DISCUSSION PAPER SERIES

No. 9420

DOES THE MARKET CHOOSE OPTIMAL HEALTH INSURANCE COVERAGE?

Jan Boone

INDUSTRIAL ORGANIZATION and PUBLIC POLICY

Centre for Economic Policy Research

www.cepr.org

Available online at: **www.cepr.org/pubs/dps/DP9420.asp**

ISSN 0265-8003

DOES THE MARKET CHOOSE OPTIMAL HEALTH INSURANCE COVERAGE?

Jan Boone, Tilburg University, CentER, TILEC and CEPR

Discussion Paper No. 9420
April 2013

Centre for Economic Policy Research
77 Bastwick Street, London EC1V 3PZ, UK
Tel: (44 20) 7183 8801, Fax: (44 20) 7183 8820
Email: cepr@cepr.org, Website: www.cepr.org

This Discussion Paper is issued under the auspices of the Centre's research programme in **INDUSTRIAL ORGANIZATION and PUBLIC POLICY**. Any opinions expressed here are those of the author(s) and not those of the Centre for Economic Policy Research. Research disseminated by CEPR may include views on policy, but the Centre itself takes no institutional policy positions.

The Centre for Economic Policy Research was established in 1983 as an educational charity, to promote independent analysis and public discussion of open economies and the relations among them. It is pluralist and non-partisan, bringing economic research to bear on the analysis of medium- and long-run policy questions.

These Discussion Papers often represent preliminary or incomplete work, circulated to encourage discussion and comment. Citation and use of such a paper should take account of its provisional character.

CEPR Discussion Paper No. 9420

April 2013

ABSTRACT

Does the market choose optimal health insurance coverage?*

Consumers, when buying health insurance, do not know the exact value of each treatment that they buy coverage for. This leads them to overvalue some treatments and undervalue others. We show that the insurance market cannot correct these mistakes. This causes research labs to overinvest in treatments that hardly add value compared to current best practice. The government can stimulate R&D in breakthrough treatments by excluding treatments with low value added from health insurance coverage. If the country is rich enough such a government intervention in a private health insurance market raises welfare.

JEL Classification: D4, I13 and I18
Keywords: cost effectiveness analysis, health insurance and pharmaceutical research and development

Jan Boone
Department of Economics
Tilburg University
Warandelaan 2
PO Box 90153
5000 LE Tilburg
THE NETHERLANDS

Email: j.boone@uvt.nl

*I am grateful for comments and suggestions by seminar participants in Barcelona, the Dutch Competition Authority (NMa), Dutch Healthcare Authority (NZa), Groningen, Leuven, Paris, Tilburg and Vienna. Financial support from the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged.

Submitted 26 March 2013

1. Introduction

Physicians worry that pharmaceutical companies introduce more and more drugs that hardly add value (compared to existing treatments); yet these new drugs are sold with substantial profits (Light and Lexchin, 2012; Deyo and Patrick, 2005). Economic theory predicts that – absent collusion– new products which are only marginally better than existing ones only make small profits. How can drugs with low value added lead to high profits? What are the effects of this on research labs' incentives to invest in R&D and dynamic efficiency? Can this explain the innovation crisis where "pharmaceutical research and development turns out mostly minor variations on existing drugs, and most new drugs are not superior on clinical measures" (Light and Lexchin, 2012).

The reason why quality and costs of new medical treatments are important is that health care costs are increasing rapidly over time in most, if not all, developed countries. We know that an important contributor to the increase in health care expenditure is the introduction of new medical technologies (Garber, 2001; Deyo and Patrick, 2005; Chernew and Newhouse, 2011; Chernew and May, 2011). To the extent that these new technologies create more utility than they cost, this form of technological progress is welfare enhancing. However, it is not clear that the market only introduces technologies where the benefits exceed the costs.

There is a difference between private and public health insurance systems. In public systems (or in the public part of the health insurance system, like the mandatory insurance in countries like Austria, Italy, Luxembourg, the Netherlands; see Jakubowski and Busse (1998) and Thomson et al. (2012)) there is a central board that decides which (new) treatments are covered by health insurance and which are not. A well known institute performing cost effectiveness (CE) analysis to determine whether a treatment should be covered by insurance is NICE in the UK. Other countries where centrally organized CE analysis plays a role include Australia, Canada, Ireland, the Netherlands and Sweden (see Dranove (2003), Drummond (2007), Garber and Sculpher (2011, pp. 487), Morton and Kyle (2011, pp. 806) and Meltzer and Smith (2011, pp. 436)).

In private health insurance markets (like the private part of the current US market), there is no central board deciding which treatments are covered by health insurance.[1] In the US, currently the (federal) Food and Drug Administration (FDA) checks whether a new treatment is safe and whether it is effective (does better than a placebo treatment) but it does not check whether it is better than existing treatments[2] nor whether it is cost effective (Deyo and Patrick (2005, chapter 12) and Danzon (2011, pp. 523)). Hence, the decision on what treatments to cover by private health insurance is left to the market.

Although private insurers can do CE analysis themselves to determine which treatments to cover, some people are concerned about leaving coverage decisions to the market. In particular, there is a worry that the market tends to over-reward medical innovations that hardly add value compared to existing treatments (Deyo and Patrick, 2005). For instance, consumers

[1]Government funded health insurance in the US can have a central board deciding which treatments are covered. See, for instance, Medicaid in Oregon (Dranove, 2003; Oberlander, 2007). With the Patient Protection and Affordable Care Act ("Obamacare") the government has to choose which treatments belong to the "essential health benefits package" (McDonough, 2011, pp. 113).

[2]This is different in Europe where the European Medicines Evaluation Agency does compare with existing treatments (Danzon, 2011, pp. 524).

don't know the value of medical treatments and may be too optimistic about the effect of a particular new treatment. Others believe that the market can correct misperceptions of uninformed consumers (Danzon and Pauly, 2001). To illustrate, an often heard argument in favor of the use of formularies by health insurers is that insurers are better at judging the value of new treatments and drugs than consumers are (Morton and Kyle, 2011, pp. 791). Hence, insurers –in the interest of their customers– exclude drugs that are not worth their costs.[3]

The question that we analyze is the following. Does the market decide efficiently (in terms of total welfare) whether a new treatment is valuable enough that it should be covered by health insurance? Or is there a role for the government in deciding which treatments can/should be covered by private health insurance (see, for instance, CBO, 2007)? In order to answer this question, we (i) analyze static and dynamic efficiency of a private health insurance market and (ii) derive conditions under which efficiency can be enhanced by introducing a government agency doing CE analysis to determine which medical innovations are allowed to be covered.[4]

The model introduces research labs investing in R&D to invent new treatments. Once a treatment is invented, labs bargain with insurers to determine the price insurers have to pay the lab if an insured patient needs the treatment (say, a new drug). Clearly, these prices are affected by the fall back option of labs in case their treatment is not covered by insurance: selling to patients on the uninsured market. The uninsured market is explicitly modeled. Finally, insurers compete on the health insurance market and consumers can decide to buy or forgo insurance.

Although it is impossible (and undesirable) to capture the richness of the real world in a model, we do capture some important aspects of health insurance markets (below we give references to motivate these assumptions). First, consumers –when buying insurance– do not necessarily know the value of each treatment that is covered by their insurance plan. The number of treatments can be huge and they may help to cure diseases that the consumer does not even know off (Liebman and Zeckhauser, 2008). Hence, a consumer can overestimate the value of her insurance covering a certain treatment. Although consumers may be wrong about the value of a particular treatment, there is little evidence that markets consistently over-value medical innovations. The model is consistent with both these observations. Second, the policy concern for covering a treatment in health insurance is not so much the disutility of consumption risk (risk aversion) but access to care. If a treatment is not covered by insurance, some (poor) people may not be able to afford the treatment when they need it. Third, when a patient needs a certain treatment, a physician helps to guide the patient's decision process. In other words, a patient needing a specific treatment is better informed about the value of this treatment than the consumer buying insurance. Finally, the offers between labs and insurers are private (in the sense of not publicly known); we consider both a bidding and an offer game (interestingly, these lead to similar equilibrium payoffs).

We find the following results. Because consumers do not necessarily know the value of each treatment, firms tend to underinvest in break-through treatments. This reduces dynamic efficiency (compared to first best). Also, if the country is "fairly rich" (to be made precise below), it may be more profitable for a lab with a break-through innovation to sell on the uninsured market. This reduces static efficiency as some poor people will not have access to

[3]Although the argument usually stops here, these (potentially misinformed) consumers still need to value the different formularies offered by insurers. We model this explicitly below.

[4]One reason why the government should do CE analysis is to avoid duplication of the analysis (CBO, 2007; Phelps, 2000). Although relevant, this is not the reason that we analyze here.

this treatment. If the government excludes innovations with low value added from insurance coverage, firms' R&D incentives increase. This happens for two reasons: (i) being covered by health insurance is now a signal about the treatment's value added and (ii) R&D labs have to work harder to qualify for health insurance. Finally, if the country is poor, some people will not be able to afford excluded low value added treatments, reducing static efficiency. Overall, total welfare increases due to government intervention in rich enough countries.

This paper is related to the literature in the following way. First, it is linked to the health (economics) literature on CE analysis which is discussed in the next section. Second, a well known result by Gaynor et al. (2000) says that static efficiency can be achieved by competitive health care and insurance markets. Garber et al. (2006) consider whether this result extents to dynamic efficiency. They show that when the coinsurance rate is set such that treatment use is efficient (static efficiency), the incentive for inventing the treatment is not. Depending on the distribution of treatment benefits, the incentive to invent treatments may well be excessive. Welfare can be increased by regulating treatment prices and limiting the duration of a patent. Lakdawalla and Sood (2009) argue in favour of public health insurance to decouple consumer prices from manufacturer prices in a way that balances static and dynamic efficiency. Jena and Philipson (2008) consider the effects of CE analysis on innovation. They argue that setting a CE threshold in a public system is equivalent to regulating treatment prices. Indeed, a lab that knows what the value added of its treatment is, sets its price so high as to just satisfy the CE threshold. Then by setting the threshold, government can balance static and dynamic efficiency effects.

The difference with our paper is the following. First, we consider a market setting without government intervention in the form of price regulation. We analyse whether there is a role for the government in determining which treatments are allowed to be covered by private health insurance. Parties then determine prices in a bilateral bargaining process. Second, in our model, consumers do not necessarily know the value of each invented treatment. Third, in the papers mentioned, the innovation is given and firms need to invest a fixed cost to invent the treatment (i.e. the focus is on the extensive margin). We consider the case where firms invest to raise the quality of their innovation (i.e. we focus on the intensive margin). By working harder and testing more, the firm can reduce complications and side effects of the treatment, make the treatment effective for different forms of a disease etc. This process of treatment improvement can be affected by the government by raising the value added threshold that a new treatment must pass before it can be covered by insurance. CE analysis is a natural instrument to affect R&D incentives in health care. Whether a treatment is covered by health insurance has a big effect on profits and hence on R&D incentives (Weisbrod, 1991; Chernew and Newhouse, 2011).

We assume that insurers know the value of the different treatments, while consumers do not (when they buy insurance). This links the paper to the literature on experts (see Dulleck and Kerschbamer, 2006, for an overview). This literature analyzes experts' incentives to provide correct services to consumers who do not know what they need. Questions analyzed include: can the market prevent under- and over-supply of services. There are two main differences with the model below. First, in our model, the consumer is not exploited in equilibrium: the insurance premium does not exceed the value of insurance for the consumer. Second, we focus on the effect of the consumer's lack of information on the bargaining between the insurer and the research lab.

Finally, Liebman and Zeckhauser (2008) stress that rational decision making is not a good

assumption when analyzing trade offs in health insurance markets. They discuss the implications for health care policy of findings in the behavioural economics literature, like status quo bias, anxiety and the way people deal with uncertain events in the future. We also deviate from full rationality, in the sense that consumers do not know the value of each treatment when buying health insurance. However, we focus on the question whether the market can correct consumers' misperceptions. And what the effect of government intervention is on static and dynamic efficiency.

Chandra and Skinner (2012) argue that part of the growth in US health care expenditure is due to the adoption of ineffective (Category II and III, in their terminology) treatments. They introduce a model to explain why ineffective treatments are used once they are covered by insurance. This paper and Skinner (2011) document that a market –left to itself– does indeed cover treatments that are not worth their costs. Their work motivates our question why consumers buy insurance coverage for such treatments. Further, we analyze the incentives to invent/introduce treatments and how government intervention changes these incentives.

This paper is organized as follows. The next section gives a brief overview of CE analysis. Then we introduce the model. Section 4 analyzes the equilibrium in simple contracts without government intervention. Then we consider the effects on static and dynamic efficiency of the government excluding low value added treatments from insurance coverage. Section 6 asks the question whether the market can exclude low quality treatments itself without government intervention. Extensions of the model are discussed and finally we conclude. Proofs of the results can be found in the appendix.

2. CE analysis

Broadly speaking, the goal of CE analysis is to maximize "health" subject to a resource constraint. In most public systems, the resource constraint takes the form of a budget constraint. The idea of CE analysis is to rank treatments in terms of the health gained per unit of the budget spent. Then starting from the top of this ranking, treatments are added to health insurance coverage until the budget is spent. Although this idea is intuitive, applying it in practice is not straightforward (Basu and Philipson, 2010).

CE analysis can be done in a number of ways. To illustrate, one distinction is the way the health benefits of treatments are measured. Although different authors use different terminology,[5] one can distinguish CE, Cost Utility (CU) and Cost Benefit (CB) analysis. All three methods work with the same concept of monetary costs (price of a treatment).[6] CE analysis measures benefits in terms of, say, "years of life gained" without putting a value on this. In other words, 3 years in perfect health is the same as 3 years living with an impairment, like not being able to walk the stairs. CU analysis tends to use QALYs (Quality Adjusted Life Year) to quantify the difference between life years with full health and with impairments. Finally, CB analysis puts a monetary value on a QALY. To illustrate, NICE tends to work with a value of 20,000–30,000 pounds per QALY (NICE, 2007). Studies based on surveys and revealed preference come up with a value between $100,000–300,000; see, for instance, Dranove (2003,

[5]For instance, the term CE analysis is also used to describe cost benefit analysis in a health care context. In this paper, we also use CE analysis in this broader sense.

[6]This in contrast to comparative effectiveness research which does not consider costs (Chandra et al., 2011).

chapter 8) and Chandra and Skinner (2012, pp. 654).

Other issues in CE analysis include how to deal with uncertainty of benefits, how to discount benefits over time, how to measure costs and how to take equity considerations into account. See Drummond et al. (2005), Garber (2000) and Garber and Sculpher (2011) for details. As our model below is rather simple in this respect, these distinctions do not matter for our analysis. There, however, is a difference on how we measure treatment costs and how this is usually done. In a public system with a budget, the relevant cost concept is the price of the treatment. The same is true when a private insurer uses CE analysis to determine which treatments to cover. However, we are interested in total welfare and then a price is a transfer from one party to another. From a total welfare point of view, the relevant concept is the production cost of a treatment. Jena and Philipson (2013); Pauly (2007) argue that the use of treatment prices instead of costs leads to suboptimal outcomes from a total welfare point of view.

As mentioned in the introduction, in many countries the government uses CE analysis to determine which treatments are covered by the public/mandatory part of health insurance. In contrast, US law explicitly prohibits the FDA (and any other federal agency) to do CE analysis (Garber and Sculpher, 2011, pp. 490/1). Hence, one would expect that private insurers then do CE analysis themselves, but they don't. Indeed, "[w]ith rare exceptions, in the United States private health insurers ... avoid explicit use of cost-effectiveness analysis to determine coverage... Safety and effectiveness, but not cost, are considered in formal coverage decision making" (Garber and Sculpher, 2011, pp. 490). Similar statements can be found in Deyo and Patrick (2005, pp. 54) and Meltzer and Smith (2011, pp. 436). As no CE analysis is done, treatments that are not worth their costs are covered by health insurance.

At first sight, this seems related to marginal or flat-of-the-curve medicine (see Dranove, 2000, pp. 36 for a discussion). Because of insurance (and many physicians not being trained to think in terms of benefits and costs) patients keep on trying treatments where the expected benefits of additional treatment do not exceed the social costs. However, this argument is conditional on insurance covering these treatments (in these circumstances). Once I have insurance that covers a certain treatment, I am indeed tempted to use the treatment even though benefits may be low and monetary costs (for society) high. Ex ante, a rational consumer prefers insurance contracts with high quality/price ratios. Hence, one would expect there to be an incentive for insurers to use CE analysis and only cover treatments with substantial value added.

The model shows why insurers do not have an incentive to do CE analysis and what the effects are if the government excludes marginal innovations (with low value added over existing treatments) from insurance coverage. We conclude this section with examples of treatments with low value added that were covered in the US health insurance market.[7] As we show below, once a marginal treatment is covered, it can command a price in excess of its social value.

Deyo and Patrick (2005) discuss a number of US cases "where a new treatment has some very small advantage over alternative treatments, but at a very high price. The number of such

[7]Two remarks on the private health insurance market in the US. First, this market is dominated by employer based group contracts; not individual contracts. However, group contracts are likely to be sensitive to individual preferences, which is what we need in the model. Second, the private market in the US is not helped by Medicare's coverage decisions that are not motivated by CE analysis (Chandra and Skinner, 2012, pp. 27). Once a treatment is covered by Medicare, it is hard for a private insurer to exclude this treatment from its own coverage. This reiterates how hard it is for insurers to communicate cost effectiveness of treatments to consumers.

treatments is enormous" (page 252). In the 1990s, US insurers covered high-dose chemotherapy with autologous bone marrow transplant to treat (advanced) breast cancer. Treatment cost per patient were $80,000 while, in fact, this new treatment had no advantage over cheaper existing chemotherapy (Mello and Brennan, 2001). Kelly et al. (2001) compare two treatments for lung cancer: paclitaxel plus carboplatin (PC) and vinorelbine and cisplatin (VC). Whereas PC was $8,000 more expensive (per patient) than VC, the survival results (and other outcome measures) are basically the same. Nonetheless PC was covered by insurance and extensively used. In the 1990s an expensive and new class of drugs –calcium-channel blockers– were used to treat high blood pressure. However, due to side effects, these were not better than existing cheaper treatments of hypertension (Fischer and Avorn, 2004).

Similar examples can be given for medical devices. Consider the use of pedicle screws to treat back pain. Comparing a spinal fusion operation with and without pedicle screws, Deyo et al. (2004, pp. 724) conclude that the incremental cost effectiveness ratio is more than $3 million per QALY. This is huge compared to the 20.000-30.000 pounds per QALY benchmark used by NICE in the UK or even the $100,000-300,000 value per life year discussed above. Skinner (2011, pp. 47) mentions the use of "very expensive treatment without known benefits for patients (like proton beam therapy for prostate cancer)". Also, it is not clear whether prostate cancer screening in apparently healthy men offers any benefits (Deyo and Patrick 2005, pp. 130 and Skinner 2011, pp. 70). Other examples include the artificial heart, Gliadel wafer (Deyo and Patrick, 2005, pp. 21,45) and "robotic" surgery tools (Chandra and Skinner, 2012, pp. 670).

Our model focuses on this type of treatments that get overvalued by the market. There are other ways in which the market (due to consumer misperception) can force health insurers to act in ways that are not socially efficient. The effects and remedies are similar to what we model below (see section 7.5).

3. Model

The model describes three markets (see figure 1). First, there is the treatment procurement market. Second, the market where research labs sell directly to (uninsured) consumers. Finally, the health insurance market where insurance companies compete for customers.

On the treatment procurement market, research labs bargain with insurers. For instance, in the pharmaceutical sector, PBMs (Pharmacy Benefit Management firms) are important players on the treatment procurement market (Scherer 2000, pp. 1325 and Dranove 2000, pp. 107). If an insurer and lab come to an agreement on prices, the insurer can cover the lab's treatment in its health insurance contracts. These contracts are sold to consumers on the health insurance market. If a treatment is not covered by a patient's insurance contract, she can buy the treatment on the uninsured market.

Treatments can be sold on the uninsured market for three reasons: (i) a consumer can decide not to buy insurance for a treatment; this treatment then needs to be bought on the uninsured market, (ii) a lab can decide to sell (only) directly to consumers and not contract with insurers and (iii) the government can decide that insurers are not allowed to cover a certain treatment. In that case, the treatment can only be sold on the uninsured market.

First, we describe consumer behavior. Then we specify how firms interact.

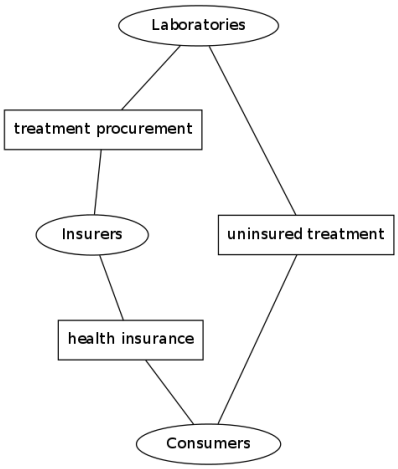

Figure 1: Three sets of players (oval) and three markets (rectangle).

3.1. utility

The market imperfection that we are interested in is that consumers do not know the value of each treatment when buying insurance. So they may over/undervalue coverage of treatments. In the medical literature, there are a number of studies documenting consumers' misperceptions of the value of treatments (see Habib et al., 2008; Mello and Brennan, 2001; Ravdin et al., 1998; Rothberg et al., 2010; Siegel et al., 2008). However, there is little evidence that consumers get the value of treatments wrong in a systematic way. Cremieux et al. (2007), Cutler and McClellan (2001), Cutler et al. (2006), Cutler (2007) and Kleinke (2001) all suggest that over the past decades in the US the growth in health care expenditure due to the introduction of new treatments has been worth it in terms of value created.[8] If consumers would consistently over-estimate the value of new treatments, treatments would be adopted where the value created does not cover costs, which is inconsistent with this evidence.[9]

A simple way to capture that consumers –when buying insurance– can overvalue some treatments but are not systematically wrong is to assume that they can get mixed up. In particular, they may incorrectly attribute the value of one treatment to another treatment. To illustrate, they have read in the newspaper that there is a new breakthrough treatment for heart failure. They read their insurance contract which reimburses a new heart treatment; is this the one they read about? Alternatively, two treatments are always used together (like beta blockers after heart surgery; Skinner, 2011, pp. 66). Hence, from past experience of family and friends, consumers know the value of this combination correctly but may attribute incorrectly value to each individual treatment.

[8]Because consumers overvalue some treatments (see examples in section 2) but are right on average, it must be the case that they tend to undervalue other treatments. Kleinke (2001) argues that there are drugs with high benefits that are not valued by the market. Volpp et al. (2008) discusses a scheme to pay people for taking their medication; suggesting that people undervalue the benefits of these treatments.

[9]Note that the evidence by Cutler and others is also consistent with consumers consistently under-estimating the value of new treatments. However, this seems unlikely given the examples mentioned in section 2. Further, policy makers worry about the high growth rate of health care expenditures which is partly caused by the development of new treatments (see Chernew and Newhouse (2011) for an overview). This seems hard to reconcile with the assumption that people tend to systematically under-estimate the value of new treatments.

More formally, we assume that consumers partition the set of all treatments into subsets. Elements in such a partition could be cancer treatment using chemo therapy, eye treatment using laser technology, etc. Consumers are rational in the sense that within a subset of treatments they know the total value that these treatments create. However, they may get confused –when buying insurance– about the value created by a particular treatment.

A simple model to capture this is the following. There are two new treatments 1 and 2 which treat different (but related) conditions.[10] Just to be clear, treatments 1 and 2 are in the same subset of the partition of all treatments (hence they "look similar" to the consumer), but the treatments do not compete: they cure different conditions or are used in different circumstances. The issue here is not that the consumer gets treated with the wrong medicine.

There is a probability $\mu_1 > 0$ that a consumer needs treatment 1 and a probability $\mu_2 >$ that she needs treatment 2.[11] To ease notation, we assume that $\mu_1 = \mu_2 = \mu$ and the consumer never needs both treatments (no co-morbidity).[12]

Treatment i raises utility by $q_i > 0$ and is produced with constant marginal cost of production c_i. We assume that the treatment is safe. That is, a government agency like the FDA excludes treatments with $q_i < 0$. The current best treatment that cures the same condition as i increases a patient's utility by q_i^o and is produced with constant marginal costs c_i^o. Hence the expected social value of innovation i equals $\mu(q_i - c_i - (q_i^o - c_i^o))$. To ease notation, we normalize $c_i = c_i^o = q_i^o = 0$. Hence the value added of treatment i equals q_i. The question is: does the price that the lab inventing i receives on the market reflect its value q_i?

To simplify further, we assume that $q_i \in \{q^l, q^h\}$ with $q^h > q^l > 0$.[13] We interpret q^h as a breakthrough treatment that offers big improvements over existing treatments. A q^l treatment is a marginal innovation that offers hardly an improvement over current best practice.

Before the consumer decides on which insurance contract to buy, she receives a signal about treatment i's quality level. The signal is either H or L indicating that i's quality is q^h or q^l resp. With probability $g \in [\frac{1}{2}, 1]$ the signals correctly reflect q_1, q_2. With probability $1 - g$, the consumer gets confused and attributes q_1 to treatment 2 and q_2 to treatment 1. Let hl denote the state of the world where one treatment has high quality and the other low quality. Then there are three states of the world: ll, lh and hh. The assumption implies that the consumer is never wrong about the state of the world; you cannot confuse the value of treatments when these values are the same. But in state lh she may attribute the wrong value to a particular treatment.[14] This simple information structure causes a reversion to the mean in terms of treatment values.

If the consumer receives signal H for treatment i and L for $j \neq i$, she updates her beliefs

[10]The reason why the conditions should be related is that it seems unlikely that a consumer would confuse a treatment against a cold with a form of heart surgery.

[11]We do not consider adverse selection problems here; hence μ_i is the same for each consumer. Also there is no moral hazard: consumers only get the treatment in case they need it.

[12]Allowing for co-morbidity complicates the expression of the profit maximizing price on the uninsured market, without adding much insight. See section 7.

[13]See the webappendix at `http://sites.google.com/site/janboonehomepage/home/webappendices` for the case where q_i is drawn from a more general distribution.

[14]If we would allow for the case where in state ll (hh) the consumer receives a signal H (L), she would also over- (under)estimate the value of treatments in states ll (hh). This further reduces labs' incentives to invest in R&D and increases the benefits of government intervention.

about quality in the following way:[15]

$$\bar{q}_i = gq^h + (1-g)q^l \quad (1)$$
$$\bar{q}_j = gq^l + (1-g)q^h \quad (2)$$

If the signal is perfectly informative, $g = 1$ and $\bar{q}_i = q_i$ in each case (no reversion to the mean). We say in this case that the consumer is perfectly informed. If the signal is uninformative $g = \frac{1}{2}$ and $\bar{q}_i = \bar{q}_j = (q_i + q_j)/2$. Note that it is always the case that $\bar{q}_i + \bar{q}_j = q_i + q_j$: consumers are not systematically wrong. They know the aggregate value of new treatments which is consistent with the evidence by Cutler and others cited above. We say that consumers over(under)value treatment i if $\bar{q}_i > (<) q_i$.

The important characteristic of this information structure, that we use below, is that for $g < 1$ there is an externality between treatments in the same set of the partition: $\bar{q}_i$ depends on q_j $(j \neq i)$.

We assume all consumers have the same partition of treatments and receive the same signal. This implies that everyone has the same valuation for insured coverage of a treatment. Hence there is no role for price discrimination in the health insurance market.

Two remarks here. First, extending the model to T different treatments (instead of 2) that the consumer cannot distinguish is straightforward but tedious as the number of cases that needs to be considered increases. This will become clear below. Second, the utility structure that we use below is linear. Hence, there can be subsets of treatments (in the consumer's partition of all treatments) for heart, lung, liver etc. conditions. The overall effect is then the sum of the effects for each case (assuming no co-morbidity). Hence, from now on, we focus on one subset of the partition of all treatments.

An agent can buy health insurance at premium σ. Consumers differ in income such that some may not be able to afford treatment on the uninsured market. However, we assume that μ is small enough that everyone can afford to buy insurance. Hence we do not consider a role for the government in reducing insurance coverage to keep health insurance affordable for everyone. Although this is an important topic as well (Garber, 2001), it is beyond the scope of this paper. See Boone (2013) for an analysis of this case.

The set of treatments covered by agent l's health insurance is denoted by ι_l. We assume that health insurance covers the costs of these treatments completely.[16] The set of treatments not covered is denoted by υ_l. Hence we have $\iota_l \cup \upsilon_l = \{1, 2\}$ and $\iota_l \cap \upsilon_l = \emptyset$.

In case agent l needs a treatment j in the set υ_l, she can decide to buy that treatment at uninsured price p_j^u. We assume that the agent cannot spend more than β_l on the uninsured market due to a budget constraint. This budget constraint is the reason why the agent buys health insurance. This is different from the standard argument for buying insurance: a concave utility function which makes the agent risk averse. The following consideration motivates this modelling choice.

As mentioned in the introduction, we are interested in access to care. There is empirical evidence to suggest that budget constraints play an important role in the decision whether or not

[15]Below we consider symmetric labs, hence there is no reason for consumers to expect a priori that, say, $q_1 = q^h$ is more likely than $q_2 = q^h$. Further, we assume that consumers cannot observe labs' R&D investments, hence these cannot be used to affect consumers' beliefs.

[16]Because in the model there are no moral hazard nor adverse selection issues on the side of patients/consumers, there is no need for insurers to offer less than full coverage.

to buy treatment when people are uninsured (or under-insured). See, for instance, Piette et al. (2004a), Piette et al. (2004b), Goldman et al. (2007), Schoen et al. (2008) and Schoen et al. (2010) for evidence of access (to health care) problems due to financial constraints. That is, even if the treatment is worth it (in the sense that q_i exceeds the price on the uninsured market, p_i^u), the agent may not be able to afford it. There may be other things that the family needs to buy and it may not be possible to borrow enough money to finance the treatment. By pooling risks, health insurance allows the agent to afford treatment if she needs it. Nyman (1999) is the first to estimate the access value of health insurance. He argues that this access value exceeds the value of health insurance for financial risk-bearing.

Further, we consider a policy where the government excludes certain treatments from health insurance coverage. On the one hand, this may improve dynamic efficiency (see below), but the trade off is that some patients cannot afford the excluded treatments.

We assume that $\beta \geq 0$ is distributed with density (cumulative distribution) function $f(F)$. This distribution of budgets that an agent can spend on the uninsured market captures the income distribution.

Finally, the focus of this paper is on the imperfection that people do not necessarily know the value of a treatment at the moment that they buy health insurance. At the moment that someone needs a specific treatment, she usually decides together with her physician. We assume that the physician reveals q_i to a patient who needs treatment i. A patient who is not insured for treatment i compares the price of i on the uninsured market, p_i^u with q_i and the budget constraint β that she faces. An insured patient will always use treatment i as the cost to the patient is zero and $q_i > 0$. Such an insured patient, after learning q_i may see that the value added of the treatment does not exceed the price the insurer has to pay; see Chandra et al. (2011, pp. 27) for an example of this case.

Although there may be incentive issues with physicians and physicians may not know the value of each treatment exactly themselves (Skinner, 2011; McGuire, 2011a), it seems reasonable to assume that the patient is better informed about q_i at the moment that she needs the treatment than at the moment that she buys insurance. Few people go over the list of insured treatments with their physician to determine the value of each single treatment. In fact, most doctors do not want to give such advice (Liebman and Zeckhauser, 2008, pp. 7).

Expected utility of a consumer l with budget β_l, set of (un)insured treatments ι_l (υ_l) and updated beliefs $\bar{q}_1, \bar{q}_2$ at the moment of buying insurance is written as:[17]

$$U_l = E\left(\sum_{j\in\iota_l} \mu q_j - \sigma + \sum_{j\in\upsilon_l | p_j^u \leq \beta_l} \max\{0, q_j - p_j^u\} | \bar{q}_1, \bar{q}_2\right). \tag{3}$$

3.2. pricing games

There are three types of agents. First, there are consumers (described above) with mass normalized to 1. Second, there are $n \geq 2$ competing health insurers. Insurer $I_i (i \in \{a, b, ..., n\})$ contracts with research labs on the price for a treatment in case one of I_i's insured needs such

[17]Note that being in state j (which means that the agent is ill in some way) can reduce the agent's utility. Since we use an additive utility structure (as in Garber et al. 2006) this is just a (negative) constant that can be added to the agent's expected utility. In other words, taking the disutility of illness into account does not affect the agent's optimal decision (while it would complicate notation).

treatment. And insurers compete in prices when selling insurance to consumers. If the insurers offer the same set ι of covered treatments, they compete in homogeneous products. If different sets ι of covered treatments are offered at different prices, the utility structure described above determines which insurer will sell in equilibrium. As usual, if consumers are indifferent between the contracts offered by a group of insurers, they choose a contract randomly and end up being distributed equally over these insurers. Further, if a consumer is indifferent between buying insurance for some treatment(s) or not at all, she buys insurance.

Finally, there are the research labs L_j with $j \in \{1, 2\}$. A lab decides on contracts with insurers and on a price p_j^u at which L_j sells its treatment to uninsured patients. The government's CE analysis takes the form of a cut off level $\underline{q} \geq 0$ such that only treatments with $q_j > \underline{q}$ are allowed to be covered by health insurance.

We need to determine the equilibrium prices charged on the three markets in figure 1. We start with the procurement market.

We denote the price per patient that I_i pays to L_j for its treatment $\tilde{p}_{ij} \geq 0$. We allow the insurer and lab to use two-part tariffs. The fixed part is denoted by t_{ij} (if $t_{ij} < 0$ the lab pays the insurer). In the equilibrium we analyze in this section, it turns out that prices are linear ($t_{ij} = 0$ for all i, j).[18]

The assumption that research labs set two prices (one for insured and one for uninsured patients) is in line with reality. This can be implemented in two ways (see e.g. Scherer (2000, pp. 1326) for details): (i) insurers buy from wholesalers at the list price and then the manufacturer issues a rebate check to the insurer for the amount of the negotiated discount (in our notation this discount equals $p_j^u - \tilde{p}_{ij}$), (ii) wholesaler supplies the insurer at the discounted price and receives a "chargeback" payment from the manufacturer to cover the difference.

Within the model, we need the labs to commit to a price $\tilde{p}_{ij}$ to insurers. If this would not happen, and an insurer would add treatment j to its set of insured treatments, L_j would set a very high price. Given that an insurer covers j in its contract, its insured will demand treatment j as the insured do not care about the price. To prevent such hold up problems, the insurer needs to contract with L_j before adding j to the set of treatments that are reimbursed by its insurance contract.

There are two contracting assumptions that look reasonable in this context. First, when an insurer has contracted with L_j and agreed on a price $\tilde{p}_{ij}$, the insurer can still ask its customers to buy treatment j on the uninsured market if this is cheaper ($p_j^u < \tilde{p}_{ij}$) and then reimburse them afterwards. It seems impossible for L_j to make sure that it only sells at p_j^u to customers without insurance. How would a customer prove to L_j that she does not have health insurance? Hence in equilibrium it will be the case that $p_j^u \geq \tilde{p}_{ij}$. Second, a contract between I_i and L_j cannot force I_i to offer coverage of treatment j. The contract is an option to buy; not a requirement for I_i to buy. In reality, we are not aware of quantity forcing contracts on this market. One explanation for this can be that at the end of the day it is not the insurer who decides on treatments but a physician. Hence, it is hard for an insurer to influence the quantity bought

[18] We assume that usage can be monitored such that $\tilde{p}_{ij}$ can indeed be used. This is clear in the pharmaceutical sector. It may be less obvious –but not impossible– in case of a medical device like an MRI scan. Also in the case of medical devices, labs can use service contracts that allow for metering. Further, use of MRI scan or surgical procedures will usually be charged by a provider (e.g. hospital) not necessarily by the inventor. To simplify the model, we have not added providers as a separate set of players in figure 1. We leave the analysis of such cases for future research.

of a certain treatment.

The timing of the game is as follows:

0. government sets minimimum standard $\underline{q}$ for treatments that can be insured
1. labs L_1, L_2 invest in R&D to invent new treatments
2. consumers learn signals and derive $\bar{q}_1, \bar{q}_2$
3. R&D labs contract privately with insurers
4. insurers decide on which treatments to cover ι and set σ (interim unobservability)
5. consumers decide whether to buy insurance and –if so– from which insurer
6. R&D labs set uninsured prices p_i^u
7. consumers fall ill and need treatment which is either covered by their insurance or not

Here we only consider stages 2 and further. Stages 0 and 1 are analyzed below. We assume in this section that $\underline{q} = 0$; that is, no government intervention. Hence each treatment can be covered by health insurance.

In stage 3, firm L_j offering treatment j that leads to utility q_j (for the patient who needs this) bargains with insurers to get her treatment covered by insurance. This bargaining is modeled in two ways: we consider both an offer and a bidding game. If consumers are perfectly informed ($g = 1$), the situation where upstream labs make offers to downstream insurers is an offer game (following the terminology in Segal and Whinston 2003, pp. 758). We extent the term offer game also to the situation where the consumer only knows $\bar{q}_1, \bar{q}_2$ (with $g < 1$) and hence there are externalities between upstream firms. In the bidding game, the downstream insurers make offers to the upstream labs.

In particular, in the offer game, the research labs make simultaneously and independently take-it-or-leave-it offers $(\tilde{p}_{ij}, t_{ij})$ to the insurer. Each insurer simultaneously and independently either accepts or rejects the offer made to it by a research lab. In the bidding game, the insurers make simultaneously and independently take-it-or-leave-it offers $(\tilde{p}, t)$ to the research labs. Then the labs decide simultaneously and independently whether or not to accept offers.

Most papers consider either an offer or a bidding game. Outcomes tend to differ in these two games. Intuitively, the party making the offer has more bargaining power and can appropriate a bigger part of the surplus generated. In our model, however, the outcomes hardly differ. This is because insurers compete in prices on the health insurance market and offer homogeneous products if they cover the same treatments. Even if insurers make the offers, the rents are competed away and appropriated by the research labs.

In line with how these negotiations work in reality (Danzon 2011, pp. 530 and Morton and Kyle 2011, pp. 790), we assume that offers are made privately.[19] To illustrate what this means: insurer I_a does not know the details of the contract that I_b has with L_j; similarly, L_j does not

[19]Two remarks on public offers. First, if offers are made publicly, consumers can –in principle– observe them as well. This creates possibilities for parties to signal quality by their contract offers. We do not think this is a relevant option in reality and signalling can be ignored with private offers. Second, from a theoretical point of view, it is straightforward to verify that the equilibrium in proposition 1 is also an equilibrium with public offers. The proof of this proposition is a bit more involved with private offers because we need to specify beliefs.

know the details of the contract that I_a has with the other lab, etc. In other words, in both games parties cannot commit to stop renegotiating contracts with other parties. That is, even if, say, L_1 would publicly announce its contract, I_a cannot be sure that L_1 does not privately offer I_b a discount on the publicly announced prices. Since the details of the contracts between insurers and pharmaceutical companies remain confidential, insurers set prices without knowing the details of each others' contracts. Rey and Vergé (2004) call this interim unobservability.

As discussed by Segal (1999), the literature using private contracts often assumes, so called, passive beliefs. With passive beliefs, if a firm receives an out-of-equilibrium offer, it assumes that all other firms did receive the equilibrium offer. This is not always a reasonable assumption and we do not assume passive beliefs here for two reasons. First, for pricing games (as we have here) Rey and Vergé (2004) derive a condition under which there is no equilibrium with passive beliefs. Indeed in our set up, an equilibrium with passive beliefs does not exist.[20] Second, it seems more reasonable that a firm receiving an out-of-equilibrium offer asks "given the offer that I have received, what are the optimal offers from this party to my competitors?" These are called wary beliefs (McAfee and Schwartz, 1994), which we use below.

Following McAfee and Schwartz (1994) and Rey and Vergé (2004), we define wary beliefs in the offer [bidding] game as follows. When $I_i[L_j]$ receives a contract offer $(\tilde{p}_{ij}, t_{ij})$, it believes that

- $L_j[I_i]$ expects it to accept this contract
- $L_j[I_i]$ offers $I_k(k \neq i)[L_k(k \neq j)]$ the contract $(\tilde{p}_{kj}, t_{kj})[(\tilde{p}_{ik}, t_{ik})]$ that maximizes its profits among the contracts that are acceptable to $I_k[L_k]$, and
- $I_k[L_k]$ reasons in the same way.

We assume that it is not possible to contract with agent i on the details of the contract with j. In the next section, we focus on simple contracts that only specify $\tilde{p}_{ij}, t_{ij}$. In section 6, we allow parties to use exclusion clauses. In that case, parties can contract on whether an insurer offers coverage of certain treatments, or not.

In period 4, the insurance companies compete in determining which treatments to cover and setting prices (insurance premium σ). In period 5, consumers decide whether and (if so) which insurance contract to buy. In period 6, research labs set prices for people without insurance. Finally, some people fall ill and need to be treated either financed by their insurance company or they buy treatment on the uninsured market (if they can afford it).

4. Market equilibrium

This section characterizes the static and dynamic efficiency of the market outcome (without government intervention). As long as the uninsured market is sufficiently unprofitable for h-treatments, the insurance market does not unravel. In this case the market outcome is efficient from a static point of view. Unless consumers are perfectly informed ($g = 1$ in equations (1), (2)) there is underinvestment in R&D in the market outcome.

On the other hand, the proof of proposition 2 –which derives bounds on equilibrium payoffs– is a bit more involved with public offers as the acceptance decision depends on an agent's own offer and on the offers to other parties. With private offers, the acceptance decision can only depend on an agent's own offer.

[20] See footnote 23 below.

4.1. static efficiency

In both the bidding and the offer game we first characterize a symmetric perfect Bayesian equilibrium. Then we derive upper and lower bounds on payoffs in any equilibrium outcome.

First, we need some notation. We define $p^* = \arg\max_p p(1-F(p))$. This is the profit maximizing uninsured price for a treatment with $q_j \geq p^*$. As patients are guided by their physicians when choosing treatment, they never buy treatment j on the uninsured market if $p_j^u > q_j$. We assume that $p(1-F(p))$ is non-decreasing in $p \in [0, p^*]$. Hence the monopoly price that maximizes profits on the uninsured market is given by $p_j^m = \min\{p^*, q_j\}$, where j denotes lab L_j or its quality level $j \in \{l, h\}$.

It is useful to introduce a smallest money unit $\varepsilon > 0$ in order to break ties. We assume that the money unit is small in two ways:

$$\varepsilon < \mu \frac{q^l}{n} \text{ and } \varepsilon < \frac{q^h - q^l}{4} \tag{4}$$

As we think of ε as a dollar or euro or even a cent, this assumption does not seem stringent.

The interesting case is the one where some people cannot afford to buy the h-treatment on the uninsured market if it is sold at the social value it creates:

$$F(q^h) > 0 \tag{5}$$

Indeed, if we would assume that $F(q^h) = 0$, there is no need to have insurance in the model (as everyone has access to all treatments without insurance) and the first best outcome can be achieved.

We want to understand whether the market biases research towards marginal l-innovations at the expense of R&D for breakthrough h-treatments. In order to analyse this, we need to know the profits that the research labs make in the market outcome.

Proposition 1 *Consider the case with $\underline{q} = 0$ and assume without loss of generality that $q_1 \leq q_2$. In a symmetric perfect Bayesian pure strategy equilibrium with wary beliefs we have the following for insurer I_i and lab L_j.*[21] *In the offer game:*

- *if $\bar{q}_2 > (1 - F(p_2^m))p_2^m$:*

$$\iota = \{1, 2\} \qquad \pi_i = 0 \qquad \sigma = \mu(\bar{q}_1 + \bar{q}_2) \tag{6}$$

$$\pi_j = \mu \bar{q}_j \qquad \tilde{p}_j = \bar{q}_j \qquad t_j = 0 \qquad p_j^u \geq \bar{q}_j \tag{7}$$

- *if $\bar{q}_2 < (1 - F(p_2^m))p_2^m$:*

$$\iota = \{1\} \qquad \pi_i = 0 \qquad \sigma = \mu q^l \tag{8}$$

$$\pi_1 = \mu q^l \qquad \tilde{p}_1 = q^l \qquad t_1 = 0 \qquad p_1^u \geq q^l \tag{9}$$

$$\pi_2 = \mu(1 - F(p_2^m))p_2^m \qquad p_2^u = p_2^m \tag{10}$$

[21] The case with $\bar{q}_2 = (1 - F(p_2^m))p_2^m$ in the offer game ($\bar{q}_2 - \varepsilon = (1 - F(p_2^m))p_2^m$ in the bidding game) is a bit tedious because of ties. As it is non-generic (small perturbations to g and thus to $\bar{q}_{1,2}$ make this disappear) it is not considered here.

In the bidding game:

- *if* $\bar{q}_2 - \varepsilon > (1 - F(p_2^m))p_2^m$*:*

$$\iota = \{1, 2\} \qquad \pi_i = 0 \qquad \sigma = \mu(\bar{q}_1 + \bar{q}_2 - 2\varepsilon) \tag{11}$$

$$\pi_j = \mu(\bar{q}_j - \varepsilon) \qquad \tilde{p}_j = \bar{q}_j - \varepsilon \qquad t_j = 0 \qquad p_j^u \geq \bar{q}_j - \varepsilon \tag{12}$$

- *if* $\bar{q}_2 - \varepsilon < (1 - F(p_2^m))p_2^m$ *and* $(1 - F(p_1^m))p_1^m \leq q_1 - \varepsilon$*:*[22]

$$\iota = \{1\} \qquad \pi_i = 0 \qquad \sigma = \mu(q^l - \varepsilon) \tag{13}$$

$$\pi_1 = \mu(q^l - \varepsilon) \qquad \tilde{p}_1 = q^l - \varepsilon \qquad t_1 = 0 \qquad p_1^u \geq q^l - \varepsilon \tag{14}$$

$$\pi_2 = \mu(1 - F(p_2^m))p_2^m \qquad p_2^u = p_2^m \tag{15}$$

The proof in the appendix shows that there are no profitable deviations from the equilibrium in the proposition.[23] As we think of the money unit ε as being small, equilibrium payoffs are basically the same in the offer and bidding games.

If expected quality $\bar{q}_2$ exceeds uninsured profits per patient for the (weakly) highest quality treatment $(p_2^m(1 - F(p_2^m)))$, both treatments are covered in equilibrium. Firms use linear contracts with $\tilde{p}_j = \bar{q}_j$ for each treatment j (and $t_j = 0$). In equilibrium, insurers make zero profits as Bertrand competition forces the premium down to what they have to pay the labs. Lab L_j makes a profit equal to $\mu\bar{q}_j$. Hence a low quality lab with $q_j < \bar{q}_j$ free rides on the high quality lab $(q_j > \bar{q}_j)$. The l-treatment gets covered by health insurance and the l-lab is paid more than its treatment is worth.

In other words, when firms use contracts that just specify prices (and no exclusion clauses) it is not possible for insurers to "correct" consumers' incorrect perception of a treatment's value (in case $\bar{q}_j \neq q_j$). In this sense, there is not much reason for private insurers to do CE analysis (as, indeed, they hardly do in the US). This incorrect perception filters through in the profits that labs earn in equilibrium.

In the model, we can assume that insurers know $q_{1,2}$. Yet, the prices that they pay research labs are determined by consumers' perceptions $\bar{q}_{1,2}$. The equilibrium suggests that if insurers would have to invest in CE analysis to find out $q_{1,2}$, they would in fact be better off investing in marketing surveys to find out how consumers value treatments. This can explain why private health insurers in the US hardly invest in CE analysis, as mentioned above.

Profit on the uninsured market can only exceed $\mu\bar{q}_j$ if one lab has a high quality treatment and the other a low quality treatment.[24] The trade off for the h-lab is then the following. On the one hand, it can sell on the uninsured market, but then $1 - F(p_h^m)$ patients cannot afford the treatment. On the other hand, when selling on the insured market, everyone can afford the treatment. But consumers value the h-treatment at $\bar{q}_j < q^h$.

[22] If $(1 - F(p_1^m))p_1^m > q_1 - \varepsilon$, then both treatments are sold on the uninsured market.

[23] To see why an equilibrium with passive beliefs in the offer game does not exist, consider $\tilde{p}_j \geq 3\varepsilon$. Then a deviating offer $\hat{p}_j = 0, \hat{t}_j = \mu(\tilde{p}_j - \varepsilon)$ is accepted by an insurer. Making this deviating offer to all $n \geq 2$ insurers is a profitable deviation for the lab. In case $\tilde{p}_j \leq 2\varepsilon$, lab can offer $\hat{p}_j = \bar{q}_j - \varepsilon, \hat{t}_j = 0$. This offer is accepted by an insurer since increasing σ by $\Delta\sigma_j = \mu(\bar{q}_j - \varepsilon)$ and selling coverage of j is more attractive for consumers than not covering treatment j.

[24] For $g \in \langle \frac{1}{2}, 1 \rangle$, $\bar{q}_j$ can take on two values: (i) $\bar{q}_j = gq_j + (1 - g)q_i$ if the signal is correct and (ii) $\bar{q}_j = gq_i + (1 - g)q_j$ if the signal is not correct.

If the profit on the uninsured market is higher, the insurance market unravels as it is not profitable for the h-lab to have its treatment covered by health insurance. The h-treatment is not covered by health insurance and only the l-treatment can be insured. If consumers see that only one treatment is covered by health insurance, they understand that the h-treatment was withdrawn from the insurance market as $p_h^m(1-F(p_h^m)) > \bar{q}_j$. They infer that the covered treatment has value q^l.

Note that the market unravels only if the country is relatively rich in the sense that some people can afford to buy the h-treatment without insurance at a high enough price that uninsured profits exceed insured profits for the h-lab. If the market does not unravel, it is efficient from a static point of view. Indeed, everyone has access to the care that they need and prices are only transfers from a total welfare point of view (but prices do affect dynamic efficiency, as we see shortly).

If the market does unravel, then the loss in static efficiency is given by $F(p_h^m)q^h$: fraction of people who cannot afford the h-treatment on the uninsured market times their utility loss. Let e denote the equilibrium probability that a lab invents a h-treatment characterized below. The probability that there is one h and one l treatment is given by $2e(1-e)$. If $p_h^m(1-F(p_h^m)) > gq^h + (1-g)q^l$, then the probability that the market unravels equals $2e(1-e)$. If $gq^h + (1-g)q^l > p_h^m(1-F(p_h^m)) > gq^l + (1-g)q^h$, the probability that the market unravels equals $2e(1-e)(1-g)$. If the market does not unravel, the market remains efficient from a static point of view.

Note that if consumers are perfectly informed ($g = 1$), each lab receives the value it creates (μq_j) as profit. In this case, the market is always efficient from a static and (as we will see below) dynamic point of view. Hence the model is set up in such a way that with fully informed consumers there is no reason for government intervention. This is due to the assumptions that everyone can afford health insurance and each innovator is a monopolist (hence there is no appropriability effect nor a business stealing effect, see Mankiw and Whinston 1986).

Finally, even if consumers are not perfectly informed ($g < 1$), in equilibrium they value their insurance contract correctly. If the market does not unravel, we have $\mu(\bar{q}_1 + \bar{q}_2) = \mu(q_1 + q_2)$. If the market does unravel, consumers understand that the insured treatment has quality q^l. This is consistent with evidence by Cutler and others cited in section 3.1 that the market does not systematically over-value new treatments (although the market does overpay for some new treatments).

The equilibrium payoffs determine the incentives for labs to invest in R&D (in stage 1 of the game). In order to be able to determine labs' R&D efforts, we need to know that labs' profits do not vary (a lot) across equilibria.

Proposition 2 *Assume that the insurance market does not unravel. The labs' equilibrium profits for the offer game satisfy* $\pi_j \in [\mu(\bar{q}_j - \varepsilon), \mu\bar{q}_j]$. *In the bidding game, assuming* $n \geq 3$ *the equilibrium profits for research labs satisfy* $\pi_j \in [\mu(\bar{q}_j - 2\varepsilon), \mu(\bar{q}_j - \varepsilon)]$.

As noted above, if $(1-F(p_j^m))p_j^m > \bar{q}_j$ then the insurance market unravels and L_j's profits can exceed $\mu\bar{q}_j$. As shown in the appendix, assuming $n \geq 3$ in the bidding game allows to avoid some ties in the proofs which simplifies the exposition. This assumption does not seem restrictive.

The intuition why profits cannot vary by much across equilibria is the following. First, total profits earned can never exceed the total value created $\mu(q_1+q_2)$. If this would happen, there is

a player making losses who then has a profitable deviation. If lab L_j would earn $\pi_j > \mu\bar{q}_j$ on the treatment procurement market, insurers can profitably deviate by dropping the contract with j. This reduces their costs by π_j while consumers reduce their valuation of such a deviating insurance contract by $\mu\bar{q}_j < \pi_j$. Hence costs fall faster than revenue (insurance premium σ). Second, a lab's profits cannot be much below $\mu\bar{q}_j$ as again there would be a profitable deviation. In the offer game, the lab itself would raise $\tilde{p}_j$ to increase its profits. In the bidding game, an insurer would try to earn monopoly profits by outbidding it's competitors in case π_j was much below $\mu\bar{q}_j$ for lab L_j.

As we view the money unit ε as being small, to ease notation we write the equilibrium profits for the research labs as in equations (7,9,10) for both the offer and the bidding game.

4.2. dynamic efficiency

To evaluate the dynamic efficiency of the equilibrium in proposition 1, we introduce a simple R&D technology to determine labs' incentives to innovate.

Labs can invest effort $e \in [0, 1]$ to invent a high quality treatment where e equals the probability that a h-treatment is found. Effort cost is given by $c(e)$ with $c(0) = 0, c'(e), c''(e), c'''(e) > 0$.[25] Hence it is increasingly more costly to find a h-treatment with higher probability. To ease the exposition, we prefer to ignore corner solutions in e; hence, we assume $c'(0) = 0, c'(1) = +\infty$.

To see whether the market is dynamically efficient, first consider the first best R&D effort e^*. First best effort solves the following maximization problem.

$$\max_e \mu(eq^h + (1-e)q^l) - c(e)$$

Under first best, everyone (who needs it) has access to a treatment. Hence the social value of a treatment q^i equals μq^i.

The first order condition for first best effort can be written as

$$c'(e^*) = \mu(q^h - q^l) \tag{16}$$

In the market outcome, we find the following.

Proposition 3 *If $g = 1$ then the market is dynamically efficient: $e^m = e^*$. For $g < 1$, the market under-invests in breakthrough treatments: $e^m < e^*$.*

Unless consumers know the value of each treatment when buying insurance, the market under-invests in R&D: $e^m < e^*$. By under-rewarding h-treatments and over-rewarding l-treatments, the probability of inventing l-treatments is higher than in first best. So, indeed, in this model the market is biased towards introducing marginal improvements over current treatments as argued in Deyo and Patrick (2005) and Light and Lexchin (2012). It is more attractive to introduce marginal improvements over current treatments than invest to find a breakthrough treatment.

[25] This is a simple way to model the intensive margin of R&D effort. Alternatively, one can think of a set-up where research ideas have an exogenous quality level q. Then the intensive margin plays a role if, say, a lab has three ideas but due to (capital or labor market) constraints can only develop two of these ideas.

5. Government intervention: CE analysis

The government policy that we consider here is that the government performs CE analysis to determine which treatments are allowed to be covered by (private) insurance. CE analysis takes the form here of a cut off value $\underline{q}$ such that treatments with value added $q \leq \underline{q}$ are not allowed to be covered by health insurance while treatments with $q > \underline{q}$ are allowed to be covered by insurance.

This is different from the current use of CE analysis by an institute like NICE. In the UK, NICE determines whether a treatment can be covered by *public* health insurance. We analyse whether in a *private* health insurance market there is a role for the government determining which treatments can be covered by health insurance. This government intervention in determining private health insurance coverage is recommended by, for instance, Diamond (1992, pp. 1242), Kleinke (2001, pp. 58) and Deyo and Patrick (2005, pp. 270). We derive conditions under which the government can raise welfare by implementing $\underline{q} = q^l$.

5.1. static efficiency

If the government excludes l-treatments from insurance coverage ($\underline{q} = q^l$), we get the following equilibrium. As above, we ignore the money unit ε.

Proposition 4 *With $\underline{q} = q^l$, a treatment is covered by insurance if and only if it has quality q^h. A lab that invents a l-treatment has profits equal to $\pi_l = \mu(1 - F(p_l^m))p_l^m$. A lab with a h-treatment earns $\pi_h = \mu q^h$.*

Because l-treatments are excluded from health insurance coverage, consumers know that any treatment covered by insurance has value added q^h.[26] Hence, consumers value an insurance contract that covers one (two) treatment(s) at $(2)\mu q^h$. Due to competition between insurers, labs can appropriate this value.

A l-treatment can only be sold on the uninsured market and because $p_l^m \leq q^l$ we find that

$$\pi_l \leq \mu q^l \tag{17}$$

Unlike the outcome with an unraveling insurance market (without government intervention), the government outcome makes sure that everyone has access to the h-treatment. However, with government intervention l-treatments are never covered by health insurance. Unless $1-F(p_l^m) = 1$, the government outcome is not efficient from a static point of view. Some people who cannot afford to buy the l-treatment on the uninsured market have to go without.

Assume $g < 1$. Comparing static efficiency in the market without and with government intervention, we find the following. If the country is relatively poor in the sense that (i) the market does not unravel ($\bar{q}_h \geq (1 - F(p_h^m))p_h^m$) and (ii) some people cannot afford to buy the l-treatment on the uninsured market ($F(p_l^m) > 0$), the market yields higher static efficiency than government intervention. If the country is relatively rich such that (i) the market does unravel

[26] The result that government CE analysis leads to h-labs' profits equal to their social value is due to the assumption that $q_j \in \{q^l, q^h\}$. In a model where q_j is continuously distributed, the imposition of $\underline{q}$ by the government raises (reduces) the conditional expectation of q_j given that treatment j is (not) covered by health insurance. Also in such a model government intervention raises R&D incentives. See the webappendix in footnote 13 for an analysis of this case.

($\bar{q}_h < (1 - F(p_h^m))p_h^m$) and (ii) everyone can afford to buy the l-treatment on the uninsured market ($F(p_l^m) = 0$), government intervention yields higher static welfare.

In the case where the market unravels and $F(p_l^m) > 0$, it is not clear which yields higher static welfare. On the one hand, the loss in welfare is smaller for the outcome with government intervention ($F(p_l^m)q^l < F(p_h^m)q^h$ because $p_l^m \leq p_h^m$ and $q^l < q^h$). On the other hand, the weight of the loss in expected static welfare equals $2e^m(1 - e^m)$ in the market outcome if $gq^h + (1 - g)q^l < p_h^m(1 - F(p_h^m))$ ($2e^m(1 - e^m)(1 - g)$ if $gq^h + (1 - g)q^l > p_h^m(1 - F(p_h^m)) > gq^l + (1 - g)q^h$) and $2[(1 - e^g)^2 + e^g(1 - e^g)]$ with government intervention. As we show shortly, $e^g > e^m$, but in general we cannot rank the expected loss in static efficiency with and without government intervention. As noted above, if consumers are perfectly informed ($g = 1$), the market outcome is first best and government intervention can only reduce welfare.

5.2. dynamic efficiency

The following result shows that government intervention of the form $\underline{q} = q^l$ leads to higher R&D effort than in the market outcome. In fact, the R&D effort can be excessive in the sense that it exceeds e^*.

Proposition 5 *With $\underline{q} = q^l$, R&D effort is given by*

$$c'(e^g) = \mu(q^h - (1 - F(p_l^m))p_l^m) \tag{18}$$

If $1 - F(q^l) = 1$ then $e^g = e^*$. Otherwise $e^g > e^*$.

If everyone can afford to buy l-treatment at its social value ($1-F(q^l) = 1$), then the outcome with government intervention is first best. The R&D incentives are first best and the outcome is also static efficient as everyone can afford to buy the l-treatment at $p_l^m = q^l$. In this case, government intervention dominates the market outcome $W^g = W^* > W^m$. Intuitively, in this case, each lab L_j earns its social value μq_j as a profit.

If $1 - F(q^l) < 1$, $\pi^l < \mu q^l$ while $\pi_h = \mu q^h$ and government intervention leads to over-investment in R&D. Government intervention removes the market bias toward l-treatments. But it is not obvious that government intervention raises (static and dynamic) welfare compared to the market outcome. If the over-investment distortion exceeds the under-investment distortion, it is better that the government does not impose $\underline{q} = q^l$. It is not hard to generate examples where this is the case. In particular, this can happen when a country is so poor that π_l is very low under government intervention and therefore the over-investment distortion is big.

Example 1 *Assume all consumers have the same income $\beta < q^l$ and the quality signal is uninformative $g = \frac{1}{2}$. Further, $c(e) = \frac{1}{2}e^2$. In this case, the market does not unravel as $p_l^m = p_h^m = \beta < q^l < \bar{q}_l = \bar{q}_h = \frac{q^h+q^l}{2}$. Expected total welfare is given by*

$$W = 2\mu(e^2q^h + e(1 - e)(q^h + q^l) + (1 - e)^2q^l) - \tfrac{1}{2}e^2$$

which is quadratic in e. We find that[27]

$$e^* = \mu(q^h - q^l)$$
$$e^m = \mu(q^h - q^l)/2$$
$$e^g = \mu(q^h - \beta)$$

Because W is quadratic, it is the case that $W^m > W^g$ if and only if $e^g - e^* > e^* - e^m$. Hence, if $q^h < 3q^l$, there exists β low enough such that government intervention stimulates R&D but reduces overall welfare W.

Summarizing we find the following. Government intervention $\underline{q} = q^l$ raises welfare compared to the market outcome if most people can afford to buy the l-treatment on the uninsured market at a price close to q^l. As most people can afford to buy the l-treatment, the static efficiency loss is small. Since most people can afford to pay q^l, π_l is close to μq^l and the over-investment distortion is small.

Hence, if a country is fairly rich (in the sense that most people can afford to pay q^l on the uninsured market), government using CE analysis to determine which treatments can be covered by health insurance tends to be welfare enhancing.

If the country is fairly poor such that (i) the insurance market does not unravel and (ii) the over-investment distortion (with government intervention) exceeds the under-investment distortion (without intervention), overall welfare is higher without government intervention.

6. Market solutions

This section considers potential market solutions to the problem that consumers do not know the value of treatments at the moment that they buy health insurance.

6.1. Exclusivity clauses

Above, we derived conditions under which government intervention improves on the market outcome by excluding l-treatments from health insurance. Question we analyze here is: if we allow for richer contracts, can the market replicate the government outcome? A full analysis of all contracting possibilities for labs and insurers is beyond the scope of this paper. However, it seems natural to analyze the use of exclusion clauses. First, the government intervention is a form of exclusion and now we allow the market parties to use a similar clause in their contracts. Second, health insurance contracts have exclusion clauses in them by specifying which treatments are and which are not insured. As a court must be able to judge the insurance contract in case of dispute between insurer and insured, the exlusion clauses in health insurance contracts need to be contractible. We now allow insurers and labs to contract on the exclusion clauses in insurance contracts. In particular, we are interested in the case where an insurer's contract with a h-lab specifies that the insurer cannot cover l-treatments. To illustrate, if treatment 1 is h and 2 is l, then the contract between insurers and L_1 states that insurers cannot cover treatment 2.

[27] Assuming that μ is small enough that we can ignore corner solutions.

To analyse this, we consider the case where the government outcome clearly dominates the market outcome. Assume that $F(q^l) = 1$ while we maintain (5). Further, the signal is uninformative ($g = \frac{1}{2}$). Then the government intervention implements first best ($W^g = W^*$) while $W^m < W^g$. Can the use of exclusion contracts by market participants bring W^m close to W^g?

As we argue below, it seems reasonable to assume that labs and insurers contract after the treatments have been invented. Then exclusion is only useful/necessary in case $\bar{q}_1 = \bar{q}_2 = (q^h + q^l)/2$: there is one l-treatment and one h-treatment. Because consumers cannot observe individual quality levels q_j, they need to form beliefs. To get an effect of exclusion, we consider the following beliefs: if only one treatment is covered in the health insurance *market*, then the h-treatment is covered by insurance.[28] We check whether the combination of exclusion clauses and this belief structure can raise dynamic efficiency in the market outcome (without government intervention).

Three remarks. First, this equilibrium is clearly not unique. If consumers' beliefs would ignore the number of covered treatments, we would be back in the analysis of section 4 and exclusion would have no effect. Second, this inference by consumers is already quite sophisticated. Perhaps, this is more than one can expect from a market. Three, we do not derive conditions under which such an equilibrium exists. We derive bounds on the equilibrium payoffs, if an equilibrium exists. Our point is: even if we assume that this can be done in a market, it is not as effective as the government setting $\underline{q} = q^l$.

The following result shows that exclusion contracts cannot achieve first best. Intuitively, the problem is that h-labs need to compensate insurers for excluding low quality treatments. This compensation to insurers reduces the profits that h-labs can earn below μq^h. As l-labs can always earn μq^l on the uninsured market ($F(q^l) = 0$ by assumption), the incentive for research effort is below the first best incentive $\mu(q^h - q^l)$ (that the government can achieve by setting $\underline{q} = q^l$).

Proposition 6 *Assume players can use exclusion clauses. Consider the case where* $q_1 = q^l, q_2 = q^h$ *and* $\bar{q}_1 = \bar{q}_2 = \bar{q} = \frac{q^h+q^l}{2}$. *Assume that if only one treatment is covered in the health insurance market, consumers believe that only the h-treatment is covered by insurance.*

If an equilibrium exists then $\pi_h \leq \mu \max\{\bar{q}, (1 - F(p_h^m))p_h^m\}$ *in the offer game and* $\pi_h \leq \mu(1 - F(p_h^m))p_h^m + \varepsilon$ *in the bidding game.*

As $F(q^h) > 0$, we see that $\pi_h < \mu q^h$ in the offer game. If we strengthen assumption (5) slightly to $F(q^h - \varepsilon) > 0$ we also have $\pi_h < \mu q^h$ in the bidding game. Since $\pi_l \geq \mu q^l$, the use of exclusion contracts does not lead to first best dynamic efficiency as $\pi_h - \pi_l < \mu(q^h - q^l)$.

The use of exclusion clauses does help insurers to raise profits. Indeed, the h-lab needs to convince insurers to exclusively cover h-treatments. An insurer can deviate and cover the l-treatment at a price $\tilde{p}_l$ close to q^l. As consumers then value each covered treatment at $\bar{q} > q^l$ (because two treatments are covered in the health insurance market), such a deviation is profitable for an insurer. The h-lab needs to leave insurers some rent ($t_h < 0$) to prevent this deviation from being profitable. This is consistent with the idea in the health economics literature that formularies are used by insurers to raise their profits.[29]

[28] Just to be clear: the beliefs do not say "if an insurer covers only one treatment, it covers the h-treatment". Such beliefs cannot be part of an equilibrium.

[29] See, for instance, McGuire (2011b, section 3.2) and references therein.

Above, we assume that parties contract at stage 3 using exclusion clauses. If we would assume that market parties can come together in stage 0, use Coasian bargaining and commit to a policy, they could implement $\underline{q} = q^l$. The contracts signed at stage 0 would then specify that insurers at stage 4 will not cover l-treatments. By raising research effort e closer to the efficient outcome, total welfare goes up and with $F(q^l) = 0$, no surplus is left to consumers. This implies that exclusion clauses with $\underline{q} = q^l$ increase total profits and hence is desirable for both labs and insurers at stage 0.

For several reasons, it is unlikely that the market can contract at stage 0 and achieve the first best outcome. First, the number of parties that bargain in period 0 may be quite large. Certainly, when taking into account that many research projects will not lead to a safe and effective treatment.[30] As it is not known ex ante which labs will be successful, all of them need to be present in period 0. The Coase theorem is unlikely to hold with that many participating parties.

Second, at time 0 parties may be cash constraint. It is usually hard to get money from the capital market to finance high risk R&D projects. Therefore, labs may not be able to compensate insurers at stage 0 for foregoing coverage of l-treatments. Hence, labs may be forced to postpone compensating payments till after the innovation. But then the compensation scheme will affect research incentives making it hard to achieve first best.

Third, it is not likely that parties can credibly commit to their strategy. Again considering the pharmaceutical market, it takes about 12 years from the idea of a new drug to actually selling it (see Danzon 2011, pp. 524, Scherer 2000, pp. 1308 and Weisbrod 1991, pp. 535). Few firms can make a commitment over such a long period. Although one might argue that the big pharmaceutical companies are long run players, many of the drugs that they sell were actually invented by small start up companies.[31] And these small companies need to commit as well. Indeed, if only the big pharmaceutical companies commit to $\underline{q} = q^l$, the start up company that invented a l-treatment will be tempted to sell to insurers directly (instead of selling out to a big pharma company) to benefit from $\bar{q} > q^l$. Also in the time frame of a decade, new insurance companies can enter the market that were not present at stage 0.

Therefore, we assume that parties in the market can only sign contracts after the treatments have been invented. That is, in stage 3 of the game presented in section 3.2.

6.2. insurer reputation

Another market based solution is that insurers build a reputation for only covering h-treatments. Consumers then value each treatment covered by this insurer at q^h.

This reputation mechanism can improve the market outcome. However, it is not as effective as government intervention. Consider an insurer with a reputation for covering only h-treatments. This insurer can deviate and include a l-treatment for coverage. Such a deviation is profitable in the short run: consumers value the additional treatment as a h-treatment while the l-lab is willing to sell at a low price. This short run gain needs to be weighed against the long run loss of reputation; say, at some point consumers may find out that they have been

[30] Danzon (2011, pp. 524) and Morton and Kyle (2011, pp. 773) note that for the pharmaceutical sector there is a less than 5% probability that pre-clinical research leads to a product that can be sold on the market.

[31] See Morton and Kyle (2011, pp. 775) on the movement of innovative activity outside large vertically integrated pharmaceutical firms. This is sometimes referred to as the shift to vertical specialization. Also see Danzon (2011, pp. 522).

cheated and switch to another insurer. This is only a loss, if the insurer with the h-reputation makes strictly positive profits. This happens if the h-labs leave a strictly positive rent to the insurer.

Hence, as in the previous subsection, we find that $\pi_h < \mu q^h$. If $\pi_l \geq \mu q^l$ the incentive for labs to invest in h-treatments is lower than first best: $\pi_h - \pi_l < \mu(q^h - q^l)$.

6.3. insurance contracts signal quality

Above we assume that insurers know the value added of each new treatment. It is natural to wonder whether insurers can signal the quality of a treatment by the conditions under which the treatment is reimbursed. For instance, could a high co-payment on a new treatment signal that the quality is high? At the moment that a patient needs a treatment, she is advised by a physician. The physician could then tell the patient that the value of the treatment is lower than the co-payment.

The problem here is that an insurer covering a low quality treatment can always mimic such behaviour. If a high co-payment would signal high quality, an insurer covering a low quality treatment can set a high co-payment as well. If the patient then decides not to use the treatment because of the co-payment, even better from the insurer's point of view.

Hence it is hard to see how an insurer can signal quality in the one-shot version of the game. In a repeated setting, reputation can play a role, as discussed above.

7. Extensions

This section discusses some extensions of the basic model.

7.1. more than two treatments

There are two ways in which we can consider more than two treatments in this model. First, we can consider more than one subset of the partition of all treatments. As mentioned, utility in equation (3) is linear. In the absence of co-morbidity, overall utility is the sum of utility for each subset and hence the analysis above applies.

Second, we can allow for more than two treatments in a subset. In this case, we also find that the government intervention excluding l-treatments raises research for breakthrough treatments. See the webappendix mentioned in footnote 13 for the analysis of this case.

7.2. co-morbidity

Above, we assumed that a patient could have one condition only and hence needed at most one treatment. If some patients need both treatments (say, with probability μ^2), this changes the profit maximizing price on the uninsured market. Above, with prices p_j^u $(j = 1, 2)$, lab j sells $\mu(1 - F(p_j^u))$ treatments on the uninsured market. Allowing for co-morbidity and assume –for simplicity– that everyone buys treatment 1, then treatment 2 is sold to $\mu(1-\mu)(1 - F(p_2^u)) + \mu^2(1 - F(p_1^u + p_2^u))$ uninsured patients. This interaction between research labs on the uninsured market is left for future research. For μ small, μ^2 is close enough to 0 that firms can ignore the co-morbidity term in their demand function on the uninsured market. Then the analysis above applies.

7.3. positive production costs

Above we normalized the marginal production costs c to zero. Assume that $c^h, c^l > 0$. Now it can happen that $q^h - c^h > q^l - c^l > 0 > \bar{q}_h - c^h$. In the market equilibrium, insurers will not cover the h-treatment as labs want a price that covers marginal production costs $\tilde{p}^h \geq c^h$. This can explain why medical innovations with high value added diffuse rather slowly (Phelps, 2000). As above, the government can exclude l-treatments from coverage. However, this is costly if $q^l - c^l$ is quite high and many people cannot afford to buy the l-treatment on the uninsured market.

An alternative is that the government mandates coverage of the h-treatment to correct consumer misperception. For instance, Kleinke (2001) suggests mandated coverage of certain drugs at the federal level.

7.4. segmented insurance market

A disadvantage of the government intervention described above is that people with low income cannot afford l-treatments as they are not allowed to be covered by health insurance. This is a loss in static efficiency because l-treatments still create value from a social point of view. Hence, one may wonder whether the government cannot correct consumers' misperceptions of treatments' value in a less costly way.

For instance, the government could publish the results of its CE analysis on a website. Consumers would then know the value of the treatments covered by health insurers. If an insurer would drop treatment j from its coverage, consumers would –after checking the CE website– reduce their valuation of this contract by μq_j. This would make sure that in equilibrium research labs receive their social value as profits.

The model above cannot address this government policy. The model is not explicit about why consumers do not know $q_{1,2}$ exactly when buying insurance. Presumably more is needed than a spreadsheet with CE results to make sure that consumers can make informed choices about the value of treatments. Medicine is complicated and a degree takes a couple of years. So it seems unlikely that just providing information will resolve this issue.

Another option is that the government segments the health insurance market. For instance, there could be mandatory basic insurance that only covers h-treatments. As in countries like Denmark, the Netherlands and Switzerland (Thomson et al., 2012), the government determines which treatments are covered for this insurance. Then there can also be a supplementary insurance market which covers other treatments. Government can then announce that each treatment covered by mandatory insurance is a h-treatment. Consumers can get coverage of l-treatments on the supplementary market.

If indeed consumers can make a clear distinction between mandatory and supplementary insurance, this two tiered insurance market can help to resolve consumer ignorance about the value added of treatments; while allowing for insurance coverage of l-treatments.

7.5. other forms of consumer ignorance

The paper is about consumers' misperception of the value of treatments, at the moment they buy health insurance. By using an imperfect signal of the value of treatments, they overvalue some treatments but are right on average. We cited evidence consistent with both these features.

However, there are other mistakes that consumers can make that are similar in nature. We discuss three of these.

First, a certain treatment may have high value for condition A but low (or no) value for condition B. Skinner (2011) and Chandra and Skinner (2012) call this Category II treatments and give examples like stents and back surgery. If the social value of a treatment is negative for condition B, it is efficient if the treatment is excluded for this condition. However, if the consumer does not know this variation in value (Rothberg et al., 2010, pp. 311), an insurer trying to limit the use of the treatment[32] may lose customers. In the spirit of the model above, this can be formalized as follows. The cost of the treatment is denoted c (assumed to be the same for both conditions), utility of the treatment is denoted q^i for condition $i = A, B$ with $q^A > c > q^B$. If the consumer is ignorant of the distinction between A and B and gets an uninformative signal ($g = \frac{1}{2}$), she values the treatment at $\bar{q}_A = \bar{q}_B = \bar{q} = (q^A + q^B)/2$. We assume that $\bar{q} > c$. If an insurer excludes the treatment for condition B, the consumer's reduction in valuation for this insurance contract is given by $\mu\bar{q} > \mu q^B$. Hence, the market cannot resolve this issue for the same reason as above. Efficiency would improve if a government agency excludes the treatment for people suffering from B. Currently, the FDA does not try to do this: "Once a drug is approved for treating one condition, doctors can legally use it for other things" (Deyo and Patrick, 2005, pp. 156).

Second, consumers may not value the information generated by a clinical trial testing a new treatment. This makes it hard for an insurer to insist on such trials (Garber 2001, pp. 79, Deyo and Patrick 2005, pp. 57 and Chandra and Skinner 2012, pp. 667). Suppose a new treatment has been invented with a benefit that can either be high q^h or low q^l with equal probability and expected benefit $\bar{q} = (q^h + q^l)/2$. Marginal production costs are denoted c. Assume that $q^h > \bar{q} > c > q^l$. There is social value in an experiment determining whether the treatment's value is high or low. However, an insurer requiring that its customers participate in a medical trial if the value of the treatment is not known, is likely to lose customers. Indeed, patients prefer to get $\bar{q} > c$ rather than run the risk of being treated with a placebo. Deyo and Patrick (2005, chapter 4) give a number of US examples of new treatments that were used under market pressure without being tested in randomized trials first. In the words of Skinner (2011) and Chandra and Skinner (2012) these are called Category III treatments. In line with the analysis above, Chandra and Skinner (2012, pp. 670) find that the "US often experiences rapid diffusion of expensive pharmaceutical treatments with uncertain benefits". The model suggests that the market cannot solve this by itself and government intervention is needed.

Finally, consumers may not know the substitution patterns between treatments. Suppose that there is a (new) treatment which creates value q^n; consumers correctly perceive this value. However, there is an existing treatment that creates value q^o when treating the same condition. If consumers are not aware of treatment o, they will reduce their valuation of an insurance contract that does not cover n but does cover o by μq^n, instead of $\mu(q^n - q^o)$. This increases the bargaining power of the research lab that invented n. Hence, even if $q^n - q^o$ is small, this

[32]It is not clear that this is always possible. However, the point is that if it would be possible –e.g. by explicitly specifying that treatment can only be used for condition A or by requiring a second opinion before paying for the treatment– it may not be profitable for an insurer to impose such restrictions. Chandra and Skinner (2012, pp. 646) give the example of stents which have substantial benefits only if the patient is treated within the first 12 hours following a heart attack. This is a contractible condition that could be used by insurers.

lab may be able to charge insurers a high price for its new treatment.

8. Conclusion

This paper starts from the premise that consumers do not perfectly know the value of each treatment at the moment that they buy health insurance. The question is whether health insurers can correct consumers' misperceptions such that treatments earn their social value as a profit.

We show in a model that captures salient features of the health insurance market –like consumers buy insurance to guarantee access to care and insurers contract privately on the treatment procurement market– that treatments with low value added can get overpaid in equilibrium. This is consistent with examples of treatments that hardly add value compared to existing treatments, but are sold at a high price (Deyo and Patrick, 2005; Light and Lexchin, 2012). This free riding of low quality treatments on high value innovations reduces dynamic efficiency.

The government can raise research efforts to find breakthrough treatments by abolishing treatments with low value added from health insurance coverage. As low quality treatments are no longer covered by health insurance, people with low income have no longer access to these treatments. This reduces static efficiency.

If the country is rich enough, such that most people can buy low quality treatments on the uninsured market at their social value, government intervention raises welfare. Hence, even in a private insurance market there is a role for a government institute like NICE to determine the treatments that can be covered by health insurance.

We argued above that the market cannot correct consumers' misperception about the value of new treatments. However, it is not obvious that the government is always able to make such a correction. For instance, in the 1980/90s the state of Oregon tried to ration Medicaid services using CE analysis. See Dranove (2003, chapter 7) for a discussion of the public pressure on the government and its consequences. It seems fair to say that politicians find it hard to resist such public pressure. One reason for the relative success of NICE can be that it is placed at arm's length from the political arena. This makes it somewhat easier to resist public demands for coverage of a new treatment.[33] Hence, we make the case for the government to establish an institute that determines which treatments can be covered by insurance, even in a private health insurance market. As with central banks, it may be a good idea that this institute enjoys independence from direct government intervention. The optimal governance structure of such an institute is left for future research.

References

Anirban Basu and Tomas J. Philipson. The impact of comparative effectiveness research on health and health care spending. Working Paper 15633, National Bureau of Economic Research, January 2010. URL http://www.nber.org/papers/w15633.

[33] See, for instance,
http://www.telegraph.co.uk/health/healthnews/7548038/Life-saving-cancer-drugs-still-not-available-on-NHS

Jan Boone. Optimal coverage in basic and supplementary health insurance. Technical report, Mimeo, 2013.

CBO. Research on the comparative effectiveness of medical treatments. Technical report, CBO, December 2007.

Amitabh Chandra and Jonathan S. Skinner. Technology growth and expenditure growth in health care. *Journal of Economic Literature*, 50(3):645–80, July 2012. doi: 10.1257/jel.50.3.645. URL http://www.aeaweb.org/articles.php?doi=10.1257/jel.50.3.645.

Amitabh Chandra, Anupam B. Jena, and Jonathan S. Skinner. The pragmatist's guide to comparative effectiveness research. *The Journal of Economic Perspectives*, 25(2):27–46, 2011. doi: doi:10.1257/jep.25.2.27. URL http://www.ingentaconnect.com/content/aea/jep/2011/00000025/00000002/art00002.

Michael E. Chernew and Dustin May. Chapter 14 - health care cost growth. In S. Glied and P. Smith, editors, *Oxford Handbook of Health Economics*, Oxford Handbook of Health Economics, pages 308 – 328. Oxford University Press, 2011.

Michael E. Chernew and Joseph P. Newhouse. Chapter one - health care spending growth. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 1 – 43. Elsevier, 2011. doi: 10.1016/B978-0-444-53592-4.00001-3. URL http://www.sciencedirect.com/science/article/pii/B9780444535924000013.

P.Y. Cremieux, D. Jarvinen, G. Long, and P. Merrigan. Pharmaceutical spending and health outcomes. Pharmaceutical Innovation: incentives, competition, and cost-benefit analysis in international perspective, chapter 12, pages 226–241. Cambridge University Press, 2007.

D. Cutler. The lifetime costs and benefits of medical technology. *Journal of Health Economics*, 26:1081–1100, 2007.

D. Cutler and M. McClellan. Is technological change in medicine worth it? *Health Affairs*, XX: 11–29, 2001.

D. Cutler, A. Rosen, and S. Vijan. The value of medical spending in the united states, 1960-2000. *New England Journal of Medicine*, 355(9):920–927, 2006.

Patricia M. Danzon. Chapter 12 - the economics of the biopharmaceutical industry. In S. Glied and P. Smith, editors, *Oxford Handbook of Health Economics*, Oxford Handbook of Health Economics, pages 520–554. Oxford University Press, 2011.

Patricia M. Danzon and Mark V. Pauly. Insurance and new technology: From hospital to drugstore. *Health Affairs*, 20(5):86–100, 2001. doi: 10.1377/hlthaff.20.5.86. URL http://content.healthaffairs.org/content/20/5/86.abstract.

R.A. Deyo and D.L. Patrick. *Hope or hype: the obsession with medical advances and the high cost of false promises*. Amacom, 2005.

R.A. Deyo, A. Nachemson, and S.K. Mirza. Spinal fusion surgery: The case for restraint. *New England Journal of Medicine*, 350(7):722–726, February 2004.

Peter Diamond. Organizing the health insurance market. *Econometrica*, 60(6):1233–1254, 1992. ISSN 00129682. URL http://www.jstor.org/stable/2951520.

D. Dranove. *The economic evolution of American health care: from Marcus Welby to managed care*. Princeton University Press, 2000.

D. Dranove. *What's your life worth? Health care rationing... who lives? who dies? and who decides?* Prentice Hall, 2003.

M. Drummond. Using economic evaluation in reimbursement decisions for health technologies: Lessons from international experience. Pharmaceutical Innovation: incentives, competition, and cost-benefit analysis in international perspective, chapter 11, pages 215–225. Cambridge University Press, 2007.

M. Drummond, M. Sculpher, G. Torrance, B. O'Brien, and G. Stoddart. *Methods for the economic evaluation of health care programmes*. Oxford University Press, third edition edition, 2005.

U Dulleck and R Kerschbamer. On doctors, mechanics and computer specialists: The economics of credence goods. *Journal of Economic Literature*, 44(1):5–42, 2006.

M.A. Fischer and J. Avorn. Economic implications of evidence-based prescribing for hypertension. *JAMA*, 291(15):1850–1856, April 2004.

A. Garber, C. Jones, and P. Romer. Insurance and incentives for medical innovation. *Forum for Health Economics & policy*, 9(2):Article 4, 2006.

Alan M. Garber. Evidence-based coverage policy. *Health Affairs*, 20(5):62–82, 2001. doi: 10.1377/hlthaff.20.5.62. URL http://content.healthaffairs.org/content/20/5/62.abstract.

Alan M. Garber and Mark J. Sculpher. Chapter eight - cost effectiveness and payment policy. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 471 – 497. Elsevier, 2011. doi: 10.1016/B978-0-444-53592-4.00008-6. URL http://www.sciencedirect.com/science/article/pii/B9780444535924000086.

A.M. Garber. Advances in cost-effectiveness analysis of health interventions. volume 1 of *Handbook of Health Economics*, chapter 4, pages 181–221. Elsevier, 2000.

M. Gaynor, D. Haas-Wilson, and W. Vogt. Are invisible hands good hands? moral hazard, competition, and the second-best in health care markets. *Journal of Political Economy*, 108 (5):992–1005, 2000.

D.P. Goldman, G.F. Joyce, and Y. Zheng. Prescription drug cost sharing: associations with medication and medical utilization and spending and health. *The Journal of the American Medical Association (JAMA)*, 298(1):61–69, 2007.

Said B. Habib, Luke Sonoda, Teik C. See, Peter J. Ell, and Ashley M. Groves. How do patients perceive the benefits and risks of peripheral angioplasty? implications for informed consent. *Journal of Vascular and Interventional Radiology*, 19(2):177–181, 2008. URL http://www.sciencedirect.com/science/article/pii/S1051044307013747.

Elke Jakubowski and Reinhard Busse. Health care systems in the eu: A comparative study. Working Paper SACO 101, European Parliament, May 1998.

A.B. Jena and T.J. Philipson. Cost-effectiveness analysis and innovation. *Journal of Health Economics*, 27(5):1224 – 1236, 2008. ISSN 0167-6296. doi: 10.1016/j.jhealeco.2008.05.010. URL http://www.sciencedirect.com/science/article/pii/S0167629608000726.

Anupam Jena and Tomas Philipson. Endogenous cost-effectiveness analysis and health care technology adoption. *Journal of Health Economics*, 32:172–180, 2013.

K. Kelly, J. Crowley, and P.A. Bunn et al. Randomized phase iii trial of paclitaxel plus carboplatin versus vinorelbine plus cisplatin in the treatment of patients with advanced non-small-cell lung cancer: A southwest oncology group trial. *Journal of Clinical Oncology*, 19 (13):3210–3218, July 2001.

J.D. Kleinke. The price of progress: Prescription drugs in the health care market. *Health Affairs*, 20(5):43–60, 2001. doi: 10.1377/hlthaff.20.5.43. URL http://content.healthaffairs.org/content/20/5/43.abstract.

D. Lakdawalla and N. Sood. Innovation and the welfare effects of public drug insurance. *Journal of Public Economics*, 93(3-4):541–548, 2009. ISSN 0047-2727. doi: 10.1016/j.jpubeco.2008.11.003. URL http://www.sciencedirect.com/science/article/pii/S0047272708001783.

Jeffrey Liebman and Richard Zeckhauser. Simple humans, complex insurance, subtle subsidies. Working Paper 14330, National Bureau of Economic Research, September 2008. URL http://www.nber.org/papers/w14330.

Donald W Light and Joel R Lexchin. Pharmaceutical research and development: what do we get for all that money? *BMJ*, 345, 8 2012. doi: 10.1136/bmj.e4348.

G. Mankiw and M. Whinston. Free entry and social inefficiency. *Rand Journal of Economics*, 17:48–58, 1986.

P. McAfee and M. Schwartz. Opportunism in multilateral vertical contracting: nondiscrimination, exclusivity, and uniformity. *American Economic Review*, 84:210–230, 1994.

J.E. McDonough. *Inside national health reform*. University of California Press, 2011.

Thomas G. McGuire. Chapter 25 - physician agency and payment for primary medical care. In S. Glied and P. Smith, editors, *Oxford Handbook of Health Economics*, Oxford Handbook of Health Economics, pages 602–623. Oxford University Press, 2011a.

Thomas G. McGuire. Chapter five - demand for health insurance. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 317 – 396. Elsevier, 2011b. doi: 10.1016/B978-0-444-53592-4.00005-0. URL http://www.sciencedirect.com/science/article/pii/B9780444535924000050.

M.M. Mello and T.A. Brennan. The controversy over high-dose chemotherapy with autologous bone marrow transplant for breast cancer. *Health Affairs*, 20(5):101–117, September/October 2001.

David O. Meltzer and Peter C. Smith. Chapter seven - theoretical issues relevant to the economic evaluation of health technologies. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 433 – 469. Elsevier, 2011. doi: 10.1016/B978-0-444-53592-4.00007-4. URL http://www.sciencedirect.com/science/article/pii/B9780444535924000074.

Fiona Scott Morton and Margaret Kyle. Chapter twelve - markets for pharmaceutical products. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 763 – 823. Elsevier, 2011. doi: 10.1016/B978-0-444-53592-4.00012-8. URL http://www.sciencedirect.com/science/article/pii/B9780444535924000128.

NICE. Briefing paper for the methods working party on the cost effectiveness threshold. Technical report, NICE, 2007.

John A. Nyman. The value of health insurance: the access motive. *Journal of Health Economics*, 18(2):141 – 152, 1999. ISSN 0167-6296. doi: 10.1016/S0167-6296(98)00049-6. URL http://www.sciencedirect.com/science/article/pii/S0167629698000496.

Jonathan Oberlander. Health reform interrupted: The unraveling of the oregon health plan. *Health Affairs*, 26(1):w96–w105, 2007. doi: 10.1377/hlthaff.26.1.w96. URL http://content.healthaffairs.org/content/26/1/w96.abstract.

M.V. Pauly. Measures of costs and benefits for drugs in cost-effectiveness analysis. Pharmaceutical Innovation: incentives, competition, and cost-benefit analysis in international perspective, chapter 10, pages 199–214. Cambridge University Press, 2007.

C.E. Phelps. Information diffusion and best practice adoption. volume 1 of *Handbook of Health Economics*, chapter 5, pages 223–264. Elsevier, 2000.

J.D. Piette, M. Heisler, and T.H. Wagner. Cost-related medication underuse among chronically ill adults: the treatments people forgo, how often, and who is at risk. *American Journal of Public Health*, 94(10):1782–1787, 2004a.

J.D. Piette, M. Heisler, and T.H. Wagner. Problems paying out-of-pocket medication costs among older adults with diabetes. *Diabetes Care*, 27(2):384–391, 2004b.

P.M. Ravdin, I.A. Siminoff, and J.A. Harvey. Survey of breast cancer patients concerning their knowledge and expectations of adjuvant therapy. *Journal of Clinical Oncology*, 16(2): 515–521, Februay, 1 1998.

Patrick Rey and Thibaud Vergé. Bilateral control with vertical contracts. *RAND Journal of Economics*, 35(4):728–746, Winter 2004. URL http://ideas.repec.org/a/rje/randje/v35y20044p728-746.html.

Michael B. Rothberg, Senthil K. Sivalingam, Javed Ashraf, Paul Visintainer, John Joelson, Reva Kleppel, Neelima Vallurupalli, and Marc J. Schweiger. Patients' and cardiologists' perceptions of the benefits of percutaneous coronary intervention for stable coronary disease. *Annals of Internal Medicine*, 153(5):307–313, 2010. doi: 10.7326/0003-4819-153-5-201009070-00005. URL +http://dx.doi.org/10.7326/0003-4819-153-5-201009070-00005.

F.M. Scherer. Chapter 25 the pharmaceutical industry. volume 1, Part B of *Handbook of Health Economics*, pages 1297 – 1336. Elsevier, 2000. doi: 10.1016/S1574-0064(00)80038-4. URL http://www.sciencedirect.com/science/article/pii/S1574006400800384.

C Schoen, R Osborn, D Squires, M M Doty, R Pierson, and S Applebaum. How Health Insurance Design Affects Access To Care And Costs, By Income, In Eleven Countries. *Health Affairs*, 29(12):1–12, 2010. ISSN 02782715. doi: 10.1377/hlthaff.2010.0862. URL http://content.healthaffairs.org/cgi/content/abstract/hlthaff.2010.0862v1.

Cathy Schoen, Sara R Collins, Jennifer L Kriss, and Michelle M Doty. How many are underinsured? trends among u.s. adults, 2003 and 2007. *Health affairs (Project Hope)*, 27(4):298–309, 2008. ISSN 1544-5208. doi: 10.1377/hlthaff.27.4.w298. URL http://www.ncbi.nlm.nih.gov/pubmed/18544591.

I. Segal. Contracting with externalities. *Quarterly Journal of Economics*, 114:337–388, 1999.

I. Segal and M.D. Whinston. Robust predictions for bilateral contracting with externalities. *Econometrica*, 71:757–791, 2003.

Corey A. Siegel, L. Campbell Levy, Todd A. MacKenzie, and Bruce E. Sands. Patient perceptions of the risks and benefits of infliximab for the treatment of inflammatory bowel disease. *Inflammatory Bowel Diseases*, 14(1):1–6, 2008.

Jonathan Skinner. Chapter two - causes and consequences of regional variations in health care. In Thomas G. Mcguire Mark V. Pauly and Pedro P. Barros, editors, *Handbook of Health Economics*, volume 2 of *Handbook of Health Economics*, pages 45 – 93. Elsevier, 2011. doi: 10.1016/B978-0-444-53592-4.00002-5. URL http://www.sciencedirect.com/science/article/pii/B9780444535924000025.

S. Thomson, R. Osborn, D. Squires, and M. Jun. *International profiles of health care systems*. Commonwealth Fund, 2012.

Kevin Volpp, George Loewenstein, Andrea Troxel, Jalpa Doshi, Maureen Price, Mitchell Laskin, and Stephen Kimmel. A test of financial incentives to improve warfarin adherence. *BMC*

Health Services Research, 8(1):272, 2008. ISSN 1472-6963. doi: 10.1186/1472-6963-8-272. URL http://www.biomedcentral.com/1472-6963/8/272.

B. Weisbrod. The health care quadrilemma: an essay on technological change, insurance, quality of care, and cost containment. *Journal of Economic Literature*, XXIX:523–552, 1991.

A. Proof of results

Proof of proposition 1

Offer game: First, we will argue that in a symmetric perfect Bayesian equilibrium it is the case that $t_j = 0$.

Let A_j denote the set of offers $(\tilde{p}_j, t_j)$ that an insurer is willing to accept and offer coverage of treatment j. Suppose –by contradiction– that there exists $(\tilde{p}_j, t_j) \in A_j$ with $t_j > 0$. Let $\bar{t}_j > 0$ denote the largest such $t_j > 0$ in A_j. Note that such a largest element exists because the total profit on the market cannot exceed μq^h (and hence an insurer never accepts $t_j > \mu q^h$) and because of the money unit ε there are only a countable number of candidate elements of A_j.

As the contracts offered are private information and we consider the case of interim unobservability, an insurer's price setting strategy in the health insurance market only depends on the details of its own contracts. Given the offer from the other lab $-j$ (which the insurer may or may not accept), an insurer raises its premium σ by $\Delta\sigma_j(\tilde{p}_j, t_j)$ if it offers coverage of treatment j.

Now define $\bar{p}_j = \arg\max_p\{\Delta\sigma_j(p, \bar{t}_j)|(p, \bar{t}_j) \in A_j\}$. This is the price $\tilde{p}_j$ that leads to the biggest increase in σ given that j is covered and given $t_j = \bar{t}_j$.

Given the offer $(\bar{p}_j, \bar{t}_j)$ to insurer I_a which I_a accepts and offers coverage of j at premium increase $\Delta\sigma_j(\bar{p}_j, \bar{t}_j)$, two situations can occur for I_a:

1. I_a expects a market share, s_a, strictly below 1. As there are two or more insurers, this implies $s_a \leq 1/2$. This means that the margin that I_a earns is at most 2ε (if it would be 3ε or more, it is optimal for I_a to reduce σ by ε and capture the whole market). Hence $\pi_a \leq \mu s_a 2\varepsilon - \bar{t}_j < 0$ because $\bar{t}_j \geq \varepsilon$ (as ε is the smallest money unit that can be used), $\mu < 1$ and $s_a \leq 1/2$.

2. I_a expects a market share equal to 1. Although the offer $\bar{p}_j, \bar{t}_j$ can be part of a deviation by L_j, we consider this deviation in the context of a pure strategy symmetric equilibrium. That is, the other lab, L_{-j}, makes the equilibrium offer $(\tilde{p}_{-j}, t_{-j})$ to all insurers who then set $\Delta\sigma_{-j}(\tilde{p}_{-j}, t_{-j})$.[34] Note that $\Delta\sigma_{-j} = 0$ if $(\tilde{p}_{-j}, t_{-j}) \notin A_{-j}$ and treatment $-j$ is not covered by health insurance. Then $s_a = 1$ together with the definition of $\bar{p}_j$ implies that other insurers received an offer with $t_j < \bar{t}_j$. However, since $(\bar{p}_j, \bar{t}_j) \in A_j$, the other insurers would have accepted offer $(\bar{p}_j, \bar{t}_j)$. As the offer with $\bar{t}_j$ strictly increases L_j's profits in this case, this is the offer that I_a expects L_j to make to the other insurers under wary beliefs. But this contradicts that I_a expects $s_a = 1$.

Hence in both cases we have a contradiction. Therefore $t_j \leq 0$.

Now we can verify whether the equilibrium in the proposition is indeed an equilibrium of the offer game. That is, there are no profitable deviations. We go over each of the possible deviations in turn. First, the case where $\bar{q}_2 > (1 - F(p_2^m))p_2^m$:

- L_j can deviate by offering $\hat{p}_j \geq \bar{q}_j + \varepsilon$. In this case, insurers receiving such a deviating offer do not offer coverage of treatment j. Indeed, to offer coverage in a profitable way requires $\Delta\sigma_j \geq \mu\hat{p}_j$ while consumers only value this coverage at $\mu\bar{q}_j < \mu\hat{p}_j$. Hence offering

[34] Deviating offers by L_j and L_{-j} cannot be coordinated as offers are made simultaneously and independently.

coverage of j in this case reduces an insurer's profits. Therefore the insurer will not offer coverage and the deviation reduces L_j's profits.

- L_j can decide not to make any offers to insurers and sell only on the uninsured market. Given the assumption here that $\mu\bar{q}_2 > \mu(1 - F(p_2^m))p_2^m$ this is not profitable for the lab with the highest q. As $q_1 \leq q_2, \bar{q}_1 \geq q_1, p_1^m \leq q_1$ such a deviation is not profitable for the other lab either.

- Offering $p_j^u < \bar{q}_j$ reduces profits as insurers will ask their customers to buy treatment j on the uninsured market (and reimburse them later).

- L_j can deviate by offering $\hat{p}_j \leq \bar{q}_j - \varepsilon$ to some (or all) insurers. As this cannot be accompanied with $t_j > 0$ (above we showed that offers with $t_j > 0$ are not acceptable to insurers), it cannot raise L_j's profits.

- Offering $t_j < 0$ clearly also does not increase L_j's profits.

- Given an offer $\tilde{p}_j = \bar{q}_j$ with $t_j = 0$, an insurer cannot gain by rejecting this offer. To see this, consider the case where I_a decided to drop j from coverage. To keep consumers indifferent between the old insurance package (with j) and the deviating package without j, I_a's premium σ has to fall by $\mu\bar{q}_j$. Hence I_a's revenues and costs fall by the same amount. Therefore this deviation is not profitable.

Next, consider equilibrium in case $\bar{q}_j < p_2^m(1 - F(p_2^m))$ (note that this cannot happen if $q_1 = q_2$). In this case, the insurance market unravels. As shown above, if both treatments would be covered, L_2 would earn $\mu\bar{q}_j$. But if L_2 decides not to sell to insurers and only sell on the uninsured market, it earns $\mu p_2^m(1 - F(p_2^m))$ which is higher by assumption. Hence, in this case, the insurance market unravels and treatment 2 is only sold on the uninsured market.

Since $\bar{q}_l \geq q^l \geq p_l^m(1 - F(p_l^m))$, l-lab has no incentive to sell on the uninsured market. Hence, if the signals are L, H and one lab sells on the uninsured market, consumer understands that insurance covers the l-treatment. No party can increase profits by deviating from the equilibrium.

Bidding game: First, consider the case with $\bar{q}_2 - \varepsilon > (1 - F(p_2^m))p_2^m$. We consider potential deviations and argue in each case that the deviation is not profitable:

- If L_2 rejects all offers from insurers, it can only sell on the uninsured market. The deviation profit then equals $\hat{\pi}_2 = \mu p_2^m(1 - F(p_2^m))$ which –by assumption– is lower than $\mu(\bar{q}_2 - \varepsilon)$. Hence this deviation is not profitable. As $p_1^m \leq q_1$ and $\bar{q}_1 \geq q_1$, the uninsured profit for L_1 cannot exceed π_1. Hence this deviation is not profitable for either lab.

- L_j can reject offers from some (but not all) insurers. This does not raise profits.

- Setting $\hat{p}_j^u < \bar{q}_j - \varepsilon$ reduces profits as insurers ask their customers to buy j on the uninsured market.

- Consider deviations for insurers. I_a sets $\hat{\sigma}_a > \mu(\bar{q}_1 + \bar{q}_2 - 2\varepsilon)$. Then all consumers strictly prefer to buy from other insurers. Setting $\hat{\sigma}_a < \mu(\bar{q}_1 + \bar{q}_2 - 2\varepsilon)$ leads to a loss. Hence deviations with σ cannot raise an insurer's profits.

- I_a drops treatment j from coverage. This allows I_a to reduce its premium with $\bar{q}_j - \varepsilon$. However, consumers' valuation of I_a's product falls by $\bar{q}_j$. Hence this is not a profitable deviation as revenue (σ) falls more than costs.

- Finally, consider deviations in I_a's offer to L_j. Note that L_j can reject the offer from I_a and accept the offers from other insurers who follow the equilibrium (as the other insurers do not know when setting σ that I_a's (deviating) offer was rejected). If I_a's offer is rejected, consumers strictly prefer health insurance from other insurers (since they offer coverage with additional perceived utility $\mu\bar{q}_j$ at an increased premium of $\mu(\bar{q}_j - \varepsilon)$) and do not buy from I_a. Hence L_j's profit after rejecting I_a's deviating offer equals $\mu(\bar{q}_j - \varepsilon)$. In other words, in response to a deviation by I_a, L_j can guarantee a profit equal to its equilibrium profit π_j. Hence a change in I_a's offer for t_j *only* cannot be profitable: $\hat{t}_j < 0$ is rejected by L_j and $\hat{t}_j > 0$ leads to a loss for I_a.

- Now, consider a deviating offer $(\hat{p}_j, \hat{t}_j)$. If $\hat{p}_j > \bar{q}_j$, I_a will raise σ_a by at least $\mu\hat{p}_j$. As consumers only value the inclusion of treatment j by $\mu\bar{q}_j$, this deviation does not raise profits.

- Next, consider $\hat{p}_j = \bar{q}_j$. If L_j would accept this deviation offer from I_a, she has two options. On the one hand, she can accept some of the offers made by the other insurers. As I_a has to raise his premium by $\mu\hat{p}_j$ to cover treatment j, while the other insurers raise their premium by $\mu(\bar{q}_j - \varepsilon) < \mu\hat{p}_j$, this is not a profitable deviation. On the other hand, she can reject all other offers. As I_a will raise σ by $\mu\hat{p}_j$ to cover treatment j, consumers are indifferent between I_a's product (with j covered) and the other insurers products (without j). Hence I_a's market share equals $1/n$ with $n \geq 2$ insurers in the market. The highest profit that L_j can earn in this case equals $\bar{q}_j/n$. However, the assumption on the smallest money unit ε (equation (4)) implies that $\mu\bar{q}_j/n < \pi_j$. Hence this is not a profitable deviation.

- Finally, consider $\hat{p}_j \leq \bar{q}_j - 2\varepsilon$. As above, if L_j accepts this offer, she can decide to accept (some of) the other offers or not. The same reasoning as before shows that this deviation is not profitable.

If $\bar{q}_2 - \varepsilon < (1 - F(p_2^m))p_2^m$, it is more profitable for L_2 to sell on the uninsured market. Hence insurance unravels and only q^l is covered in equilibrium. The reasoning here is the same as for the offer game. *Q.E.D.*

Proof of proposition 2

In the *offer game* the highest equilibrium profit that lab L_j can earn when the insurance market does not unravel equals $\mu\bar{q}_j$. As shown in the proof of proposition 1, if lab j makes an offer with $\tilde{p}_j > \bar{q}_j$ the insurer will drop j from its health insurance. To keep consumers indifferent it will reduce its premium σ by $\mu\bar{q}_j$. In this way, its costs decrease faster than its revenues and hence not covering j leads to higher profits for the insurer.

Does there exist an equilibrium in the offer game where $\pi_j \leq \mu(\bar{q}_j - 2\varepsilon)$? No, because L_j can always offer insurers $\tilde{p}_j = \bar{q}_j - \varepsilon$. Insurers are always willing to accept this offer (no matter what they believe about competitors), cover treatment j and increase σ by $\mu\bar{q}_j$. Since consumers are indifferent between insurance without coverage of j and insurance with coverage

of j that is $\mu\bar{q}_j$ more expensive in terms of its premium, the insurer can never lose accepting such an offer (and it may win).

To verify the claim for the *bidding game*, we check both inequalities:

1. $\pi_j \leq \mu(\bar{q}_j - \varepsilon)$. Suppose not, that is

   - either $\pi_j > \mu\bar{q}_j$: then there exists an insurer I_i with market share $s_i > 0$ such that $s_i\mu\tilde{p}^i_j + t^i_j > s_i\mu\bar{q}_j$. But then I_i has a profitable deviation: drop the contract with L_j and reduce σ_i by $\mu\bar{q}_j$. This deviation does not affect s_i but I_i's profits increase as I_i's costs decrease faster than its revenues;
   - or $\pi_j = \mu\bar{q}_j$: then there exists a set I of insurers such that $\sum_{i\in I} s_i\mu\tilde{p}^i_j + t^i_j = \mu\bar{q}_j$ and $\sum_{i\in I} s_i = 1$ (if this would not be the case, there exists $i \in I$ with $s_i\mu\tilde{p}^i_j + t^i_j > s_i\mu\bar{q}_j$; however, this is not possible as shown in the previous point). Further, it must be the case that $s_i\mu\tilde{p}^i_j + t^i_j = s_i\mu\bar{q}_j$ for each $i \in I$. Let $i^* = \arg\max_{i\in I} s_i$. It follows that $s_{i^*} \geq 1/n$. We show that insurer I_{i^*} has a profitable deviation. Suppose that I_{i^*} offers $\hat{p}_j = \bar{q}_j - \varepsilon, \hat{t}_j = 0$ to L_j. If L_j accepts this offer, indeed I_{i^*} has a profitable deviation and L_j earns (at least) $\hat{\pi}_j = \mu(\bar{q}_j - \varepsilon)$. We show that L_j cannot credibly reject I_{i^*}'s deviating offer. If L_j rejects the offer, its profits cannot exceed $(1 - s_{i^*})\mu\bar{q}_j \leq \frac{n-1}{n}\mu\bar{q}_j$. However, $\hat{\pi}_j > \frac{n-1}{n}\mu\bar{q}_j$ because of assumption (4) on the smallest money unit ε. Hence L_j cannot credibly reject I_{i^*}'s offer and hence I_{i^*} has a profitable deviation.

   Hence in each case there is a contradiction. Therefore $\pi_j \leq \mu(\bar{q}_j - \varepsilon)$.

2. $\pi_j \geq \mu(\bar{q}_j - 2\varepsilon)$. Suppose not, that is assume –by contradiction– that $\pi_j \leq \mu(\bar{q}_j - 3\varepsilon)$. We write $\pi_j = \mu(\bar{q}_j - y\varepsilon)$ with $y \geq 3$ and total insurer profits equal $\sum_{i=a}^{n} \pi_i = \mu x\varepsilon$ with $x \leq y$. We go over the following cases:

   - $x = 0$: I_a can deviate and offer to L_j: $\hat{p}_{aj} = \mu(\bar{q}_j - 2\varepsilon), \hat{t}_{aj} = 0$. L_j accepts this offer and rejects all other offers. Profit maximizing premium choice for I_a then leads to $\hat{s}_a = 1$. L_j's profits increase by accepting this deviating offer as $\mu(\bar{q}_j - 2\varepsilon) > \mu(\bar{q}_j - y\varepsilon)$ for $y \geq 3$. Further, $\hat{\pi}_a \geq \mu\varepsilon > 0 = \pi_a$.
   - $x = 1$: there exists at least one insurer (say, I_a) with $\pi_a \leq \mu\varepsilon/n$. Same deviation for I_a as under the previous point ($\hat{p}_{aj} = \mu(\bar{q}_j - 2\varepsilon)$) raises I_a's profits and is acceptable to L_j.
   - $x \geq 2$: There exists at least one insurer (say, I_a) with $\pi_a \leq x\mu\varepsilon/n$. If I_a deviates by cutting its premium by ε, its profits equal $\hat{\pi}_a = \mu(x-1)\varepsilon > \mu x\varepsilon/n \geq \pi_a$ (with strict inequality because $n \geq 3$).[35]

Q.E.D.

Proof of proposition 3 With $g \in \langle \frac{1}{2}, 1\rangle$ there can be three cases, depending on the signal (with $g = 1$ there is only one case and with $g = \frac{1}{2}$ there are two cases: cases 1 and 3 below).

[35] In principle, it could be the case that insurers earn profits via $t_j < 0$ (instead of the margin $\sigma - \mu\tilde{p}_j$). However, this clearly cannot be an equilibrium. First, suppose that all insurers have $t_j < 0$, then L_j should reject some of these offers and save on the transfer t_j to insurers. Second, if some insurer does not have $t_j < 0$, it can deviate by setting $\hat{p}_{aj} = \mu(\bar{q}_j - 2\varepsilon), \hat{t}_{aj} = 0$; as in the case of $x = 0$.

Case 1 $gq^l + (1-g)q^h > (1-F(p_h^m))p_h^m$: L_j solves

$$\max_e \mu\big(e_{-j}(eq^h + (1-e)((g^2+(1-g)^2)q^l + 2g(1-g)q^h)) + (1-e_{-j})(e((g^2+(1-g)^2)q^h + 2g(1-g)q^l) + (1-e)q^l)\big) - c(e)$$

The first order condition for e can be written as

$$c'(e) = \mu((g^2+(1-g)^2)q^h - (1-2g(1-g))q^l)$$

Hence we find for the market outcome in this case that

$$c'(e^m) = \mu(1-2g(1-g))(q^h - q^l) < \mu(q^h - q^l)$$

if $g \in [\frac{1}{2}, 1\rangle$ and thus $e^m < e^*$.

Case 2 $(1-g)q^l + gq^h > (1-F(p_h^m))p_h^m > gq^l + (1-g)q^h$: L_j solves

$$\max_e \mu\big(e_{-j}(eq^h + (1-e)(g(gq^l + (1-g)q^h) + (1-g)q^l)) + (1-e_{-j})(e(g(gq^h + (1-g)q^l) + (1-g)(1-F(p_h^m))p_h^m) + (1-e)q^l)\big) - c(e)$$

Hence equilibrium e^m is determined by

$$\begin{aligned} c'(e^m) &= \mu(g^2(q^h - q^l) + (1-g)((1-F(p_h^m))p_h^m - q^l) + e^m(1-g)(q^h - (1-F(p_h^m))p_h^m)) \\ &\leq \mu(g^2(q^h - q^l) + (1-g)(q^h - q^l)) \end{aligned}$$

because $e^m \leq 1$. Hence, we see that $g < 1$ implies that $e^m < e^*$.

Case 3 $(1-F(p_h^m))p_h^m > (1-g)q^l + gq^h \geq gq^l + (1-g)q^h$: L_j solves

$$\max_e e_{-j}\mu\left(eq^h + (1-e)q^l\right) + (1-e_{-j})\mu\left(e(1-F(p_h^m))p_h^m + (1-e)q^l\right) - c(e)$$

Equilibrium e^m is determined by

$$c'(e^m) = \mu((1-F(p_h^m))p_h^m - q^l + e^m(q^h - (1-F(p_h^m))p_h^m))$$

Since $(1-F(p_h^m))p_h^m < q^h, e^m < 1$, we find that $e^m < e^*$. *Q.E.D.*

Proof of proposition 4 With $\underline{q} = q^l$, l-treatment can only be sold on the uninsured market. Profits of an l-lab equal $\pi_l = \mu(1-F(p_l^m))p_l^m$.

With this government policy, consumers understand that insured treatments are h-treatments. Hence, the updated value of insured treatments equals $\bar{q}_j = q^h$. It then follows from proposition 1 that $\pi_h = \mu q^h$ (ignoring ε in the bidding game). *Q.E.D.*

Proof of proposition 6 Use of exclusion clauses is only relevant in the case with one l-treatment and one h-treatment. Allowing for such clauses in the *offer game*, the h-lab offers contracts to insurers with the clause that the insurer is not allowed to cover the l-treatment. Hence if an insurer –accepting such a contract– does cover the l-treatment, it is not possible to cover the h-treatment as its h-contract is void and the h-lab can charge $\tilde{p}_h = +\infty$.

Suppose that the claim in the proposition is not true. That is, assume –by contradiction– that the equilibrium profit for the h-lab satisfies:

$$\pi_h \geq \mu(\bar{q} + \varepsilon) \tag{A.1}$$

in the case where $\bar{q} > (1 - F(p_h^m))p_h^m$.

We claim that the equilibrium profit of every insurer I_i must satisfy $\pi_i \geq \mu(\bar{q} - q^l - \varepsilon)$. Suppose not. That is, assume that there is an insurer, say I_a, with $\pi_a < \mu(\bar{q} - q^l - \varepsilon)$. Then the l-lab can make I_a an offer $\hat{p}_l = q^l + \varepsilon, \hat{t}_l = 0$. This offer –if accepted– raises l-lab's profits above its profits on the uninsured market (μq^l). If I_a accepts and sells coverage of the l-treatment (only) at $\hat{\sigma}_a = \mu\bar{q}$ then all consumers buy from I_a because the other insurers sell at $\sigma > \mu\bar{q}$ (and given that both treatments are covered on the insurance market, consumers value coverage of a treatment at $\mu\bar{q}$). Hence I_a's profits equal $\hat{\pi}_a = \mu(\bar{q} - q^l - \varepsilon) > \pi_a$. Hence in equilibrium it must be the case for each insurer that $\pi_i \geq \mu(\bar{q} - q^l - \varepsilon)$.

It follows that the h-lab's equilibrium profit satisfies $\pi_h \leq \mu q^h - n\pi_i$. Combining this inequality with equation (A.1) and the inequality for π_i, we find

$$\mu(\bar{q} + \epsilon) \leq \pi_h \leq \mu(q^h - n(\frac{q^h - q^l}{2} - \varepsilon))$$

which can be rewritten as

$$\frac{q^h - q^l}{2} \leq \varepsilon$$

which contradicts the assumption on the smallest money unit in equation (4).

Now we turn to the *bidding game*, where insurers make offers to labs with exclusionary clauses. That is, insurer I_i offers the h-lab a contract that specifies that I_i will not sell coverage of the l-treatment. First, note that offers by insurers to the l-lab with $\tilde{p}_l = q^l, t_l = 0$ are always accepted by l-labs (even though insurers do not offer coverage of the l-treatment in an exclusive equilibrium). Now assume –by contradiction– that

$$\pi_h \geq \mu(1 - F(p_h^m))p_h^m + 2\varepsilon \tag{A.2}$$

This implies that the h-treatment is covered by health insurers and that they charge $\sigma > \mu q^l$ because $p_h^m \geq q^l$ (as $F(q^l) = 1$ by assumption).

Suppose that the insurer with the highest share in h-lab's profits (denoted I_a with $s_a \geq 1/n$) deviates and reduces t_a by ε. If h-lab rejects I_a's offer, I_a will offer coverage of l-treatment (with $\tilde{p}^l = q^l, t^l = 0$) and set a premium $\hat{\sigma}_a = \min\{\sigma, \mu\bar{q}\} > \mu q^l$ which implies that I_a's market share will not fall (because $\hat{\sigma}_a \leq \mu\bar{q}$ and $\hat{\sigma}_a \leq \sigma$) and increase if $\hat{\sigma}_a < \sigma$.

Because I_a was the biggest contributor to h-lab's profits and I_a does not lose market share, profit of the h-lab cannot exceed $\frac{n-1}{n}\pi_h$. This implies that it is not credible for h-lab to reject I_a's deviating offer because $\frac{n-1}{n}\pi_h < \pi_h - \varepsilon$ (by equation (4)). It follows that I_a has a profitable deviation. Therefore in equilibrium equation (A.2) cannot hold.

Q.E.D.