one: $I_{\gamma_k \leq c^{ij}}$ is multiplied by 1 in the first line of (9)). That is, this does not free up money to buy more health insurance and get access to new treatments. Thus, $[(1 - I_{\gamma_k \leq c^{ij}}) \lambda^{ij} + I_{\gamma_k \leq c^{ij}}]$ gives ij 's return of a reduction in γ_k . Hence, if there is only one type, basic insurance should cover any treatment for which the agent buys supplementary insurance. The return λ^{ij} is the same for each of these treatments. Hence, it is not necessary to cover treatments that have high CE scores themselves. For treatments covered by basic insurance, γ_k should not fall below c^{ij} as this reduces the return on γ_k from $\lambda^{ij} > 1$ to 1.

²⁵ In a system where mandatory basic insurance is not financed by the government but sold by private insurers, the relevant constraint would be that low income people can afford to buy the basic insurance package. This leads to a similar analysis.

Table 2
Value, costs and probabilities for treatments K in Example 1.

K	v_{1k}	δ_k	θ_k^h	θ_k^l
1	30	10	$\frac{1}{10}$	$\frac{1}{10}$
2	20	10	0	0
3	15	10	$\frac{1}{5}$	$\frac{1}{5}$

Now consider more than one type ij . Then the sum of CS_k^{ij} over ij adds up to 1. Hence, we rank on the basis of a weighted average. Consider a type ij and treatment k with $\gamma_k > c^{ij}$. If ij has no insurance for k ($\rho_k^{ij} = 0$), she has no access to this treatment and a small reduction in γ_k has no effect on ij 's utility: ij 's cost share equals zero for this treatment, $CS_k^{ij} = 0$. If k is insured ($\rho_k^{ij} > 0$) then the cost share is positive, $CS_k^{ij} > 0$, and reducing γ_k leads to a reduction in ij 's insurance premium. The money saved on the premium can be used to buy insurance for other treatments. At the margin, this additional money has return equal to the health shadow value λ^{ij} . As above, if $\gamma_k < c^{ij}$ the return for type ij of reducing γ_k is only 1.

As long as there is one type with $\lambda^{ij} > 1$, Eq. (9) shows that there is an effect of budget constraints on the priority of treatments for basic insurance. If $\lambda^{ij} = 1$ for all types ij , the term in square brackets in (9) equals 1 for all ij and k . In words, access to care has no role to play anymore. We are back in the model of Boone (2015) where priority is determined by adverse selection effects captured by v^{il} – the second term in (9).

Treatments with high priority have the following characteristics: (i) covered by supplementary insurance ($\rho_k^{ij} > 0$) for types with the highest shadow value λ^{ij} ; (ii) especially used by this type: government's "bang for the buck" is higher, the higher is CS_k^{ij} . If $CS_k^{ij} = 1$, every euro spent on treatment k is spent on type ij with highest λ^{ij} . If other types use this treatment k , they add their shadow value (λ or 1) to the average; thereby reducing it.

From this observation, we get the result that the government should never cover cheap treatments in basic insurance. This can be seen as follows. If $c^{ij} = c$ for all ij , reducing γ_k below c is not optimal due to Assumption (5). This would yield a return equal to 1, while (5) implies that there are treatments that can be covered with a higher return. Hence $\gamma_k < c$ has a low ranking in Eq. (9). If not all c^{ij} are the same, $\gamma_k < c^{ij}$ for some ij is possible. However, basic insurance does not cover treatments with (very) low costs: $\delta_k < \min_{ij} c^{ij}$ as long as (5) holds. Since both the planner's objective and the budget constraint are linear in γ_k , γ_k can at max. have one of five different values: either v_{0k} (fully covered by basic insurance) or one of the values c^{ij} . It should also not fall below v_{0k} as $\gamma_k < v_{0k}$ induces inefficient use. If there would be no moral hazard ($v_{0k} = 0$), the government can reduce the co-payment of a treatment to 0.

Eq. (9) shows the importance of prevalence of a disease. Let λ^{mn} denote the highest λ^{ij} among all types. Note that λ^{ij} does not vary with k , it is the weight CS_k^{ij} that varies with k . Hence, basic insurance should cover treatments k where type mn has a high cost share. This is achieved by covering treatments k for diseases that are particularly prevalent under mn (compared to other types) and for which mn buys insurance $\rho_k^{mn} > 0$ and thus $CS_k^{mn} > 0$. As the next example illustrates, these are not necessarily treatments with a high CE score. Hence, the idea in the CE literature that the government should cover treatments with the highest CE score (Drummond et al., 2005; Gold et al., 1996) is not correct in a context where supplementary insurance is available.

Consider the extreme situation where a type, say lh , has such a low budget that it cannot afford any supplementary insurance (all money is needed to pay for co-payment γ_k in basic insurance). Then the logic above changes in the following way. This type will tend to have high λ . The only way to increase health for this type is to subsidize directly highly cost effective treatments used by this type. But also here we have the trade off: treatment k with high v_k/δ_k is attractive to cover by basic insurance but treatments with lower v/δ that are better targeted at this type lh can yield a higher health gain for the government as less of the budget is spent on types with lower λ .

Example 1. This example illustrates two points made above: (i) IC constraints do not necessarily bind with second degree price discrimination and (ii) basic insurance should not cover treatments with the highest CE score. To make these points, we assume that there are three treatments: $K = \{1, 2, 3\}$. These treatments do not suffer from moral hazard: $v_{0k} = \psi_k = 0$ for each $k \in K$. Co-payment equals $c = 0$ for each type. Table 2 gives the value created by each treatment, its cost and the probability that the treatment is needed for each θ type. People with low income have $\beta^l = 1$, $\phi^l = 0.5$ and with high income: $\beta^h = 2$, $\phi^h = 0.5$.

First, consider the case where there is no government budget ($B = 0$). Market equilibrium features: $\rho_1^l = \rho_1^h = 1$ and $\rho_2^l = \rho_2^h = \rho_3^l = \rho_3^h = 0$.²⁶ Hence the low income types are pooled on a contract with premium $\sigma^{lh} = \sigma^{lh} = \beta^l = 1$: due to the budget constraint there is no adverse selection for β^l types. This contract covers the treatment with the highest CE score: $v_{11}/\delta_1 = 3$. For the high income types we find the following: $\rho_1^{hh} = 1$, $\rho_2^{hh} = 1/2$, $\rho_3^{hh} = 0$ at a premium equal to $\sigma^{hh} = \beta^h = 2$; $\rho_1^{hl} = 1$, $\rho_2^{hl} = 0$, $\rho_3^{hl} = 1/2$ at a premium equal to $\sigma^{hl} = \beta^h = 2$. These contracts are incentive compatible.

Now suppose that the government has a small budget $B > 0$; which treatments should be covered by basic insurance? Based on the idea from the CE literature that the government should cover treatments with highest CE scores, one option is to cover (part of) treatment 1. This would free up money for all types to buy additional supplementary insurance. The θ^h types (both with low and high income) buy more coverage of treatment 2, the θ^l types buy more coverage of treatment 3. Hence, the health gain per unit of budget B spent is a weighted average of $v_{12}/\delta_2 = 2.0$ and $v_{13}/\delta_3 = 1.5$.

Eq. (9) suggests to cover treatment 2 in basic insurance. Marginal utility of income equals 2 for θ^h and 1.5 for θ^l . Hence, basic insurance should cover the treatment that is predominantly used by θ^h which is treatment 2. For $B > 0$ small enough, this yields a health gain of 2.0 per unit of the budget B spent. Hence covering treatment 1 by basic insurance is, in fact, suboptimal.

²⁶ The general characterization of supplementary insurance is given in Proposition 2 below. In this simple example it is routine to verify that these values of ρ_k^{ij} constitute an equilibrium.

Recall that the planner's goal (7) is to maximize welfare (efficiency) without any distributive objective. But in an access to care model, this efficiency objective endogenously gives the planner an incentive to redistribute towards types with high λ . To maximize the health impact of the budget B , basic insurance should cover/subsidize treatments used by types that *at the margin* buy treatments with a high CE score. As the example shows this is not the same as basic insurance covering treatments with high CE scores (as they may – also – be used by types with a low return λ at the margin). Section 7 derives a sufficient condition under which high λ types (“most deserving”) are, in fact, people with low income and low health (“most needy”).

Going back to Proposition 1, the ranking of treatments for basic insurance is based on the sum of two terms in Eq. (9). The second term shows that basic insurance should cover treatments with serious adverse selection effects. The gain from reducing adverse selection is given by the shadow price of (IC), v^{il} , times il 's cost share in k times the adverse selection term. The expression for AS_k^l in the appendix shows that it is made up of terms $\theta_k^h/\theta_k^l > 1$.

The higher θ_k^h/θ_k^l , *ceteris paribus*, the bigger the adverse selection term if the value of a reduction in γ_k is higher for ih when choosing her own contract than when mimicking il . Indeed, this reduction in γ_k makes it less attractive for ih to buy il 's contract and this effect is stronger the higher θ_k^h/θ_k^l ; this can be seen as follows. By buying the il contract, the premium for the θ^h type is determined by θ_k^l instead of θ_k^h . This discount on the premium increases with θ_k^h/θ_k^l . Hence, θ_k^h/θ_k^l captures the attractiveness for ih of buying the il contract. This attractiveness is reduced if the government lowers γ_k as it reduces the premium differential between the two contracts. Roughly speaking, $(\theta_k^h - \theta_k^l)\gamma_k$ falls with a reduction in γ_k . If mimicking becomes less attractive, the distortion in the il contract can be reduced. The value of reducing this distortion equals v^{il} .

As mentioned, in a standard model with moral hazard and adverse selection this adverse selection effect is the deciding factor in choosing treatments for basic insurance (Boone, 2015). In an access to care model, basic insurance should cover treatments that (i) help to reduce the adverse selection inefficiency and (ii) are predominantly used by agents with high λ^{ij} . The next section shows how adverse selection affects the equilibrium in the supplementary market.

6. Supplementary insurance

As shown in Proposition 1, treatment k 's CE score is not leading when it comes to basic insurance. Here we show that CE scores are important for supplementary insurance. We define the following CE score for treatment k with co-payment c^{ij} and cost on the private market γ_k .

$$s_k(\gamma_k, c^{ij}) = \begin{cases} \frac{v_{1k} - \gamma_k}{\gamma_k - c^{ij}} & \text{if } c^{ij} \geq v_{0k} \\ \frac{v_k - \gamma_k}{\gamma_k - c^{ij}} & \text{if } c^{ij} < v_{0k} \end{cases} \quad (10)$$

where we take c^{ij} as given; Section 8 endogenizes c^{ij} . Treatments that are not covered by basic insurance have $\gamma_k = \delta_k$ and Eq. (1) then implies $s_k < 0$ for $c^{ij} < v_{0k}$.

Proposition 2. *Given c^{ij} , supplementary insurance does not cover treatments with $\gamma_k < c^{ij}$. For θ^h agents, insurers rank treatments on the basis of their CE score*

$$s_k(\gamma_k, c^{ih}) \quad (11)$$

For θ^l types, insurers rank treatments on the basis of the adjusted CE score

$$s_k(\gamma_k, c^{il}) \left(1 - v^{il} \left[1 + \frac{1 + s_k(\gamma_k, c^{il})}{s_k(\gamma_k, c^{il})} \left(\frac{\theta_k^h}{\theta_k^l} - 1 \right) \right] \right) \quad (12)$$

Supplementary insurance for type ij covers the highest ranking treatments until agent ij 's budget $\beta^i - c^{ij}$ runs out.

With $v^{il} = 0$, both θ^l and θ^h types use supplementary insurance to cover treatments with the highest CE scores (10). Eq. (5) rules out the case where enough budget is available to cover all valuable treatments. Hence, agents choose the treatments that give the highest (health) utility per unit of their budget spent. In an insurance model without budget constraints, all valuable treatments can be covered and hence there is no reason to prioritize on the basis of CE scores (Boone, 2015). Treatments with $c^{ij} < v_{0k}$ will be used inefficiently and have a low health gain v_k in Eq. (10). Hence, priority to cover such treatments in supplementary insurance is low.

Although this is not mentioned in handbooks on CE analysis, health insurers offering supplementary insurance should correct the CE score for the co-payment c^{ij} . The correct CE score is not simply v/δ ,²⁷ but the expression in Eq. (10). This is a minor adjustment but it does give rise to an interesting inverse-U relation between cost of treatment and coverage by insurance. To see this, fix $c^{ij} \geq v_{0k}$ and $v_{1k} > c^{ij}$ and let γ_k increase from v_{0k} . Then for $\gamma_k < c^{ij}$, the treatment should not be covered by supplementary insurance: people can pay the treatment out-of-pocket. That is, relatively cheap treatments should not be covered by supplementary insurance at all. Treatments with $\gamma_k > c^{ij}$ close to c^{ij} will certainly be covered by insurance as the ratio in (10) is very high for such treatments. A small investment (in insurance premium) $\gamma_k - c^{ij}$ leads to a big health gain. Then as γ_k keeps rising (for given v_{1k}), the CE score deteriorates and k drops in priority. For γ_k high enough, the treatment will not be covered by insurance. A second observation is that CE scores actually differ between persons with different co-payments c^{ij} . Hence there is not a unique CE score for a treatment.

Finally, treatments with severe moral hazard problems ($v_{0k} > c^{ij}$) and $\gamma_k = \delta_k$ are not covered by supplementary insurance. Hence in this model, moral hazard affects the extensive margin of insurance. In a standard model, moral hazard with respect to treatment k leads to marginally lower insurance by requiring a higher co-payment $c_k^{ij} \geq v_{0k}$ (Pauly, 1968; Zeckhauser, 1970). That is, moral hazard affects the

²⁷ Clearly, ranking on the basis of $(v - \delta)/\delta$ or inverse ranking on the CE ratio δ/v gives the same results as ranking on the basis of v/δ .

Table 3Value, costs and probabilities for treatments K in Example 2.

K	v_{1k}	δ_k	θ_k^h	θ_k^l
1	50	10	$\frac{4}{10}$	$\frac{2}{10}$
2	40	10	$\frac{5}{10}$	$\frac{1}{10}$
3	30	10	$\frac{3}{10}$	$\frac{1}{10}$
4	20	10	$\frac{4}{10}$	$\frac{4}{10}$
5	15	10	$\frac{1}{10}$	$\frac{1}{10}$

Table 4

Market outcome in Example 2.

	β^l						β^h					
	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	λ	ρ_1	ρ_2	ρ_3	ρ_4	ρ_5	λ
θ^l	$\frac{3}{4}$	0	0	$\frac{5}{8}$	0	$\frac{5}{4}$	1	0	$\frac{1}{6}$	1	$\frac{1}{3}$	$\frac{6}{5}$
θ^h	1	0	0	0	0	4	1	$\frac{1}{2}$	0	0	0	4

intensive margin only, not the extensive margin. In a model with binding budget constraints, c^{ij} does not vary with k and hence $v_{0k} > c^{ij}$ implies that treatment k is not covered at all.

The effect of $v^{il} > 0$ on the treatment ranking for supplementary insurance for θ^l is as follows. For $v^{il} > 0$ there is a “correction term” in Eq. (12). This correction disappears if $\theta_k^h = \theta_k^l$: treatments that are used equally by θ^h and θ^l types, do not cause adverse selection issues and hence do not need to be distorted in supplementary insurance.²⁸ However, if $\theta_k^h > \theta_k^l$, the use of this treatment will be distorted in order to separate the θ^l from the θ^h contract. The distortion takes the form of pushing this treatment down the ranking for θ^l . Hence θ^l may end up without coverage of this treatment even though its CE score (10) is quite high. Treatments k with high $(1 + s_k)/s_k$ (i.e. low s_k) get distorted more in the sense of being pushed down further in the ranking (for given θ_k^h/θ_k^l). To illustrate, if $v - \gamma$ is high, θ^h does not want to miss out on this treatment and a small distortion on this treatment is enough to stop θ^h from buying the θ^l -contract. If $\gamma_k - c$ is small, the gain in insurance premium of mimicking is small for the θ^h type (this gain is of the form $(\theta_k^h - \theta_k^l)(\gamma_k - c^{il})$). In either case, s_k is rather high and treatment k does not need to be distorted much to get separation.

7. Redistribution

The following lemma gives a sufficient condition for basic insurance to focus on the “most needy” agents: low income and high risk. That is, basic insurance should cover treatment of conditions that are prevalent under low income/high risk agents.

Lemma 1. Assume that $C^{ij} = c^{ij} = c$ for all types ij . Then $\lambda^{lh} \geq \lambda^{hh}$ and $\lambda^{ih} \geq \lambda^{il}$ for $i \in \{l, h\}$.

The co-payment c can be the same for all types for two reasons. First, the government mandates co-payment c for supplementary insurance. Second, all types choose the same co-payment (see next section).

The lemma shows that we can rank the health shadow prices $dV^{ij}/d\beta^i = \lambda^{ij}$. Type lh has the highest shadow value of all four types. Further, within an income class i the θ^h type has higher value λ than the θ^l type. This happens for two reasons. First, an agent with either a higher income or lower probabilities θ_k can cover more treatments with supplementary insurance. As the treatments are ranked, this implies that the agent with a higher budget and/or lower risk covers lower ranked treatments at the margin. Thus, the marginal value falls. Second, θ^l types use a distorted ranking (12) and hence, at the margin, cover treatments with lower value.

Comparing the shadow prices for two θ^l types (one with high, the other with low income) is not straightforward due to the interaction of the budget and IC constraints. As β increases (from β^l to β^h) the budget constraint relaxes. Because $dV^{ih}/d\beta^i > 1$, we have $d(V^{ih} - \beta^i)/d\beta^i > 0$. Hence also the IC constraint in (6) is relaxed in response to an increase in β^i . As both the budget and the IC constraints are relaxed, it is natural to expect that both shadow prices (λ^{il} , v^{il}) fall as well.²⁹

Combining Proposition 1 and Lemma 1, basic insurance’s focus on treatments k with high prevalence among high risk types, leads to a focus on treatments where θ_k^h/θ_k^l is high. Basic insurance covering treatments with high θ_k^h/θ_k^l for which θ^h types buy insurance, helps to target government resources at high risk types. This frees up money for these people to buy additional insurance which leads to high health gains at the margin (high λ).

This observation is significant in the light of the CE literature. As discussed in the introduction, the CE literature tends to view distribution as orthogonal to maximizing health gain per unit of the government budget spent. In an access to care model, by maximizing total welfare (7), the planner endogenously takes distribution into account. Government budget should be targeted at agents with highest λ .

The next example illustrates the case where for the θ^l types both IC and budget constraints are binding.

Example 2. Consider the set of five treatments in Table 3. Assume $v_{0k} = \psi_k = c = 0$ for each $k \in K$. Hence treatment 1 has the highest CE score v/δ and treatment 5 the lowest. Assume $\beta^l = 4$, $\beta^h = 6.5$. We start by assuming $B = 0$.

Contracts offered in the market are given in Table 4. Note that we find $\lambda^{il} \leq \lambda^{lh}$: as θ^h spends her income on treatments with the highest CE scores, her marginal health return to income is higher; this is true in each income class $i \in \{l, h\}$. The high income θ^l type has lower values for λ and v than the low income θ^l type as an increase in β relaxes both the budget and the IC constraints.

²⁸ Note that v^{il} does not vary with k .

²⁹ Although we do not use this result below, for completeness Lemma 3 in Appendix C derives necessary and sufficient conditions for λ^{il} , v^{il} to fall as income increases.

Table 5Value, costs and probabilities for treatments K in Example 3.

K	v_{1k}	v_{0k}	δ_k	θ_k^h	θ_k^l
1	150	1.9	10	0.2	0.0
2	15	0	10	0.1	0.1
3	15	0	10	0.1	0.1

If the government increases its budget $B > 0$, it should subsidize treatment 2, not treatment 1 with the highest CE score. The reason is that treatment 2 is better targeted at type θ^h because $\rho_2^l = 0$ for $i = l, h$. Hence each unit of the budget is spent on a type with $\lambda = 4$.³⁰

8. Optimal co-payments

Up till now we have taken c^{ij} as given. In this section, we endogenize the co-payment. In an RS style model, θ^l types face higher co-payments than θ^h : in this way insurers separate high from low risk types. In an access to care model, there are two ways to separate types: higher co-payments and covering a different set of treatments. The following result derives sufficient conditions for θ^l to prefer higher co-payments than θ^h as is the case in RS.

We use the following notation. We compare two arbitrary co-payments $\underline{c}$ and $\bar{c} > \underline{c}$. Let $\underline{V}^{ij}$ denote the (max.) utility for ij from a contract with $\underline{c}$. And $\bar{V}^{ij}$ from a contract with co-payment $\bar{c}$. The sets in Eq. (3) are denoted by $\underline{K}(\bar{K})$ for the low (high) co-payment.

Proposition 3. Assume that (i) $\underline{K}_i = \emptyset$ and (ii) for each $k \in \bar{K}_n \cap \underline{K}_e$ it is the case that $\theta_k^h = \theta_k^l$. Then we have the following

$$\begin{aligned} \bar{V}^{hh} - \underline{V}^{hh} &\geq \bar{V}^{lh} - \underline{V}^{lh} \\ \bar{V}^{il} - \underline{V}^{il} &\geq \bar{V}^{ih} - \underline{V}^{ih} \end{aligned}$$

for each income group $i \in \{l, h\}$.

The proposition compares the change in utility under the high and low co-payment for different types. It shows high risk types with high income have a stronger preference for high co-payments than high risk types with low income. Further, within an income class, low risk types prefer the higher co-payment more than high risk types.

Increasing c leads to the following trade-off at the individual level. As c increases, more money $C^{ij} = c$ needs to be set aside to cover the co-payment c and hence less money is available to spend on the insurance premium which needs to be reduced to stay within the agent's budget. This has the following implications for the treatments k that are covered. As c increases, more –relatively cheap– treatments ($\gamma_k \leq c$) can be covered without insurance (the set K_n). However, due to the reduced premium, coverage of some expensive treatments needs to be reduced. At the margin, the loss of health due to reductions in ρ^{ij} is given by λ^{ij} . In other words, the opportunity cost of increasing C^{ij} is given by λ^{ij} . This is weighed against the observation that an increase in c implies that more efficient treatments become available: $\bar{K}_e \cup \bar{K}_n \supseteq \underline{K}_e \cup \underline{K}_n$.

As implied by Lemma 1, for a given c the λ^{ij} effect works in the direction of θ^l preferring higher c than θ^h : opportunity cost of increasing c is lower. Similarly, a high income θ^h prefers higher c than a low income θ^h type.

This intuition is correct under two assumptions. Example 3 below shows that the result does not hold if assumption (i) is not satisfied. It is straightforward to see that a violation of assumption (ii) can lead to a situation where θ^h prefers a higher c than θ^l . Assume that by increasing c only one thing happens: treatment k moves from K_e to K_n as agents have higher C^{ij} . If for this treatment $\theta_k^h > \theta_k^l = 0$, the increase in c has some value for θ^h (as k can now be financed directly out of C^{ij}). But because $\theta_k^l = 0$, this increase in c has not much value for the θ^l type. Hence, it can happen that θ^h has a higher preference for this increase in c than θ^l . If assumption (ii) holds, this cannot happen.

The θ^l agent has an additional reason for high c : it facilitates separation from θ^h (standard RS argument). As can be seen in Eq. (IC), higher c^{il} relaxes the IC constraint – for given sets K_i, K_e, K_n .

Example 3. Consider the treatments in Table 5. We consider only one income class with $\beta = 2$. For treatment 1 we assume $\psi_1 = 0.95$. Covering treatment 1 destroys value unless the co-payment is at least equal to 1.9. Insurance covering 1 with $c = 1.9$ costs θ^h : $0.2 * (1 - 0.95)(10 - 1.9) + 1.9 = 2$ (approx.) and yields higher utility for θ^h than buying insurance covering treatments 2 and 3.

At $\underline{c} = 0$ it is the case that $\underline{K}_i = \{1\}$; hence assumption (i) in Proposition 3 is not satisfied. Type θ^h strictly prefers $\bar{c} = 1.9$ above $\underline{c}$. In this example, θ^h has a preference for higher c than θ^l . In contrast to models in the line of RS, in an access to care model θ^h can have a higher co-payment (for all treatments) than θ^l .

Finally, if co-payments differ between contracts, different types set aside different amounts of money C^{ij} . The following example illustrates that the result in Lemma 1 does not necessarily hold if C^{ij} differ. The consequence can be that the government wants to redistribute towards low risk types to maximize the health impact of budget B .

Example 4. Consider the treatments in Table 6 and assume $B = 0$. We focus on one income class with budget $\beta = 2$. Treatments 1 and 3 do not feature moral hazard problems, but treatment 2 does. We assume that $\psi_2 = 4/5$. Hence $\psi_2 * 1 + (1 - \psi_2) * 40 < \delta_2$: treatment 2 can only be covered with a co-payment equal to 1.

It is optimal for θ^h to buy insurance coverage for treatment 1. Hence, she chooses a contract with $c^h = 0$, $\rho_1^h = 1$. She spends her whole budget on this. If she would get a small increase in income, she cannot afford the co-payment to make treatment 2 efficient and hence buys insurance for treatment 3: $\lambda^h = 3$.

Type θ^l buys coverage for treatment 2 with $c^l = 1$, $\rho_2^l = 4/30$. Hence increasing θ^l 's income marginally, allows her to extend her coverage of treatment 2: $\lambda^l = 4 > \lambda^h$.

³⁰ This is the value of λ^{lh} for a small increase in β^l .

Table 6
Value, costs and probabilities for treatments K in Example 4.

K	v_{1k}	v_{0k}	δ_k	θ_k^h	θ_k^l
1	50	0	10	$\frac{2}{10}$	0
2	40	1	10	$\frac{3}{4}$	$\frac{3}{4}$
3	30	0	10	$\frac{1}{10}$	0

The intuition why the government may decide to prioritize funding of treatments used by low risk types is the following. If c^{ij} are allowed to differ between types, it can happen that θ^l prefers a high co-payment and in this way makes a treatment k available in an efficient way. Suppose this treatment is highly cost effective (high v_k/δ_k) if consumption in state 0 is avoided by means of a high co-payment. If θ^l can but θ^h cannot afford this high co-payment, the health return to income can be higher for θ^l than for θ^h . Without the co-payment, the treatment is not cost effective and the government cannot make this treatment available to θ^h . If $B > 0$, basic health insurance should target treatments that are mainly used by θ^l in this case because $\lambda^l > \lambda^h$.

Hence with the model above, redistribution from low income/high risk types to high income/low risk types is possible. It is an empirical question to determine which type has the highest return to an additional dollar spent on health care.

9. Policy implications

A number of countries feature a combination of public basic insurance and private supplementary insurance. A standard model based on moral hazard and adverse selection suggests that basic insurance should cover treatments that suffer most from adverse selection (Boone, 2015). However, many governments use cost effectiveness as criterion for inclusion in basic health insurance (Thomson et al., 2012). This suggests that the standard framework misses important aspects of the problem.

We introduce an access to care model. In this model CE analysis does play a role. But basic insurance should not simply cover treatments with the highest CE scores. The main insight is that prevalence is important, even though the planner's objective function does not feature redistribution as an explicit goal.

When deciding on (changes in) the basic insurance package, two empirical issues need to be resolved. First, which types feature –at the margin– the highest health return to an increase in their budget. Second, which diseases (and their treatments) are prevalent among these types. Basic insurance should focus on coverage of these treatments.

For the first step, the data set that is used to estimate the risk adjustment model for a country is useful. It allows for the identification of types as in the risk adjustment system. The more detailed this system is, the better the planner can target. To illustrate, the Dutch risk equalization scheme features among others 38 age/gender groups, 20 groups based on past use of pharmaceuticals, 13 groups based on past diagnoses, 12 groups of social economic status (van de Ven and Schut, 2011). This creates a detailed grid of types for the Dutch population.

The question is the following. Consider a treatment k that type i currently covers with supplementary insurance. If treatment k would be covered by basic insurance, this frees up financial resources for type i ; how would these be spent? Would this be spent to create (more) access for i to treatment $\tilde{k}$; if so, what is $\tilde{k}$'s return in terms of health? CE analysis is needed to determine $\tilde{k}$'s health return. If, on the other hand, the freed up resources are spent on skiing holidays and/or cigarettes, the health return may be low. In this way, the government can determine the types with the highest health return per dollar spent on basic insurance.

For the second step, epidemiological studies are needed that identify the prevalence of diseases. Such studies do not only consider which conditions/symptoms tend to be correlated but also consider the prevalence of diseases among social classes, education groups, regions etc. (see e.g. Conen et al., 2009; Dalstra et al., 2005; Leyland, 2005; Van Loon et al., 1998). Hence this type of study can be done, although we are not aware of studies using the detailed type space used in risk adjustment.

Once the types with the highest health return are identified and the treatments that are more prevalent among these than among other types, these treatments should be covered by basic health insurance. This gives the government the biggest health gain per dollar spent on basic health insurance.

Finally, should the government intervene in the health insurance market at all by offering basic insurance? In a standard insurance model, the welfare gain of such intervention is the reduction of the adverse selection inefficiency: under-insurance for low risk types. As argued by, for instance, (Pauly, 2008, pp. 121) the adverse selection inefficiency is rather small and hence competitive health insurance markets do a good job. Then there is not much need for government intervention. Indeed, if insurers are allowed to risk rate, the market outcome is efficient in the standard model and there is no reason at all for a welfare maximizing government to intervene.

However, in an access to care model the return to government intervention is much higher: it is related to the health benefit of treatments that (low income/high risk) agents cannot afford. This is true both with and without risk rating by insurers. Section 3 cited papers documenting that the uninsured in the US go without essential (and highly cost effective) health care. Then the return to government intervention is magnitudes higher than in a model without budget constraints.

One extension of the model is the following. We assumed that basic insurance was universal: accessible for everyone on the same terms. This is, for instance, the case in the Netherlands. However, other countries vary the co-payment of basic insurance (γ_k in the model above) with agent characteristics like income, age and health status.³¹ If the planner in the model above can vary payments γ_k in basic insurance by type ij , would this increase welfare? Indeed, it would allow the planner to target resources better. If the health return to income (λ) is higher for the low income, high risk (lh) type, the planner can set γ_k^{lh} lower than for other types. This introduces trade offs that we did not model above. To illustrate, if the co-payment in basic insurance varies with income, this introduces a marginal tax on income. Hence, targeted health insurance must be weighed against distortions in the labour market. As long as everyone's co-payment lies below the social

³¹ See, for instance, https://en.wikipedia.org/wiki/Prescription_charges (accessed March 30, 2018) for a description of the NHS. I thank an anonymous referee for pointing this out.

treatment cost, the main result above still applies: cover treatments in basic insurance that are mainly used by agents with a high health return on income.

Another extension of the model is left for future research: the assumption that the agent's condition is verifiable by a physician. There is an asymmetry here that we do not deal with in this paper. On the one hand, an agent can claim to have a disease and the physician can verify whether this is indeed the case. On the other hand, if an agent has two conditions she may decide to hide the symptoms of one of the conditions.³² Above we assume that the physician can verify all conditions that the agent suffers from. One can argue that it is easier for a physician to verify whether the patient suffers from a disease that she claims she has than to find the condition that the patient is hiding. Taking this asymmetry into account when designing the optimal insurance system is left for future research.

Appendix A. Proof of results

Lemma 2. *With second degree price discrimination, θ^h and θ^l are not pooled into one contract (in equilibrium) where at least one treatment k is covered with $\theta_k^h > \theta_k^l$.³³ At max. two IC constraints are binding: lh wants to mimic ll and hh wants to mimic hl .*

The proof of this lemma uses standard arguments and can be found in [Appendix C](#).

Proof of proposition 1 The government sets the values for γ_k . Given these values, the market offers equilibrium contracts that maximize V^{ih} , V^{il} as given by Eqs. (4) and (6) resp. However, it is easier to think of the planner as setting γ_k , c^{ij} and ρ_k to maximize welfare (7) while taking the same budget constraints and IC constraints into account that the market does. This allows us to use the envelop theorem as explained in [Appendix B](#).

Let μ denote the Lagrange multiplier on constraint (GBC). In order to derive the first order condition for γ_k , we use the indicator function defined in the main text and the following indicator function:

$$I_k^{ij} = \begin{cases} 1 - \psi_k & \text{if } k \in K_e^{ij} \\ 1 & \text{if } k \in K_i^{ij} \\ 0 & \text{otherwise} \end{cases} \quad (\text{A.1})$$

The function I_k^{ij} gives the fraction of times that an insured treatment is used (conditional on the condition occurring). Treatment $k \in K_e^{ij}$ is only used in state 1 and $k \in K_i^{ij}$ is used in both states 0 and 1.

The first order condition for γ_k can be written as

$$\begin{aligned} & \sum_i F^i \left[\sum_j \phi^j \theta_k^j \left(I_k^{ij} \rho_k^{ij} \lambda^{ij} + (1 - \psi_k) I_{\gamma_k \leq c^{ij}} \right) \right] \\ & - V^{il} \phi^l \left(\theta_k^{il} \rho_k^{il} \hat{\lambda}^{il} + (1 - \psi_k) I_{\gamma_k \leq c^{il}} \theta_k^{ih} - \theta_k^{ih} \rho_k^{ih} I_{\gamma_k \leq c^{ih}} \lambda^{ih} + I_{\gamma_k \leq c^{ih}} (1 - \psi_k) \right) \\ & = \mu \sum_i \sum_j F^i \phi^j \theta_k^j \left(I_k^{ij} \rho_k^{ij} + I_{\gamma_k \leq c^{ij}} (1 - \psi_k) \right) \end{aligned} \quad (\text{A.2})$$

where $\hat{\lambda}^{il}$ denotes the return for the mimicking ih type of re-optimizing il 's contract after γ_k is reduced –which is optimized for il , not for ih .

Note that γ_k does not appear in this condition. In other words, the planner's objective function is linear in γ_k . Thus we find a bang-bang solution for γ_k : it either equals v_{0k} , c^{ll} , c^{lh} , c^{hl} , c^{hh} or k is not covered: $\gamma_k = \delta_k$ (if the right hand side –costs– exceeds the left hand side –benefits). For the last treatment k to be covered (before the government hits its budget constraint) the left hand side equals the right hand side. The government then sets γ_k equal to the relevant c^{ij} or v_{0k} and the probability that the treatment is covered $\rho_k^g \in (0, 1]$ such that the budget B is exactly spent.

Dividing Eq. (A.2) by the total cost of a reduction in γ_k , $\sum_i \sum_j F^i \phi^j \theta_k^j \left(I_k^{ij} \rho_k^{ij} + I_{\gamma_k \leq c^{ij}} (1 - \psi_k) \right)$, we find that the government ranks treatments using expression (9). The government then covers the treatments that rank highest on this measure until the budget B is spent. To see this, we define the cost share as follows:

$$CS_k^{ij} = \frac{F^i \phi^j \theta_k^j (I_k^{ij} \rho_k^{ij} + (1 - \psi_k) I_{\gamma_k \leq c^{ij}})}{\sum_{i'} \sum_{j'} F^{i'} \phi^{j'} \theta_k^{i'} (I_k^{i'j'} \rho_k^{i'j'} + (1 - \psi_k) I_{\gamma_k \leq c^{i'j'}})} \quad (\text{A.3})$$

with $F^l = F$, $F^h = 1 - F$, $\phi^l = \phi$, $\phi^h = 1 - \phi$. Thus, for each treatment k we take the average of the returns $[(1 - I_{\gamma_k \leq c^{ij}}) \lambda^{ij} + I_{\gamma_k \leq c^{ij}}]$ over ij with weight CS_k^{ij} . The denominator of CS_k^{ij} is the total cost for the government of reducing γ_k by 1 euro. The numerator of CS_k^{ij} is ij 's part of this cost. In other words, we take the weighted average of the return to a reduction in γ_k where type ij 's weight equals her share in the cost of this reduction in γ_k .

³² To illustrate, assume that treatment for condition 1 has an affordable co-payment only if the agent does not suffer from condition 2 as well. An agent suffering from conditions 1 and 2, may then decide to hide condition 2 to get access to treatment 1.

³³ It is possible that θ^h and θ^l are pooled on a contract where the treatments covered feature $\theta_k^h = \theta_k^l$: see Example 1. Given their budget constraint, there is no information asymmetry for these types.

We define the adverse selection term as

$$AS_k^i = \frac{\frac{\theta_k^h}{\theta_k^l} [I_k^{ih} \rho_k^{ih} \lambda^{ih} + (1 - \psi_k)(I_{\gamma_k \leq c^{ih}} - I_{\gamma_k \leq c^{il}})] - \rho_k^{il} I_k^{il} \hat{\lambda}^{il}}{I_k^{il} \rho_k^{il} + (1 - \psi_k) I_{\gamma_k \leq c^{il}}} \quad (\text{A.4})$$

Two terms determine the extent to which a reduction in γ_k reduces the adverse selection inefficiency. First, there is

$$\frac{\theta_k^h}{\theta_k^l} I_k^{ih} \rho_k^{ih} \lambda^{ih} - \rho_k^{il} I_k^{il} \hat{\lambda}^{il} \quad (\text{A.5})$$

which gives the effect of a reduction in γ_k on ih 's incentive to mimic il . In particular, a reduction in γ_k increases ih 's utility by $\theta_k^h I_k^{ih} \rho_k^{ih} \lambda^{ih}$: the reduction in ih 's insurance premium times ih 's health shadow value λ^{ih} . If ih mimics il , the fall in premium is given by $\theta_k^l \rho_k^{il} I_k^{il}$. But, because il 's insurance contract is not optimized for ih , the return on this additional income is not λ^{ih} (nor does it equal λ^{il}). As the premium falls, the contract is re-optimized for il which gives a mimicking return $\hat{\lambda}^{il}$ for ih . If (A.5) is positive, the reduction in γ_k reduces the incentive to mimic and hence increases the priority of k to be covered by basic insurance.

In a standard model with moral hazard and adverse selection this is the deciding factor in choosing treatments for basic insurance (Boone, 2015). There the effect always works in the direction of covering treatments with highest θ_k^h / θ_k^l by basic insurance. Here, however, because agents are budget constrained, $\theta_k^h / \theta_k^l > 1$ can work in the opposite direction for some treatment k . Due to budget constraints, there can be treatments k with $\rho_k^{il} > \rho_k^{ih}$ making the expression in (A.5) negative. Then a reduction in γ_k makes mimicking more attractive as it gives mimicking ih cheaper access to k . $v^{il} > 0$ implies a lower ranking for treatments where (A.5) is negative.

The second term focuses on uninsured care and can also encourage mimicking. This happens if

$$I_{\gamma_k \leq c^{ih}} - I_{\gamma_k \leq c^{il}} < 0 \quad (\text{A.6})$$

Uninsured care is only used in state 1 and can be afforded if $\gamma_k \leq c$ for that contract.³⁴ Hence, for a treatment k with $\gamma_k > c^{ih}$ that is not insured in ih 's contract and features $\gamma_k \leq c^{il}$, a reduction in γ_k makes mimicking more attractive as access to k becomes cheaper in the mimicking contract.

Although we discussed Eqs. (A.5) and (A.6) separately, the sum is relevant. That is, it can also happen that k is insured under ih 's contract ($\rho_k^{ih} > 0$) but bought without insurance under il 's contract ($\rho_k^{il} = 0$, $I_{\gamma_k \leq c^{il}} = 1$) etc. Q.E.D.

Proof of proposition 2 Note that treatments with $\gamma_k \leq c^{ij}$ do not need to be insured as they are financed out of C^{ij} when needed.

First, consider a type θ^h : ρ^{ih} is chosen to maximize (4). It is routine to verify that the first order condition for ρ^{ih} can be written as

$$\rho^{ih} \begin{cases} = 1 & \text{if } s_k(\gamma_k, \rho^{ih}) > \lambda^{ih} \\ \in [0, 1] & \text{if } s_k(\gamma_k, \rho^{ih}) = \lambda^{ih} \\ = 0 & \text{if } s_k(\gamma_k, \rho^{ih}) < \lambda^{ih} \end{cases} \quad (\text{A.7})$$

Put differently, treatments are ranked according to $s_k(\gamma_k, c^{ij})$. Treatments with highest rank are insured ($\rho_k^{ih} = 1$) until the agent's budget β^i runs out.

For θ^l , ρ^{il} maximizes (6). With a similar reasoning as above, the first order condition then implies that treatments are ranked on the basis of

$$s_k(\gamma_k, c^{il}) - v^{il} \left(\frac{\theta_k^h}{\theta_k^l} (s_k(\gamma_k, c^{il}) + 1) - 1 \right) \quad (\text{A.8})$$

It is routine to verify that this equation can be written as (12). Q.E.D.

Proof of lemma 1 The inequality $\lambda^{ih} \geq \lambda^{il}$ follows from the following observations. Both θ^h types rank treatments on the basis of $s_k(\gamma_k, c)$ and cover the highest ranking treatment by supplementary insurance. If k denotes the last treatment covered by ih , then

$$\lambda^{ih} = s_k(\gamma_k, c) \quad (\text{A.9})$$

As they have the same value for $C^{ij} = c$ by assumption, they have (efficient) access to the same treatments. The θ^h type with the higher income can cover more treatments and hence has a lower λ .

Similarly, for θ^l type with k the last treatment for which she buys insurance:

$$\lambda^{il} = s_k(\gamma_k, c) \left(1 - v^{il} \left(\left(\frac{\theta_k^h}{\theta_k^l} - 1 \right) \frac{s_k(\gamma_k, c) + 1}{s_k(\gamma_k, c)} + 1 \right) \right) \quad (\text{A.10})$$

Within the same income class β^i , we have $\lambda^{ih} \geq \lambda^{il}$. This follows from Eqs. (A.9) and (A.10) in two steps. First, for given treatment k the right hand side of (A.9) is at least as big as the right hand side of (A.10). Hence, if both types covered the same treatment at the margin, the claim is true. Second, θ^l covers a treatment at the margin which is lower in priority than θ^h because: (i) θ^l buys coverage for treatments k at a premium based on probabilities $\theta_k^l \leq \theta_k^h$ and hence can cover more treatments, (ii) θ^l buys reduced (distorted) coverage of high priority treatments (based on (A.9)) to get separation from the θ^h contract and hence spends more money (than θ^h) on treatments with lower priority s_k . Q.E.D.

³⁴ Lemma 2 shows that agents do not buy contracts aimed at people with lower income.

Proof of proposition 3 In order to derive a “marginal” condition, we introduce the following “lottery contract”. Fix two values $\bar{c} > \underline{c} \geq 0$. The agent reserves an amount $c_\alpha = \alpha\bar{c} + (1-\alpha)\underline{c}$ for co-payments. This c_α is the price of a lottery ticket which is linked to health insurance in the following way. Before learning the realization of the lottery, the agent buys insurance coverage ρ_k . After the insurance is bought –but before treatment is used– the payoff from the lottery is realized. The lottery pays out $\bar{c}(\underline{c})$ with probability $\alpha(1-\alpha)$. If the lottery pays $\bar{c}(\underline{c})$, the agent faces co-payment $\bar{c}(\underline{c})$ when using treatment k . The relevant treatment sets are denoted $\bar{K}_{e,i,n}(\underline{K}_{e,i,n})$. For given α , if there are no IC constraints, agent ij solves

$$\begin{aligned} V_{\alpha}^{ij} = & \max_{\rho_k} \beta^i + \alpha \left(\sum_{k \in \bar{K}_e} \theta_k^j (1 - \psi_k) \rho_k (v_{1k} - \gamma_k) + \sum_{k \in \bar{K}_i} \theta_k^j \rho_k (v_k - \gamma_k) + \sum_{k \in \bar{K}_n} \theta_k^j (1 - \psi_k) (v_{1k} - \gamma_k) \right) \\ & + (1 - \alpha) \left(\sum_{k \in \underline{K}_e} \theta_k^j (1 - \psi_k) \rho_k (v_{1k} - \gamma_k) + \sum_{k \in \underline{K}_i} \theta_k^j \rho_k (v_k - \gamma_k) + \sum_{k \in \underline{K}_n} \theta_k^j (1 - \psi_k) (v_{1k} - \gamma_k) \right) \\ & - \lambda^{ij} \left(\alpha \left(\sum_{k \in \bar{K}_e} \theta_k^j (1 - \psi_k) \rho_k (\gamma_k - \bar{c}) + \sum_{k \in \bar{K}_i} \theta_k^j \rho_k (\gamma_k - \bar{c}) \right) \right. \\ & \left. + (1 - \alpha) \left(\sum_{k \in \underline{K}_e} \theta_k^j (1 - \psi_k) \rho_k (\gamma_k - \underline{c}) + \sum_{k \in \underline{K}_i} \theta_k^j \rho_k (\gamma_k - \underline{c}) \right) - (\beta - c_\alpha) \right) \end{aligned} \quad (A.11)$$

Clearly, $c = \bar{c}$ leads to higher utility than $c = \underline{c}$ if and only if $V_{\alpha=1}^{ij} > V_{\alpha=0}^{ij}$. But because the optimal ρ_k 's are different at $\alpha = 1$ than at $\alpha = 0$, the condition $V_{\alpha=1}^{ij} > V_{\alpha=0}^{ij}$ does not easily convey the relevant trade-offs. Therefore, we consider $\alpha \in [0, 1]$ and evaluate the marginal effect $d\alpha$. Then we can use (with a slight abuse of notation) $V_{\alpha=1}^{ij} - V_{\alpha=0}^{ij} = \int_0^1 dV_{\alpha}^{ij}/d\alpha d\alpha$. The envelope theorem implies

$$\begin{aligned} \frac{dV_{\alpha}^{ij}}{d\alpha} = & \sum_{k \in \bar{K}_e \cap \bar{K}_i} \theta_k^j \rho_k^{ij} (\psi_k (\gamma_k - v_{0k}) + \lambda^{ij} (\gamma_k - \underline{c} - (1 - \psi_k) (\gamma_k - \bar{c}))) \\ & + \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k^j (1 - \psi_k) \left((1 - \rho_k^{ij}) (v_{1k} - \gamma_k) + \rho_k^{ij} \lambda^{ij} (\gamma_k - \underline{c}) \right) \\ & + \sum_{k \in \bar{K}_n \cap \bar{K}_i} \theta_k^j \left((1 - \psi_k) (1 - \rho_k^{ij}) (v_{1k} - \gamma_k) + \rho_k^{ij} (\gamma_k - v_{0k}) + \rho_k^{ij} \lambda^{ij} (\gamma_k - \underline{c}) \right) \\ & - \lambda^{ij} (\bar{c} - \underline{c}) \\ = & \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k^j (1 - \psi_k) \left((1 - \rho_k^{ij}) (v_{1k} - \gamma_k - \lambda^{ij} (\gamma_k - \underline{c})) - \lambda^{ij} (\bar{c} - \gamma_k) \right) \\ & - (1 - \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k^j (1 - \psi_k)) \lambda^{ij} (\bar{c} - \underline{c}) \end{aligned} \quad (A.12)$$

The first equality follows from the following two observations. First, $\bar{K}_j \supseteq K_j$ for $j = e, n$ because $\bar{c} > \underline{c}$. Second, treatments $k \in \bar{K}_j \cap K_j$ drop out in the expression for $dV_{\alpha}^{ij}/d\alpha$. For the second equality, we use $\underline{K}_i = \emptyset$.

The remaining summation is over treatments k that are used efficiently at $\underline{c}$ and can be financed without insurance at $\bar{c}$. If such a treatment is not insured at $\underline{c}$ (i.e. $\rho_k^{ij} = 0$), then the increase in c makes the treatment available which leads to utility gain $v_{1k} - \gamma_k$ when used. If the treatment was already covered by insurance (i.e. $\rho_k^{ij} = 1$), it is used both at $\underline{c}$ and at $\bar{c}$. The only effect of increasing α is then to save money on the insurance premium $\theta_k^j (1 - \psi_k) (\gamma_k - \underline{c})$.

This positive effect of $d\alpha > 0$ needs to be weighed against the increase in c_α as α increases. The cost of this increase $dc_\alpha/d\alpha = \bar{c} - \underline{c}$ is valued at the marginal utility of income λ^{ij} . Higher λ^{ij} makes an increase in α less attractive. However, in order to compare $dV_{\alpha}^{ij}/d\alpha$ across types, we need to take the different values of ρ_k^{ij} into account as well.

First, we compare $dV_{\alpha}^{hh}/d\alpha$ with $dV_{\alpha}^{lh}/d\alpha$. Using (A.12), we write

$$\begin{aligned} \frac{d(V^{hh} - V^{lh})}{d\alpha} = & \left(1 - \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k^h (1 - \psi_k) \right) (\lambda^{lh} - \lambda^{hh}) (\bar{c} - \underline{c}) \\ & + \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k^h (1 - \psi_k) \left[(1 - \rho_k^{hh}) (v_{1k} - \gamma_k - \lambda^{hh} (\gamma_k - \underline{c})) - \lambda^{hh} (\bar{c} - \gamma_k) \right] \\ & - \left[(1 - \rho_k^{lh}) (v_{1k} - \gamma_k - \lambda^{lh} (\gamma_k - \underline{c})) - \lambda^{lh} (\bar{c} - \gamma_k) \right] \geq 0 \end{aligned}$$

where the inequality follows from the following considerations. First, it follows from Lemma 1 that $\lambda^{lh} - \lambda^{hh} \geq 0$. As $k \in \bar{K}_n$ implies $\bar{c} > \gamma_k$, a sufficient condition for the inequality to hold is

$$(1 - \rho_k^{hh}) (v_{1k} - \gamma_k - \lambda^{hh} (\gamma_k - \underline{c})) \geq (1 - \rho_k^{lh}) (v_{1k} - \gamma_k - \lambda^{lh} (\gamma_k - \underline{c}))$$

As both θ^h types use the same ordering (11) of treatments and the β^h type has more money to spend, we find that $\rho_k^{hh} \geq \rho_k^{lh}$. Hence, we need to consider three cases. First, $\rho_k^{hh} = \rho_k^{lh} = 1$; then the inequality is satisfied. Second, $\rho_k^{hh} = 1 > \rho_k^{lh} = 0$, then the right hand side is negative and the inequality is satisfied. Third, $\rho_k^{hh} = \rho_k^{lh} = 0$, then the inequality follows from $\lambda^{lh} \geq \lambda^{hh}$.

Next consider the comparison within income class β^i of θ^l and θ^h . We want to show that θ^l has a stronger preference for higher c than θ^h . We do this in two steps. First, we assume that no IC is binding. Then we consider the case where $v^{il} > 0$. With $v^{il} = 0$ we have

$$\begin{aligned} \frac{d(V^{il} - V^{ih})}{d\alpha} &= \left(1 - \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k(1 - \psi_k)\right) (\lambda^{ih} - \lambda^{il})(\bar{c} - \underline{c}) \\ &+ \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k(1 - \psi_k) [(1 - \rho_k^{il})(v_{1k} - \gamma_k - \lambda^{il}(\gamma_k - \underline{c}))\lambda^{il}(\bar{c} - \gamma_k) \\ &- ((1 - \rho_k^{ih})(v_{1k} - \gamma_k - \lambda^{ih}(\gamma_k - \underline{c})) - \lambda^{ih}(\bar{c} - \gamma_k))] \geq 0 \end{aligned}$$

where we have used the assumption that $\theta_k^h = \theta_k^l = \theta_k$ for each $k \in \bar{K}_n \cap \underline{K}_e$. Using Lemma 1, we have $\lambda^{ih} \geq \lambda^{il}$. As above, a sufficient condition for the inequality to hold is

$$(1 - \rho_k^{il})(v_{1k} - \gamma_k - \lambda^{il}(\gamma_k - \underline{c})) \geq (1 - \rho_k^{ih})(v_{1k} - \gamma_k - \lambda^{ih}(\gamma_k - \underline{c}))$$

Note that the right hand side of this inequality is non-positive. Hence, with $v_{1k} - \gamma_k - \lambda^{il}(\gamma_k - \underline{c}) \geq 0$ the inequality is satisfied. Consider $v_{1k} - \gamma_k - \lambda^{il}(\gamma_k - \underline{c}) < 0$ (which implies $\rho_k^{il} = 0$). Then $\lambda^{ih} \geq \lambda^{il}$ implies that $\rho_k^{ih} = 0$ and

$$v_{1k} - \gamma_k - \lambda^{ih}(\gamma_k - \underline{c}) < v_{1k} - \gamma_k - \lambda^{il}(\gamma_k - \underline{c})$$

and again the inequality above is satisfied.

Now we turn to the case where $v^{il} > 0$. This introduces the following additional term:

$$\begin{aligned} &+ v^{il} \left(\beta^i + \alpha \left(\sum_{k \in \bar{K}_e} \theta_k^h(1 - \psi_k)\rho_k(v_{1k} - \gamma_k) + \sum_{k \in \bar{K}_i} \theta_k^h\rho_k(v_k - \gamma_k) + \sum_{k \in \bar{K}_n} \theta_k^h(1 - \psi_k)(v_{1k} - \gamma_k) \right) \right. \\ &+ (1 - \alpha) \left(\sum_{k \in \bar{K}_e} \theta_k^l(1 - \psi_k)\rho_k(v_{1k} - \gamma_k) + \sum_{k \in \bar{K}_i} \theta_k^l\rho_k(v_k - \gamma_k) + \sum_{k \in \bar{K}_n} \theta_k^l(1 - \psi_k)(v_{1k} - \gamma_k) \right) \\ &\left. - \hat{V}^{ih} \right) \end{aligned} \quad (\text{A.13})$$

where the value for ih of mimicking il is given by

$$\begin{aligned} \hat{V}_\alpha^{ih} &= \beta^i + \alpha \left(\sum_{k \in \bar{K}_e} (1 - \psi_k)\rho_k^{il}(\theta_k^h v_{1k} - \theta_k^l \gamma_k - \Delta\theta_k \bar{c}) \right. \\ &+ \sum_{k \in \bar{K}_i} \rho_k^{il}(\theta_k^h v_k - \theta_k^l \gamma_k - \Delta\theta_k \bar{c}) + \sum_{k \in \bar{K}_n} (1 - \psi_k)\theta_k^h(v_{1k} - \gamma_k) \\ &+ (1 - \alpha) \left(\sum_{k \in \bar{K}_e} (1 - \psi_k)\rho_k^{il}(\theta_k^h v_{1k} - \theta_k^l \gamma_k - \Delta\theta_k \underline{c}) \right. \\ &\left. + \sum_{k \in \bar{K}_i} \rho_k^{il}(\theta_k^h v_k - \theta_k^l \gamma_k - \Delta\theta_k \underline{c}) + \sum_{k \in \bar{K}_n} (1 - \psi_k)\theta_k^h(v_{1k} - \gamma_k) \right) \end{aligned} \quad (\text{A.14})$$

Hence we find that

$$\frac{d\hat{V}_\alpha^{ih}}{d\alpha} = \sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k(1 - \psi_k)(1 - \rho_k^{il})(v_{1k} - \gamma_k)$$

Hence, differentiating equation (A.13) with respect to α gives

$$v^{il} \left(\sum_{k \in \bar{K}_n \cap \underline{K}_e} \theta_k(1 - \psi_k)(\rho_k^{il} - \rho_k^{ih})(v_{1k} - \gamma_k) \right) \geq 0$$

where the inequality follows from the following considerations. First, if $\rho_k^{ih} = 0$, the inequality holds. Second, consider $\rho_k^{ih} = 1$. This implies that treatment k ends up high in the ranking (10). But because $\theta_k^h = \theta_k^l$ for the treatments in this summation, the ranking in (12) is high as well and θ^l has more money to spent as insurance is cheaper for the better risk. Thus we see that $\rho_k^{ih} = 1$ implies that $\rho_k^{il} = 1$ and the inequality is satisfied. Intuitively, as α increases, treatment k is more likely to end up in $\bar{K}_n$ instead of $\underline{K}_e$; this makes mimicking less attractive. Q.E.D.

Appendix B. Using the envelop theorem

When we consider the effects of the government's choice γ_k in Section 5, we focus on the direct effects; ignoring the effects of γ_k on ρ_k^{ij} , c^{ij} , C^{ij} . This is due to the envelop theorem. However, it may not be obvious that the envelop theorem can be applied in this context. The following proposition derives an envelop theorem result in the context of IC constraints. We first introduce some notation.

Let $u^i(c^i, \gamma)$ denote type $i \in \{l, h\}$'s utility as a function of i 's choice vector c^i and parameter vector γ set by the government. The vector c^i now incorporates the variables ρ_k^{ij} , c^{ij} , C^{ij} . We have $c^i \in \mathbf{C}$ where the set $\mathbf{C}$ includes constraints like $\rho_k \in [0, 1]$. The function u^i can be a Lagrangian with budget constraint and multiplier λ^i . If type j mimics i , j 's utility is written as $\hat{u}^j(c^i, \gamma)$. Finally, v^i denotes the Lagrange multiplier on j 's IC constraint not to mimic i .

In our analysis, the following assumption is satisfied.

Assumption 1. Type h may want to mimic l , but l does not mimic h .

In our model this is correct within an income class. Types do not mimic across income class. Then we have

$$U^h(\gamma) = \max_{c^h \in \mathbf{C}} u^h(c^h, \gamma) \quad (\text{B.1})(\text{B.2})$$

Now we have the following envelop result.

Proposition 4. With assumption 1 we find that

$$\frac{dU^h(\gamma)}{d\gamma} = \frac{\partial u^h(c^h, \gamma)}{\partial \gamma} \quad (\text{B.3})(\text{B.4})$$

Appendix C. Lemmas 2 and 3

Proof of lemma 2 Suppose –by contradiction– that θ^h and θ^l with the same income β are pooled into one contract that covers $(\rho_k > 0)$ at least one treatment k with $\theta_k^h > \theta_k^l$. Then utility of type θ^l can be written as

$$\begin{aligned} u^l = \beta &+ \sum_{k \in K_e^{ij}} (1 - \psi_k) \rho_k (\theta_k^l (v_{1k} - C^{ij}) - \theta_k (\gamma_k - C^{ij})) \\ &+ \sum_{k \in K_i^{ij}} \rho_k (\theta_k^l (\psi_k v_{1k} + (1 - \psi_k) v_{0k} - C^{ij}) - \theta_k (\gamma_k - C^{ij})) \\ &+ \sum_{k \in K_n^{ij}} (1 - \psi_k) \theta_k^l (v_{1k} - \gamma_k) \end{aligned} \quad (\text{C.1})$$

where $\theta_k = \phi \theta_k^l + (1 - \phi) \theta_k^h$ is the pooled probability that someone (in this pool) needs treatment k and $C^{ij} = c^{ij}$.

We claim that there exists a vector $\hat{\rho}$ with $\hat{\rho}_k \leq \rho_k$ for each k such that $u^h = \hat{u}^h$ where

$$\begin{aligned} \hat{u}^h = \beta &+ \sum_{k \in K_e^{ij}} (1 - \psi_k) \hat{\rho}_k (\theta_k^h (v_{1k} - C^{ij}) - \theta_k^l (\gamma_k - C^{ij})) \\ &+ \sum_{k \in K_i^{ij}} \hat{\rho}_k (\theta_k^h (\psi_k v_{1k} + (1 - \psi_k) v_{0k} - C^{ij}) - \theta_k^l (\gamma_k - C^{ij})) \\ &+ \sum_{k \in K_n^{ij}} (1 - \psi_k) \theta_k^h (v_{1k} - \gamma_k) \end{aligned} \quad (\text{C.2})$$

This follows from the mean value theorem. At $\hat{\rho}_k = \rho_k$ for each k we have that $\hat{u}^h \geq u^h$. Further, $\hat{\rho}_k = 0$ for each k leads to $\hat{u}^h \leq u^h$. By continuity there exists $\hat{\rho}$ such that $\hat{u}^h = u^h$. Further, we can choose $\hat{\rho}_k < \rho_k$ for at least one treatment k where $\theta_k^h > \theta_k^l$ due to the assumption that there is at least one treatment with $\rho_k > 0$ and $\theta_k^h > \theta_k^l$.

In words, there exists a deviating contract for θ^l such that θ^l is indifferent between the original pooling contract and this deviation contract. That is, if this deviating contract is offered and priced at the θ^l probabilities, type θ^h will not switch to it. The following shows that $\hat{u}^l > u^l$: θ^l strictly prefers the deviation contract.

We can write

$$\begin{aligned} \hat{u}^j - u^j &= \sum_{k \in K_e^{ij}} (1 - \psi_k) [(\hat{\rho}_k - \rho_k) (\theta_k^j (v_{1k} - C^{ij}) - (\hat{\rho}_k \theta_k^l - \rho_k \theta_k) (\gamma_k - C^{ij}))] \\ &+ \sum_{k \in K_i^{ij}} [(\hat{\rho}_k - \rho_k) (\theta_k^j (v_k - C^{ij}) - (\hat{\rho}_k \theta_k^l - \rho_k \theta_k) (\gamma_k - C^{ij}))] \end{aligned} \quad (\text{C.3})$$

where treatments $k \in K_n^{ij}$ drop out. It follows that

$$(\hat{u}^h - u^h) - (\hat{u}^l - u^l) = \sum_{k \in K_e^{ij}} (1 - \psi_k) (\hat{\rho}_k - \rho_k) (\theta_k^h - \theta_k^l) (v_{1k} - C^{ij}) + \sum_{k \in K_i^{ij}} (\hat{\rho}_k - \rho_k) (\theta_k^h - \theta_k^l) (v_k - C^{ij}) \quad (\text{C.4})$$

Table C.1
12 possible IC constraints

mimics →	ll	lh	hl	hh
ll	X	X	too expensive	too expensive
lh	$IC_{lh,ll}$	X	too expensive	too expensive
hl	no problem	X	X	X
hh	X	no problem	$IC_{hh,hl}$	X

Because $\hat{\rho}_k \leq \rho_k$ and $\theta_k^h \geq \theta_k^l$ with at least one treatment where it is the case that $\hat{\rho}_k < \rho_k$ and $\theta_k^h > \theta_k^l$, we have that

$$\hat{u}^l - u^l > \hat{u}^h - u^h = 0 \quad (C.5)$$

In words, θ^l strictly prefers contract $\hat{\rho}$ to contract ρ . This implies that there exists a deviating contract that is strictly preferred by θ^l while θ^h prefers the pooling contract. Because $\hat{u}^l > u^l$ the deviating contract can be offered in a profitable way. In equilibrium, types θ^h and θ^l cannot be pooled on a contract with $\rho_k > 0$ for at least one treatment with $\theta_k^h > \theta_k^l$.

Next consider the IC constraints. Table C.1 shows all potentially relevant constraints. We do not need to consider the diagonal elements where a type mimics itself. We argue that of the 12 remaining constraints in the table, at max. two are binding.

First, note that a high income type who mimics a low income type with the same θ is actually not a problem in the sense that it does not affect the profitability of the contract (contract is based on the correct probabilities θ).

Second, due to (5), each type spends her whole budget on insurance premium and co-payment C^{ij} . Hence, contracts can be chosen such that low income types cannot mimic high income types: they cannot afford the high income contracts. These are the “too expensive” entries in the table.

We just showed that θ^l and θ^h cannot be pooled into one contract. The same proof shows that θ^l does not want to mimic θ^h . In particular, for each contract for θ^h it is straightforward to find a contract for θ^l which θ^l strictly prefers while θ^h does not.

Finally, type hh does not want to mimic ll ; this can be seen as follows:

$$u^{hh} \geq \hat{u}^{lh} \geq u^{lh} \geq \hat{u}^{ll}$$

where the first inequality follows from the fact that θ^h with high income does not want to mimic θ^h with low income (as there are treatments to insure which the low income type cannot afford); the second inequality follows from the fact that the high income type has higher income (even if hh mimics lh which gives hh utility $\hat{u}^{lh}$, higher income leads to higher consumption and hence higher utility than u^{lh}); the last inequality follows from $IC_{lh,ll}$ and the fact that hh and lh have the same θ^h and hence the same preferences.

Hence we only need to take into account two IC constraints: $IC_{lh,ll}$ and $IC_{hh,hl}$: the θ^h type wants to mimic the θ^l type with the same income. *Q.E.D.*

The following result extends Lemma 1 to the ranking of λ^{il} , v^{il} .

Lemma 3. Consider type θ^l with income β^l . Denote the two marginal treatments (last treatments to be covered) k_1, k_2 with shadow prices λ^{ll} , v^{ll} . For θ^l with income β^h we denote the two marginal treatments k_3, k_4 where we order k_3, k_4 such that $v_{k_3}/\gamma_{k_3} \geq v_{k_4}/\gamma_{k_4}$. Then a necessary and sufficient condition for $\lambda^{hl} \leq \lambda^{ll}$ is

$$\frac{\theta_{k_4}^l (v_{1k_4} - \gamma_{k_4})}{\theta_{k_4}^h (v_{1k_4} - c) - \theta_{k_4}^l (\gamma_{k_4} - c)} - \frac{\theta_{k_3}^l (v_{1k_3} - \gamma_{k_3})}{\theta_{k_3}^h (v_{1k_3} - c) - \theta_{k_3}^l (\gamma_{k_3} - c)} \leq \left(\frac{1}{\frac{\theta_{k_4}^h}{\theta_{k_4}^l} \frac{v_{1k_4} - c}{\gamma_{k_4} - c} - 1} - \frac{1}{\frac{\theta_{k_3}^h}{\theta_{k_3}^l} \frac{v_{1k_3} - c}{\gamma_{k_3} - c} - 1} \right) \lambda^{ll} \quad (C.6)$$

A necessary and sufficient condition for $v^{hl} \leq v^{ll}$ is

$$\left(\frac{v_{1k_3} - c}{\gamma_{k_3} - c} \frac{\theta_{k_3}^h}{\theta_{k_3}^l} - \frac{v_{1k_4} - c}{\gamma_{k_4} - c} \frac{\theta_{k_4}^h}{\theta_{k_4}^l} \right) v^{ll} \geq \frac{v_{1k_3} - \gamma_{k_3}}{\gamma_{k_3} - c} - \frac{v_{1k_4} - \gamma_{k_4}}{\gamma_{k_4} - c} \quad (C.7)$$

Proof of lemma 3 Comparing two θ^l types (one with high, the other with low budget β^i) is a bit more involved than comparing two θ^h types because of the interaction of the budget and IC constraints. To see what happens with shadow prices v, λ as budget β^i increases, we derive the dual problem of an agent's optimization problem. By assumption $C^{ij} = c$ for all types and hence the set K_n is the same for all types. It follows from Proposition 1 that either $\gamma = c$ or $\gamma_k = \delta_k$. For treatments with $\delta_k > c$, assumption (1) implies that $\rho_k^{ij} = 0$ for each $k \in K_i$; where $C^{ij} = c$ implies that the sets in (3) can be written without superscript ij .

Eq. (6) can now be written as the a linear programming problem:

$$\max_{\rho_k \geq 0} \sum_{k \in K_e} \rho_k \theta_k^l (1 - \psi_k) (v_{1k} - c) \quad (C.8)$$

subject to

$$\rho_k \leq 1 \text{ foreach } k \in K_e \quad (C.9)(BC)(IC)$$

where we made the following simplifications to ease notation. The right hand side of equation (BC) should be $\beta^i - c$. But because $C^{ij} = c$ for all types, we dropped the c . Similarly, the objective function equals c plus the expression in (C.8), but we drop c . As all types have the same set K_n , we dropped the term $\sum_{k \in K_n} (1 - \psi_k) \theta_k^h (v_{1k} - \gamma_k)$ in equation (IC).

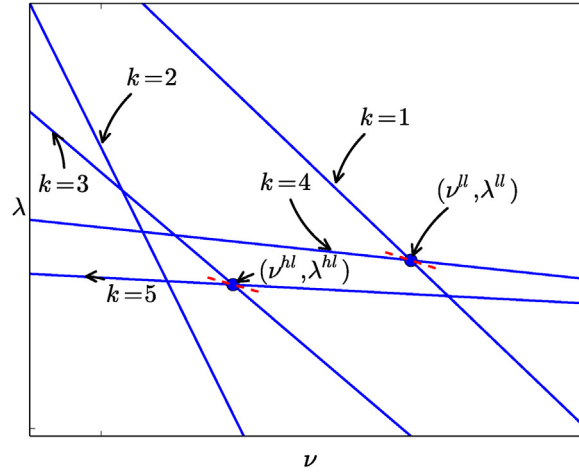


Fig. 2. Dual problem of Example 2.

The dual of problem (C.8) is given by (see, for instance, Brinkhuis and Tikhomirov, 2005)

$$\min_{\zeta_k, \lambda, v \geq 0} \sum_{k \in K_e} \zeta_k + \beta^i \lambda + (V^{ih} - \beta^i) v \quad (C.10)$$

subject to

$$\zeta_k + \theta_k^l (1 - \psi_k)(\gamma_k - c) \lambda + (1 - \psi_k)(\theta_k^h (v_{1k} - c) - \theta_k^l (\gamma_k - c)) v \geq \theta_k^l (1 - \psi_k)(v_{1k} - c) \quad (C.11)$$

where ζ_k is the shadow price corresponding to the constraint $\rho_k \leq 1$, λ corresponds to (BC) and v to (IC).

For $\beta^i \geq 0$ close to 0, $V^{ih} - \beta^i \geq 0$ is close to zero as well. It is then optimal to set $\zeta_k = 0$ for each k in optimization problem (C.10). Hence, λ and v are chosen such that all inequalities (C.11) hold with $\zeta_k = 0$. If $\lambda, v > 0$, they are determined by 2 treatments k_1, k_2 such that $\zeta_{k_1} = \zeta_{k_2} = 0$ and equation (C.11) holds with equality for k_1 and k_2 .³⁵

As β^i and $V^{ih} - \beta^i$ increase, at some point it becomes optimal to lower λ, v such that $\zeta_{k_1}, \zeta_{k_2} > 0$ and $\lambda, v > 0$ are determined by two treatments k_3, k_4 with $\zeta_{k_3} = \zeta_{k_4} = 0$ and (C.11) holding with equality for these two treatments. In the primal problem this corresponds to β^i high enough that $\rho_{k_1} = \rho_{k_2} = 1$.

This is illustrated in Fig. 2 based on example 2. The solid lines correspond to equations (C.11) holding with equality and $\zeta_k = 0$ for each of the 5 treatments $k \in K$:

$$\lambda = \frac{v_{1k} - c}{\gamma_k - c} - \left(\frac{\theta_k^h}{\theta_k^l} \frac{v_{1k} - c}{\gamma_k - c} - 1 \right) v \quad (C.12)$$

The dashed lines correspond to the slope of the objective function (C.10) in (v, λ) space. For low values of β^i and $V^{ih} - \beta^i$, it is optimal to set $\zeta_k = 0$ for each k . For (v^{ll}, λ^{ll}) all inequalities (C.11) are satisfied and they are binding for treatments $k_1 = 1$ and $k_2 = 4$. At higher values β^i and $V^{ih} - \beta^i$, it is optimal to set $\zeta_1, \zeta_4 > 0$ and the other $\zeta_k = 0$. At (v^{hl}, λ^{hl}) the binding constraints (C.11) correspond to treatments $k_3 = 3$ and $k_4 = 5$.

We can now derive a necessary and sufficient condition for $\lambda^{hl} \leq \lambda^{ll}$. Consider the treatment with the lower intercept $(v_{k_4} - \gamma_{k_4})/(\gamma_{k_4} - c)$ in equation (C.12). The intersection of this line with $\lambda = \lambda^{ll}$ should lie to the left of the intersection of the k_3 -line with λ :

$$\frac{\theta_{k_4}^l (v_{1k_4} - \gamma_{k_4})}{\theta_{k_4}^h (v_{1k_4} - c) - \theta_{k_4}^l (\gamma_{k_4} - c)} - \lambda^{ll} \frac{1}{\frac{\theta_{k_4}^h}{\theta_{k_4}^l} \frac{v_{1k_4} - c}{\gamma_{k_4} - c} - 1} \leq \frac{\theta_{k_3}^l (v_{1k_3} - \gamma_{k_3})}{\theta_{k_3}^h (v_{1k_3} - c) - \theta_{k_3}^l (\gamma_{k_3} - c)} - \lambda^{ll} \frac{1}{\frac{\theta_{k_3}^h}{\theta_{k_3}^l} \frac{v_{1k_3} - c}{\gamma_{k_3} - c} - 1} \quad (C.13)$$

This can be written as Eq. (C.6).

Similarly, we can derive the necessary and sufficient condition for $v^{hl} \leq v^{ll}$. In words, at v^{ll} the line corresponding to k_3 should lie below the line corresponding to k_4 . Then the intersection point of the two lines is to the left of v^{ll} . This can be written as:

$$\frac{v_{1k_3} - \gamma_{k_3}}{\gamma_{k_3} - c} - \left(\frac{\theta_{k_3}^h}{\theta_{k_3}^l} \frac{v_{1k_3} - c}{\gamma_{k_3} - c} - 1 \right) v^{ll} \leq \frac{v_{1k_4} - \gamma_{k_4}}{\gamma_{k_4} - c} - \left(\frac{\theta_{k_4}^h}{\theta_{k_4}^l} \frac{v_{1k_4} - c}{\gamma_{k_4} - c} - 1 \right) v^{ll} \quad (C.14)$$

which is equivalent to Eq. (C.7).Q.E.D.

³⁵ If either $\lambda = 0$ or $v = 0$, the other shadow price is determined by $\zeta_{k_1} = 0$ and (C.11) holding with equality for k_1 . As we are interested in the case where both (BC) and (IC) are binding, we focus on $\lambda, v > 0$. It is straightforward to extend the analysis below for the case where one shadow price equals 0.

References

- Bijlsma, Michiel, Boone, Jan, Zwart, Gijsbert, 2014. Competition leverage: how the demand side affects optimal risk adjustment. *RAND J. Econ.* 45 (4), 792–815.
- Blomqvist, Ake, 2011. Chapter 12 – public-sector health care financing. In: Glied, S., Smith, P. (Eds.), *Oxford Handbook of Health Economics*. Oxford University Press, pp. 257–284.
- Boadway, Robin, Leite-Monteiro, Manuel, Marchand, Maurice, Pestieau, Pierre, 2006. Social insurance and redistribution with moral hazard and adverse selection. *Scand. J. Econ.* 108 (2), 279–298, 1467–9442.
- Boone, Jan, 2015. Basic versus supplementary health insurance: moral hazard and adverse selection. *J. Public Econ.* 128, 50–58.
- Boone, Jan, Schottmüller, Christoph, 2017. Health insurance without single crossing: why healthy people have high coverage. *Econ. J.* 127 (599), 84–105.
- Brinkhuis, Jan, Tikhomirov, Vladimir, 2005. *Optimization: Insights and Applications*. Princeton University Press.
- Chandra, Amitabh, Skinner, Jonathan S., 2012 July. Technology growth and expenditure growth in health care. *J. Econ. Lit.* 50 (3), 645–680.
- Chernew, Michael E., Encinosa, William E., Hirth, Richard A., 2000. Optimal health insurance: the case of observable, severe illness. *J. Health Econ.* 19 (5), 585–609.
- Chernew, Michael E., Rosen, Allison B., Mark Fendrick, A., 2007. Value-based insurance design. *Health Aff.* 26 (2), w195–w203.
- Cohn, J., 2007. *Sick: the Untold Story of America's Health Care Crisis-and the People Who Pay the Price*. Harper Perennial.
- Colombo, F., Tapay, N., 2004. Private health insurance in oecd countries: the benefits and costs for individuals and health systems, Working Paper 15, OECD.
- Conen, David, Glynn, Robert J., Ridker, Julie E., Buring, Paul M., Albert, Michelle A., 2009. Socioeconomic status, blood pressure progression, and incident hypertension in a prospective cohort of female health professionals. *Eur. Heart J.* 30, 1378–1384.
- Cutler, David M., Zeckhauser, Richard J., 2000. Chapter 11 the Anatomy of Health Insurance. Volume 1. Part A of *Handbook of Health Economics*. Elsevier, pp. 563–643.
- Dalstra, J.A.A., Kunst, A.E., Borrell, C., Breeze, E., Cambois, E., Costa, G., Geurts, J.J.M., Lahelma, E., Van Oyen, H., Rasmussen, N.K., Regidor, E., Spadea, T., Mackenbach, J.P., 2005. Socioeconomic differences in the prevalence of common chronic diseases: an overview of eight european countries. *Int. J. Epidemiol.* 34, 316–326.
- DeMarzo, Peter M., Fishman, Michael J., Hagerty, Kathleen M., 2005. Self-regulation and government oversight. *Rev. Econ. Stud.* 72 (3), 687–706.
- Drummond, M., Sculpher, M., Torrance, G., O'Brien, B., Stoddart, G., 2005. *Methods for the Economic Evaluation of Health Care Programmes*, 3rd edition. Oxford University Press.
- Ellis, Randall P., Manning, Willard G., 2007. Optimal health insurance for prevention and treatment. *J. Health Econ.* 26 (6), 1128–1150.
- Faure-Grimaud, Antoine, Laffont, Jean-Jacques, Martimort, David, 1999. The endogenous transaction costs of delegated auditing. *Eur. Econ. Rev.* 43 (4), 1039–1048.
- Finkelstein, A., McGarry, K., 2006. Multiple dimensions of private information: evidence from the long-term care insurance market. *Am. Econ. Rev.* 96 (4), 938–958.
- Frijters, P., Haisken-DeNew, J.P., Shields, M.A., 2005. The causal effect of income on health: Evidence from german reunification. *J. Health Econ.* 24 (5), 997–1017.
- Garber, Alan M., Sculpher, Mark J., 2011. Chapter eight – cost effectiveness and payment policy. In: McGuire, Thomas G., Pauly, Mark V., Barros, Pedro P. (Eds.), *Handbook of Health Economics*, volume 2. Elsevier, pp. 471–497.
- Gold, M.R., Siegel, J.E., Russell, L.B., Weinstein, M.C., 1996. *Cost-Effectiveness in Health and Medicine*. Oxford University Press.
- Gravelle, H., Sutton, M., 2009. Income, relative income, and self-reported health in britain 1979–2000. *Health Econ.* 18 (2), 125–145.
- Hoel, Michael, 2007. What should (public) health insurance cover? *J. Health Econ.* 26 (2), 251–262, ISSN 0167-6296.
- Leyland, A.H., 2005. Socioeconomic gradients in the prevalence of cardiovascular disease in scotland: the roles of composition and context. *J. Epidemiol. Community Health* 59, 799–803.
- McDonough, J.E., 2011. *Inside National Health Reform*. University of California Press.
- Meltzer, David O., Smith, Peter C., 2011. Chapter seven – theoretical issues relevant to the economic evaluation of health technologies. In: McGuire, Thomas G., Pauly, Mark V., Barros, Pedro P. (Eds.), *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*. Elsevier, pp. 433–469.
- E. Mossialos and S. Thomson. Voluntary health insurance in the european union. Working paper, World Health Organization, 2004.
- Munkin, M.K., Trivedi, P.K., 2010. Disentangling incentives effects of insurance coverage from adverse selection in the case of drug expenditure: a finite mixture approach. *Health Econ.* 19 (9), 1093–1108, ISSN 1099-1050.
- Netzer, Nick, Scheuer, Florian, 2010. Competitive screening in insurance markets with endogenous wealth heterogeneity. *Econ. Theory* 44 (2), 187–211.
- Nyman, J.A., 2003. *The Theory of Demand for Health Insurance*. Stanford University Press.
- Nyman, John A., 1999. The value of health insurance: the access motive. *J. Health Econ.* 18 (2), 141–152.
- Pauly, M.V., 2008. Adverse Selection and Moral Hazard: Implications for Health Insurance Markets. Incentives and Choice in Health Care, Chapter 5. MIT Press, pp. 103–129.
- Pauly, M.V., 1968. The economics of moral hazard: Comment. *Am. Econ. Rev.* 58 (3), 531–537.
- 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.
- Schoen, C., Osborn, R., Squires, D., Doty, M.M., Pierson, R., Applebaum, S., 2010. How Health Insurance Design Affects Access To Care And Costs, By Income, In Eleven Countries. *Health Aff.* 29 (12), 1–12.
- Schoen, Cathy, Collins, Sara R., Kriss, Jennifer L., Doty, Michelle M., 2008. How many are underinsured? trends among u.s. adults, 2003 and 2007. *Health Aff. (Project Hope)* 27 (4), 298–309.
- Schokkaert, Erik, van de Voorde, Carine, 2011. Chapter 15 – user charges. In: Glied, S., Smith, P. (Eds.), *Oxford Handbook of Health Economics*. Oxford University Press, pp. 329–353.
- Schottmüller, Christoph, 2013 July. Cost incentives for doctors: A double-edged sword. *Eur. Econ. Rev.* 61, 43–58.
- Smart, M., 2000. Competitive insurance markets with two unobservables. *Int. Econ. Rev.* 41 (1), 153–170.
- Smith, Peter C., 2005. User charges and priority setting in health care: balancing equity and efficiency. *J. Health Econ.* 24 (5), 1018–1029.
- Smith, Peter C., 2007. Chapter 5 – provision of a public benefit package alongside private voluntary health insurance. In: Preker, A.S., Scheffler, R.M., Bassett, M.C. (Eds.), *Private voluntary health insurance in development: friend or foe?* The World Bank, pp. 147–167.
- Stabile, Mark, Thomson, Sarah, 2014. The changing role of government in financing health care: An international perspective. *J. Econ. Lit.* 52 (2), 480–518.
- Thomson, S., Osborn, R., Squires, D., Jun, M., 2012. International Profiles of Health Care Systems. Commonwealth Fund.
- van de Ven, Wynand P.M.M., Schut, Frederik T., 2011. Chapter 17 – guaranteed access to affordable coverage in individual health insurance markets. In: Glied, S., Smith, P. (Eds.), *Oxford Handbook of Health Economics*. Oxford University Press, pp. 380–404.
- Van Loon, A. Jeanne M., Alexandra Goldbohm, R., Van den Brandt, Piet A., 1998. Socioeconomic status and stomach cancer incidence in men: results from the netherlands cohort study. *J. Epidemiol. Community Health* 52, 166–171.
- Wakker, P., Thaler, R., Tversky, A., 1997. Probabilistic insurance. *J. Risk Uncertain.* 15, 7–28.
- Wambach, Achim, 2000. Introducing Heterogeneity in the Rothschild-Stiglitz Model. *J. Risk Insurance* 67 (4), 579–591.
- Zeckhauser, Richard J., 1970. Medical insurance: A case study of the tradeoff between risk spreading and appropriate incentives. *J. Econ. Theory* 2 (1), 10–26.
- Zweifel, Peter, 2011. Chapter 13 – voluntary private health insurance. In: Glied, S., Smith, P. (Eds.), *Oxford Handbook of Health Economics*. Oxford University Press, pp. 285–307.