Health Insurance without Single Crossing: why healthy people have high coverage*

Jan Boone and Christoph Schottmüller†

January 16, 2014

Abstract

Standard insurance models predict that people with high (health) risks have high insurance coverage. It is empirically documented that people with high income have lower health risks and are better insured. We show that income differences between risk types lead to a violation of single crossing in an insurance model where people choose treatment intensity. We analyze different market structures in this setting and show the following: If insurers have some market power, the violation of single crossing caused by income differences can explain the empirically observed outcome. In contrast to other papers, our results do not rely on differences in risk aversion between types.

Keywords: health insurance, single crossing, competition

JEL classification: D82, I11

---

*We thank Cedric Argenton, Eric van Damme, Humberto Moreira, Florian Schütt and seminar participants at Tilburg University for useful comments. Financial support from the Dutch National Science Foundation (VICI 453.07.003) is gratefully acknowledged.

†Boone: Department of Economics, University of Tilburg, P.O. Box 90153, 5000 LE, Tilburg, The Netherlands; Tilec, CentER and CEPR; email: j.boone@uvt.nl. Schottmüller: Department of Economics, University of Copenhagen, Øster Farimagsgade 5, building 26, DK-1353 Copenhagen K, Denmark; email: christoph.schottmuller@econ.ku.dk.

1. Introduction

A well documented problem in health insurance markets with voluntary insurance like the US is that people either have no insurance at all or are underinsured.[1] Standard insurance models–inspired by the seminal work of Rothschild and Stiglitz (1976) (RS) and Stiglitz (1977)–predict that *healthy* people have less than perfect insurance or–in the extreme–no insurance at all. However, both popular accounts like Cohn (2007) and academic work like Schoen et al. (2008) show that people with low health status are over-represented in the group of uninsured and underinsured.[2] We develop a model to explain why sick people end up with little or no insurance. We do this by adding two well documented empirical observations (discussed below) to the RS model: (i) richer people tend to be healthier and (ii) health is a normal good. Technically speaking, introducing the latter two effects can lead to a violation of single crossing in the model.

There is another indication that the standard RS framework with single crossing does not capture reality in the health insurance sector well. The empirical literature that is based on RS does not unambiguously show that asymmetric information plays a role in health insurance markets. One would expect that people have private information about their health risks—think for example of preconditions, medical history of parents and other family members or life style. However, some papers, like for example Cardon and Hendel (2001) or Cutler et al. (2008), do not find evidence of asymmetric information while others do, e.g. Bajari et al. (2005) or Munkin and Trivedi (2010). The test for asymmetric information employed in these papers is the so called "positive correlation test," i.e. testing whether riskier types buy insurance contracts with higher coverage.[3]

We show that an insurance model with a violation of single crossing is capable of explaining why healthy people have better insurance than people with a low health status. In particular, the positive correlation property no longer holds if single crossing is violated. Consequently, testing for this positive correlation can no longer be viewed as a test for

[1] In empirical studies, underinsurance is defined using indicators of financial risk. To illustrate, one definition of underinsurance used by Schoen et al. (2008) is "out-of-pocket medical expenses for care amounted to 10 percent of income or more". In our theoretical model, underinsurance refers to less than socially optimal/efficient insurance.

[2] In the words of Schoen et al. (2008, pp. w303): "underinsurance rates were higher among adults with health problems than among healthier adults".

[3] "Risk" is in structural estimation papers—broadly speaking—interpreted as a parameter on which the distribution of health shocks depends.

asymmetric information.

Single crossing means that people with higher health risks have a higher willingness to pay for marginally increasing coverage, e.g. reducing copayments. If this property holds for all possible coverage levels, a given indifference curve of a high risk type can cross a given indifference curve of a low risk type at most once. A rough intuition for why the stylized facts above can lead to a violation of single crossing is given by the following: at full coverage (indemnity insurance that pays for all medical costs), high risk (low health) types will tend to spend more on treatments than low risk types. Hence, a small reduction in coverage, leads to a bigger loss in utility for high risk types. Now consider health insurance with low coverage where the insured faces substantial copayments. Because health is a normal good, it is possible that the rich-healthy type spends more on treatment than the low income, low health type. Put differently, a rich-healthy type might utilize the insurance more conditional on falling ill. In that case, a small change in coverage can have a bigger effect on the utility of the healthy type than of the low health agent. The healthy type will therefore have a higher willingness to pay for a marginal increase in coverage than the low health type. This violates single crossing.

We show the following results. In insurance models without single crossing, higher health risks are not necessarily associated with more coverage while this prediction is inevitable with single crossing. More specifically, we analyze in the same framework a setting of perfect competition as well as settings with market power. If insurance companies have market power, high risk types might have less coverage in equilibrium than low risk types. This is not the case if the insurance market is perfectly competitive: there would always be a profitable pooling contract in such a situation. If firms have market power, they do not offer this pooling contract because profits from low risk types are lower in the pooling contract. It should be noted that in equilibria in which high risk individuals have low insurance coverage, their insurance coverage is below first best. This leads to different policy implications than suggested by the literature on advantageous selection; see section 5.

The starting point for our paper is the positive correlation property which is established in various forms in the theoretical literature. The most general treatment is Chiappori et al. (2006). Their main focus is a positive correlation between coverage and expenditure claims while we are interested in the correlation between patient risk and coverage. Although Chiappori et al. (2006, pp. 787) note that they do not assume single crossing, our model is not a special case of their framework. In particular, in their model higher risk types have higher expenditure claims (in expectation). This is not necessarily the case in our model because of the utilization effect analyzed in our paper. Put differently, we analyze a situation where the agent has a treatment choice after the risk realizes while Chiappori et al. (2006) analyze a model where the agent can take an action that influences the risk distribution before the risk realizes.

The literature on violations of single crossing is relatively scarce and has so far not dealt with ex post decisions, e.g. treatment decisions made after the risk realizes. There are three papers analyzing perfectly competitive insurance markets with 2×2 types: people differ in two dimensions and both dimensions can either take a high or a low value. In Smart (2000) and Wambach (2000), the two dimensions are risk and risk aversion. Netzer and Scheuer (2010) model an additional labor supply decision and the two dimensions are productivity and risk. All papers have a pooling result, i.e. if single crossing does not hold two of the four types can be pooled. Only in Netzer and Scheuer (2010) there can be equilibria where some low risk types have more coverage than some high risk types. However, the wealthiest types have the lowest coverage in their model. This contrasts with the empirical observation in the health insurance sector mentioned above. In Smart (2000) and Wambach (2000), the high risk/high risk aversion type receives full coverage and the (low, low) type gets partial coverage. The $(high, low)$ and $(low, high)$ type can be pooled on an intermediate coverage level. Although two types with different risks are pooled, the positive correlation property still holds (weakly) in those models. The pooling itself is a result of the fact that some high risk types are less risk averse than some low risk types. Given that high risk types are likely to be poor in the health insurance context, even this pooling result appears unlikely to apply in the health insurance sector.

Jullien et al. (2007) take a different approach to answer the question why high risk types might have lower coverage in insurance markets. They use a model where types differ in risk aversion and single crossing is satisfied. Hence, types with higher risk aversion will have more coverage in equilibrium. At the same time, more risk averse agents might engage more in preventive behavior. If types are still separated in equilibrium and risk aversion differences remain the driving force, high risk aversion types will exhibit less risk (due to prevention) and higher coverage. Similar explanations for "advantageous selection" as in Jullien et al. (2007) can be found in Hemenway (1990) and De Meza and Webb (2001). While differences in risk aversion can explain the observed outcome of some insurance markets, e.g. Jullien et al. (2007) mention car insurance, this explanation does not easily fit the stylized facts of the health insurance market. We come back to this in section 5.

Since risk in the health sector is exogenously different for different persons, e.g. due to genetics, we follow RS and take a different starting point than Jullien et al. (2007). We assume risk differences instead of risk aversion differences. The result that high risk people have low coverage is in our paper not the result of low risk aversion. The driving force is the violation of single crossing caused by empirically documented income differences between high risks and low risks; see section 2. This is also in line with empirical evidence in Fang et al. (2008) who show that income is a source of advantageous selection in the medigap insurance market.

In the following section, we explain by use of a small model why consumers' preferences for health insurance violate single crossing. Section 3 introduces a general insurance model in which equilibria under perfect competition, monopoly and oligopoly are derived. In section 4, we illustrate the setup and the results with two numerical examples. Section 5 relates our results to the advantageous selection literature and section 6 concludes. Proofs are relegated to the appendix.

2. Income and health

We present a model where SC is violated because income affects treatment choices and differs between types. Unlike previous papers, e.g. Wambach (2000), De Meza and Webb (2001) and in some sense also in Netzer and Scheuer (2010), we do not assume that risk aversion depends on income or wealth. We do not see differences in risk aversion as a natural explanation for under-insurance problems in health care; see section 5.

The idea of our model is that partial coverage contracts require people to finance a part of the costs of treatment out of their own pocket. In this case, low income agents may decide to choose cheaper treatment or forgo treatment altogether. This effect is documented in the medical literature, see for example Piette et al. (2004b), Piette et al. (2004a) or Goldman et al. (2007). Put differently, the fact that health is a normal good can lead to a violation of single crossing. The reason is that poor, high risk types do not utilize the insurance fully when copayments are substantial. Therefore, their willingness to pay for a marginal increase in coverage can be lower than the one of rich, low risk types who utilize the insurance fully.

This utilization effect is well established in the medical literature. By extrapolating from their sample to the US population Piette et al. (2004a, p. 1786) conclude that "2.9 million of the 14.1 million American adults with asthma (20%) may be cutting back on their asthma medication because of cost pressures." They also document for a number of chronical conditions that people from low income groups are much more likely to report foregoing prescribed treatment due to costs.[4] Further examples can, for instance, be found in Piette et al. (2004b), Goldman et al. (2007), Schoen et al. (2010) or Schoen et al. (2008, pp. w305) who report that "[b]ased on a composite access indicator that included going without at least one of four needed medical care services, more than half of the underinsured and two-thirds of the uninsured reported cost-related access problems".

The utilization effect leads to a violation of single crossing if richer people face lower health risks, i.e. income and health risk are negatively correlated. This is also well documented in the empirical health literature, see for example Frijters et al. (2005), Finkelstein and McGarry (2006), Gravelle and Sutton (2009) or Munkin and Trivedi (2010). Potential explanations for this correlation between income and health include the following. High income people are better educated and hence know the importance of healthy food, exercise etc. Healthy food options tend to be more expensive and therefore better affordable to high income people. Or (with causality running in the other direction) healthy people are more productive and therefore earn higher incomes.

To illustrate how the described features of the health sector can lead to a violation of single crossing and also to give an example for models encompassed by the general model of section 3, we present a simple model of health insurance demand. We assume that consumers have one of two risk types $\theta \in \{\theta^h, \theta^l\}$ and that the risk type of a consumer is his private information. A type θ consumer faces a health shock $s \in [0, 1]$ with distribution (density) function $F(s|\theta)(f(s|\theta))$. We take $s = 1$ as the state in which the agent is healthy and needs no treatment. Lower health states s correspond to worse health. The assumption that the θ^h type has a higher health risk than the θ^l type can now be stated as $F(s|\theta^h) > F(s|\theta^l)$ for each $s \in \langle 0, 1\rangle$. In words, low health states s are more likely for the θ^h type than for the θ^l type.

A consumer can buy a health insurance contract. Insurance contracts consist of an

[4] For most chronic diseases people with income less than \$ 20000 are roughly 2 (5) times more likely to forgo prescribed treatment due to costs than people with an income between \$ 20000 and \$ 40000 (more than \$ 60000); see table 3 in Piette et al. (2004a) for details.

insurance premium p and coverage q. The most straightforward interpretation is that the consumer has a copayment rate of $1-q$ but it will become clear that the same effects as below are also present in contracts with deductibles.

Once an agent receives a health shock $s<1$, she can increase her health by tretament $\tau \in [0, 1-s]$. That is, any treatment between no treatment and full recovery is available. Low income consumers with partial coverage, i.e. $q<1$, may decide to choose cheaper treatment than if they had full insurance.[5]

Let $w(\theta)$ denote the wealth (or income) of a type θ agent. The agent maximizes expected utility

$$\begin{aligned} u(q,p,\theta) &= \int_0^1 \{v(w(\theta)-p-(1-q)\tau(s,q,\theta), s+\tau(s,q,\theta))\} dF(s|\theta) \\ & \text{where } \tau(s,q,\theta) \text{ is defined as:} \\ \tau(s,q,\theta) &= \arg\max_{\tau\in[0,1-s]} v(w(\theta)-p-(1-q)\tau, s+\tau) \end{aligned} \tag{1}$$

where $v(y,x)$ is the utility function of an agent which depends on consumption of other goods (y) and health (x). Denoting partial derivatives by subscripts, we assume that $v(y,x)$ satisfies $v_y, v_x > 0, v_{yy}, v_{xx} < 0$ and that health is a normal good: $v_{xy} \geq 0$. That is, utility increases in both health and consumption of other goods at a decreasing rate. As income increases, people's preference for health increases as well. In line with the empirical literature cited above, we assume that income and health status θ are negatively correlated: $w(\theta^h) \leq w(\theta^l)$.

The first order condition for an interior solution $\tau(s,q,\theta) \in [0, 1-s]$ can be written as[6]

$$v_x(w(\theta)-p-(1-q)\tau(s,q,\theta), s+\tau(s,q,\theta)) = (1-q)v_y(w(\theta)-p-(1-q)\tau(s,q,\theta), s+\tau(s,q,\theta)). \tag{2}$$

To see the implications of this model for single crossing, consider the slope of the indifference curves in (q,p)-space:

$$p_q(q,u,\theta) = -\frac{u_q}{u_p} = \frac{\int_0^1 v_y(w(\theta)-p-(1-q)\tau(s,q,\theta), s+\tau(s,q,\theta))\tau(s,q,\theta) dF(s|\theta)}{\int_0^1 v_y(w(\theta)-p-(1-q)\tau(s,q,\theta), s+\tau(s,q,\theta)) dF(s|\theta)} \tag{3}$$

[5]Implicitly, we assume that contracts cannot be contingent on treatment choice. Given the problems of verifiability of treatment and quantity choice as well as the possibility of doctor and patient to "collude" against the insurance (see Ma and McGuire (1997) for an analysis of these problems), this seems not unreasonable.

[6]If the left hand side of (2) is higher (lower) than the right hand side for all τ, the boundary solution $\tau = 1-s$ ($\tau = 0$) results.

In words, the slope p_q equals the weighted average of treatment $\tau(s,q,\theta)$ over the states s with weight

$$\frac{v_y(w(\theta)-p-(1-q)\tau(s,q,\theta),s+\tau(s,q,\theta))f(s|\theta)}{\int_0^1 v_y(w(\theta)-p-(1-q)\tau(s,q,\theta),s+\tau(s,q,\theta))dF(s|\theta)} \tag{4}$$

on state s (where the weights integrate to 1).

We will now show in two steps how single crossing can be violated in this model. Single crossing means that one type's indifference curves are steeper than the other type's indifference curves at any contract (q,p). First, we will show that at full coverage, i.e. at $q=1$, the high risk type has the steeper indifference curve. In a second step, it is shown that the low risk type can have the steeper indifference curve at low coverage levels. Hence, single crossing does not hold. Along the way, we show that wealth differences causing different utilization in case of falling ill are the reason for the violation of single crossing, i.e. single crossing will be satisfied if both types have the same wealth.

As the optimal treatment in the case of full coverage is clearly $\tau=1-s$ (irrespective of type), it is easy to calculate p_q at full coverage:

$$p_q(1,u,\theta)=\frac{\int_0^1 v_y(w(\theta)-p,1)(1-s)dF(s|\theta)}{\int_0^1 v_y(w(\theta)-p,1)dF(s|\theta)}=\int_0^1(1-s)dF(s|\theta)$$

where the last equality follows from the fact that $v_y(w(\theta)-p,1)$ is constant in s. The stochastic dominance assumption implies that θ^h puts more weight on low s states (where $1-s$ is high) compared to θ^l. Hence, under these assumptions, the high risk type has the steeper indifference curve at full coverage.

Single crossing is satisfied if there are no wealth differences between types, i.e. $w(\theta^h)=w(\theta^l)$. The idea is that without wealth differences (2) yields the same optimal treatment for both types. Put differently, $\tau(s,q,\theta)$ is independent of θ. If patients choose more treatment in worse health states, single crossing will be satisfied: due to stochastic dominance, θ^h types have higher weight (4) on low health states. In these low health states, treatment $\tau(s)$ is high. This follows from equation (2) which—using the implicit function theorem—implies

$$(-(1-q)^2v_{yy}+2(1-q)v_{yx}-v_{xx})\frac{d\tau}{ds}=v_{xx}-(1-q)v_{yx}. \tag{5}$$

The assumptions on v imply that $d\tau/ds\geq 0$. Clearly, this weak inequality also holds at boundary solutions. Because high risks place a higher weight on states in which treatment is high, p_q in (3) is higher for θ^h than for θ^l types for all $q\in[0,1]$ in case of equal wealth.

If high risk types have lower incomes, they will choose a lower treatment than low risk types at partial coverage contracts. This follows from equation (2), since

$$\frac{d\tau}{dw} = \frac{-(1-q)v_{yy} + v_{xy}}{-(1-q)^2 v_{yy} + 2(1-q)v_{yx} - v_{xx}} > 0. \tag{6}$$

Hence, if $\tau(s, q, \theta^l) \in [0, 1-s]$ is an interior maximum, the θ^h type tends to choose lower treatment τ. In words, since a fraction $1-q$ of the treatment cost has to be paid by the insured, a low income θ^h patient may choose cheaper treatment than the richer θ^l type (as health is a normal good). Since she does not utilize the insurance as much as the (rich) low risk type, type θ^h has a lower marginal willingness to pay for insurance coverage at low coverage levels. However, for high levels of coverage, i.e. q close to 1, wealth differences matter less in the treatment choice because the patient does not have to pay (much) out of pocket for the treatment. Consequently, single crossing can be violated.

This model—where agents differ in income and treatment choice is endogenous—can generate the violation of single crossing mentioned above. In the following section, we will analyze a general model that allows for violations of single crossing.

3. Insurance model

This section introduces a general model of (health) insurance that allows us to consider both the case where single crossing (SC) is satisfied and the case where it is not satisfied (NSC). After describing the demand side of the insurance market, we consider three alternatives for the supply side: perfect competition, monopoly and oligopoly.

3.1. Demand side model

Following RS, we consider an agent with utility function $u(q, p, \theta)$ where $q \in [0, 1]$ denotes coverage or generosity of her insurance contract,[7] $p \geq 0$ denotes the insurance premium and $\theta \in \{\theta^l, \theta^h\}$ with $\theta^h > \theta^l > 0$ denotes the type of consumer.[8] Higher θ denotes a higher risk in the sense of higher expected costs in case $q^h = q^l = 1$; see below. This could, for instance, be the case due to chronic illness or higher risk due to a genetic precondition. We make the following assumptions on the utility function.

[7] Apart from literal coverage—where $1-q$ denotes the agent's copayments—q could, for example, be interpreted as $1/(1 + \text{deductible})$. Note that in models without moral hazard both parameters are similar in the sense that people with high expected expenditure are affected by co-payments and deductibles more relative to people with low expected expenditure.

[8] We follow RS in assuming that there are only two types. For an analysis of a violation of single crossing with a continuum of types θ, see Araujo and Moreira (2003, 2010) and Schottmüller (2011).

Assumption 1 *The utility function* $u(q,p,\theta)$ *is continuous and differentiable. It satisfies* $u_q > 0, u_p < 0$. *We define the indifference curve* $p(q,u,\theta)$ *as follows:*

$$u(q,p(q,u,\theta),\theta) \equiv u \tag{7}$$

We assume that these indifference curves $p(q,u,\theta)$ *are differentiable in* q *and* u *with* $p_q = -u_q/u_p > 0, p_u = 1/u_p < 0$.
Further, the crossing at $q = 1$ *satisfies:*

$$p_q(1,u^h,\theta^h) > p_q(1,u^l,\theta^l) \tag{C1}$$

for all $u^l \geq \bar{u}^l = u(0,0,\theta^l), u^h \geq \bar{u}^h = u(0,0,\theta^h)$.

In words, utility u is increasing in coverage q and decreasing in the premium p paid for insurance. For given type θ and utility level u, the indifference curve $p(q,u,\theta)$ maps out combinations (q,p) that yield the same utility. Because higher coverage leads to higher utility, p has to increase to keep utility constant. Hence, indifference curves are upward sloping in (q,p) space ($p_q > 0$). Increasing u—for a given coverage level q—requires a lower price. Thus, raising u shifts an indifference curve downwards ($p_u < 0$).

Type $k \in \{h,l\}$ buys insurance if it leads to a higher utility than her outside option $\bar{u}^k$. This outside option is given by the "empty insurance contract": $q = p = 0$.

At full coverage ($q = 1$), a marginal reduction in coverage q should be compensated by a bigger decrease in the premium p for θ^h compared to θ^l. This reflects the fact that the θ^h type faces higher expected health care expenditures, i.e. she is the high risk type. At full coverage, other factors that could differ by types like willingness to pay for treatment do not play a role. In this sense, this assumption "defines" what higher θ means: at full coverage, higher θ types face higher expected costs. With the same idea we assume that expected costs for the insurer of a contract with $q = 1$ is higher for the θ^h than for the θ^l type: $c(1,u^h,\theta^h) > c(1,u^l,\theta^l)$ for all $u^h \geq \bar{u}^h, u^l \geq \bar{u}^l$. Intuitively, u should not matter for health care consumption at full coverage and the high risk type will use the insurance more.

To allow for income effects, for instance in treatment choice, the cost function depends on u. However, we assume two regularity conditions.

Assumption 2 *For each type* $k \in \{h,l\}$ *and* $q \in [0,1]$, *we assume that*

- $c_u(q,u^k,\theta^k) \geq 0$ *for* $u^k \geq \bar{u}^k$,

- $c(1, u^k, \theta^k) = c(1, \tilde{u}^k, \theta^k)$, *for* $u^k, \tilde{u}^k \geq \bar{u}^k$.

In words, as the income of the agent increases (which *ceteris paribus* leads to higher utility), the agent has more money to spend on treatment. As the insurer pays a fraction $q \geq 0$ of these treatments, this leads to (weakly) higher costs for the insurer. Second, costs at full coverage ($q = 1$) do not vary in utility. Intuitively, if $q = 1$ treatments are for free for the agent and there is no reason to forgo treatments, irrespective of the level of $u^k \geq \bar{u}^k$.

Because of (C1), the single crossing condition (also called sorting, constant sign or Spence-Mirrlees condition) reads

$$p_q(q, u^h, \theta^h) > p_q(q, u^l, \theta^l) > 0 \tag{SC}$$

$q \in [0, 1]$ and $u^h \geq \bar{u}^h$, $u^l \geq \bar{u}^l$ such that $p(q, u^h, \theta^h) = p(q, u^l, \theta^l)$. The intuition is the following. Suppose an indifference curve of type θ^h intersects with an indifference curve of type θ^l in some point (p, q). Then (SC) implies that the slope of the θ^h indifference curve will be higher. It follows that these two indifference curves can intersect only once.

We consider both the case where (SC) is satisfied and the case where it is violated (NSC). In both the SC and NSC cases, we maintain the assumption that $q = 1$ is the efficient insurance level (EI) for each type $\theta \in \{\theta^l, \theta^h\}$. Hence, we do not consider the case where insurance leads to inefficiency by inducing over-consumption of treatments.

Assumption 3 *For a given utility level* u^k*, welfare (and therefore profits) are maximized at full coverage, i.e.*

$$\max_{q \in [0,1]} p(q, u^k, \theta^k) - c(q, u^k, \theta^k) \tag{EI}$$

is uniquely maximized by $q = 1$ *for each* $k \in \{h, l\}$ *and* $u^k \geq \bar{u}^k$.

This basically means that the insurance motive, i.e. transferring risk from a risk averse agent to a risk neutral insurer, is not overruled by other considerations. To illustrate, we do not assume that the low income agent's preference for health/treatment is so low that foregoing insurance would be socially optimal. Put differently, we assume that full insurance is socially desirable. Underinsurance—with no insurance as extreme case—results therefore not from first best but from informational distortions and price discrimination motives.

Our motivation for making this assumption is twofold. First, this assumption simply normalizes the socially efficient insurance level in the same way as in RS. Hence, we only deviate from the RS set up by allowing for both SC and NSC. Second, we want to argue that under realistic assumptions, θ^h types have less than full insurance. If the first best insurance level is actually below one, then this result would follow rather trivially. Another way of putting this is to say that a θ^h type would buy full insurance if she chose from all actuarially fair insurance contracts. In this sense, the answer to our question "why healthy people have high coverage" is not simply that unhealthy people cannot afford actuarially fair insurance. The result that $q^h < 1$ is not directly driven by utility and cost functions. As mentioned before, we are interested in adverse selection issues, not moral hazard. It should also be noted that—unlike Jullien et al. (2007)—we do not model prevention. In our framework, consumers make a treatment decision after falling ill. Prevention efforts are taken ex ante (before falling ill) and may be negatively affected by generous coverage ($q = 1$). In a model where $q = 1$ leads either to excessive health care consumption or to under-investment in prevention efforts, it is almost trivial to show that $q^h < 1$ is optimal.[9] In our model, types are separated because the θ^h type prefers the cheap low coverage insurance above the expensive generous insurance contract.

To illustrate that the assumptions encompass not only models with endogenous treatment choice as in section 2 but also standard models of the insurance literature, we show that the RS setup satisfies all of our assumptions.

Example 1 *In the RS setup, an agent faces with probability θ a monetary loss D. She has initial wealth w and expected utility $u(q, p, \theta) = \theta v(w-(1-q)D-p)+(1-\theta)v(w-p)$ where $v' > 0$ and $v'' < 0$. Using the implicit function theorem, $p_q(1, u, \theta) = \theta D$ and therefore (C1) is satisfied. Note that (C1) will also be satisfied if w depends on θ. The insurer is risk neutral and has profits $p-\theta qD$. As profits do not depend on u (or w), assumption 2 is trivially satisfied. Since the agent is risk averse and the insurer is risk neutral, assumption 3 is also satisfied.*

[9] As shown in propositions 2 and 3, we get $q^h < q^l$ in case the l-type is willing to pay more than the h-type for efficient coverage. If due to either moral hazard or prevention, first best coverage satisfies $q^* < 1$, it becomes easier to get $q^h < q^l$. Indeed, because of assumption (C1), at q close to 1 the indifference curve for the h-type is steeper than for the l-type. As health is a normal good, this is the case to a lesser extent (or not at all) at $q^* < 1$. Hence it is easier to satisfy $p(q^*, u^l, \theta^l) > p(q^*, u^h, \theta^h)$ than it is to get $p(1, u^l, \theta^l) > p(1, u^h, \theta^h)$.

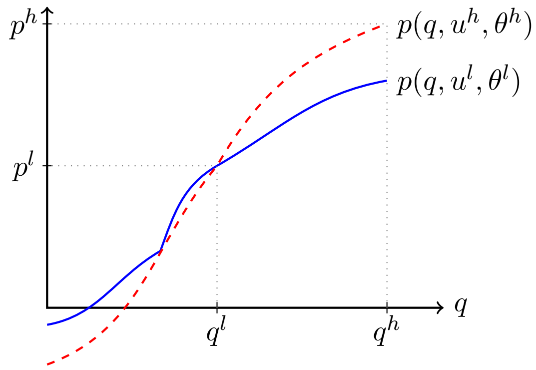

Figure 1: incentive compatibility constraints

3.2. Supply side

An insurer offers a menu of two contracts; one contract for each type. The contract of type θ^k consists of a coverage level q^k and a price p^k resulting in utility level $u^k = u(q^k, p^k, \theta^k)$. The two contracts can be identical (pooling) or differ from each other (separating). In case of separating, the contracts have to satisfy the incentive compatibility (IC) constraints for each type:

$$p(q^l, u^l, \theta^l) \geq p(q^l, u^h, \theta^h) \qquad (IC_h)$$

$$p(q^h, u^h, \theta^h) \geq p(q^h, u^l, \theta^l) \qquad (IC_l)$$

where u^k is the utility level of type θ^k when buying the contract (q^k, p^k). The first constraint implies that the contract intended for θ^h (i.e. $(q^h, p(q^h, u^h, \theta^h))$) lies on a (weakly) lower indifference curve for θ^h than the contract that is meant for the θ^l type $(q^l, p(q^l, u^l, \theta^l))$. That is, the inequality implies $u(q^h, p^h, \theta^h) \geq u(q^l, p^l, \theta^h)$ where $p^i = p(q^i, u^i, \theta^i)$ with $i \in \{h, l\}$. This is illustrated in figure 1 where the inequality is binding: both contracts lie on the θ^h indifference curve (dashed line). Similarly, the second inequality implies that $u(q^l, p^l, \theta^l) \geq u(q^h, p^h, \theta^l)$. In figure 1, (IC_l) is satisfied with inequality: The θ^l indifference curve through the (q^l, p^l) contract is below p^h at q^h.

Irrespective of the mode of competition and whether (SC) holds, we have the following result for the models that we use below where profit maximizing insurers offer contracts simultaneously and independently.

Lemma 1 *At least one type has full coverage. If the types are separated under the optimal contract scheme $(q^l, p^l), (q^h, p^h)$ with $q^l \neq q^h$, then at most one incentive constraint binds.*

3.2.1. Perfect competition

The literature on insurance models considers mostly perfect competition.[10] We show that perfect competition implies $q^h = 1$ even if (SC) is not satisfied. Hence, in our model, market power on the insurance side is needed to obtain equilibria with $q^h < 1$.

Following the RS definition of the perfect competition equilibrium, we require that (i) each offered contract makes non-negative profits and (ii) given the equilibrium contracts there is no other contract yielding positive profits.

Proposition 1 *In a perfectly competitive insurance market, the high risk type buys a full coverage contract, i.e. $q^h = 1$, in equilibrium.*

The proposition shows that even with violations of single crossing, high risk types will obtain (weakly) higher coverage than low risk types in the RS equilibrium.

However, as is well known, existence of equilibrium is not guaranteed in the RS framework. When the only separating equilibrium is broken by a pooling contract, a RS equilibrium does not exist. Wilson (1977) (see also Miyazaki (1977), Spence (1978) and Netzer and Scheuer (2011)) offers a slightly different model of perfect competition in an insurance market where equilibrium always exists. The main difference between RS and Wilson (1977) is that Wilson uses an equilibrium notion where every firm expects that all contracts that become unprofitable due to the firm's offered contracts will be withdrawn immediately. Netzer and Scheuer (2011) specify an extensive form game that generates this equilibrium as a subgame perfect equilibrium that is robust to the introduction $\varepsilon > 0$ withdrawal costs. We denote these contracts "MW contracts".

As shown by Miyazaki (1977) and Netzer and Scheuer (2011), MW contracts are—adapted to our notation—the solution to the following optimization problem:[11]

$$\max_{p^h,q^h,p^l,q^l} u(q^l, p^l, \theta^l) \qquad (P_{MW})$$

[10]See Jack (2006) and Olivella and Vera-Hernández (2007) for exceptions using a Hotelling model to formalize market power on the insurer side of the market. These papers assume that (SC) is satisfied and hence find efficient insurance for the θ^h type.

[11]Note that lemma 1 does not apply here, as firms do not make their contract offers simultaneously and independently.

subject to the constraints

$$p^l \geq p(q^l, u(q^h, p^h, \theta^h), \theta^h) \tag{8}$$

$$p^h \geq p(q^h, u(q^l, p^l, \theta^l), \theta^l) \tag{9}$$

$$\phi(p^h - c(q^h, u(q^h, p^h, \theta^h), \theta^h)) + (1-\phi)(p^l - c(q^l, u(q^l, p^l, \theta^l), \theta^l)) \geq 0 \tag{10}$$

$$p^h \leq c(q^h, u(q^h, p^h, \theta^h), \theta^h) \tag{11}$$

$$q^i \in [0,1], \quad p^i \in [0, w^i] \qquad \text{for } i = h, l \tag{12}$$

where ϕ is the share of θ^h types in the population. The first two constraints are incentive compatibility constraints. Constraint (10) is a non-negative profits constraint (in expectation). Cross-subsidization from θ^h to θ^l types is prohibited by (11).

Netzer and Scheuer (2011) derive a number of properties of optimization problem (P_{MW}). For us the relevant properties are: (i) this optimization problem always has a unique solution, (ii) this solution is a Wilson equilibrium and (iii) if constraint (11) binds, the solution coincides with the RS contracts. In this sense, the MW contracts generalize the RS equilibrium. If the RS equilibrium does not exist, the MW contracts include a cross subsidization from the low risk to the high risk type. As our setup is not exactly the same as in Netzer and Scheuer (2011), we prove these results for our setup in the webappendix.[12]

The question is: does this characterization of the "perfect competition" outcome allow for $q^h < 1$? The answer is no.

Lemma 2 *In any MW contract, $q^h = 1$.*

Hence, we need to deviate from perfect competition (either in the RS or the MW sense) to get $q^h < q^l$. Put differently, the positive correlation property holds in these models of perfect competition irrespectively of single crossing. To explain violations of the positive correlation property which are pointed out in the empirical literature, it is necessary to deviate from the perfect competition assumption.

Indeed, recent research for the US, see Dafny (2010), shows that health insurers have market power. More generally, in most countries where health insurance is provided by private companies, these firms tend to be big due to economies of scale in risk diversification. Hence, one would expect that these firms have some market power.

[12] see `http://sites.google.com/site/janboonehomepage/home/webappendices`

3.2.2. Monopoly

Consider an insurance monopolist. In this case, the positive correlation property can be violated in the (NSC) case.

Proposition 2 *The type with the highest willingness to pay for full coverage, i.e. the type θ^k with highest $p(1, \bar{u}^k, \theta^k)$, obtains a full coverage contract in an insurance monopoly. Either her incentive compatibility or her individual rationality constraint is binding (or both). The other type's individual rationality constraint is binding.*

Let θ^k denote the type with the highest willingness to pay for full coverage. It follows from the proposition that θ^k obtains a contract $(q, p) = (1, p^k)$ for some $p^k \leq p(1, \bar{u}^k, \theta^k)$. The monopoly outcome is now pinned down by the choice of p^k. If $p^k = p(1, \bar{u}^k, \theta^k)$, then both individual rationality constraints are binding. In this case, type θ^{-k} might be excluded, i.e. θ^{-k} gets the contract $(0, 0)$. If $p^k = p(1, \bar{u}^{-k}, \theta^{-k})$, both types are pooled. If $p^k \in \langle p(1, \bar{u}^{-k}, \theta^{-k}), p(1, \bar{u}^k, \theta^k)\rangle$, the equilibrium separates the types and θ^{-k} gets an insurance contract with partial coverage. The optimal p^k depends on the share of θ^k types in the population.

A direct implication of proposition 2 is that high risk types will always have full coverage if single crossing is satisfied. To see this, note that the indifference curve corresponding to $\bar{u}^k$ (that is the individual rationality constraint) goes through the origin $(p, q) = (0, 0)$ for both types. With (SC) the indifference curve of the high risk type is steeper and lies therefore above the individual rationality constraint of the low risk type for all coverage levels.

Without single crossing this is no longer the case. We will give a numerical example below where the low risk type θ^l has the higher willingness to pay for full coverage. Therefore, the low risk type will receive full coverage. If the types are separated, which depends on the share of each type in the population, we find that $q^h < q^l = 1$. Put differently, the positive correlation property no longer holds in the monopoly setup.

The intuition for why market power is needed for $q^h < 1$ is the following. Under perfect competition, any candidate equilibrium where $q^h < 1$ allows a profitable deviation. An insurer who deviates by offering $\hat{q}^h = 1$ and the corresponding price $(p(1, u^h, \theta^h))$ on the θ^h type's indifference curve makes a positive profit on the θ^h type because of assumption 3. If θ^l types decide to buy this deviation contract as well, the insurer will make positive profits on these types as well because $c(1, u^l, \theta^l) < c(1, u^h, \theta^h)$. Hence, a perfect competition outcome with $q^h < 1$ is destroyed by such a profitable deviation.

For a monopolist, the situation is different. Again starting from $q^h < 1$ and considering a deviation to $\hat{q}^h = 1$, attracting θ^l types on the deviating contract is bad news if this reduces the profits the monopolist makes on the θ^l contract. If the loss on the θ^l types exceeds the gain on the θ^h type (due to assumption 3), $q^h < 1$ is an equilibrium outcome for the monopolist.

Of course, a monopolistic market structure is an extreme case and not entirely realistic in the health insurance market. However, the idea that the positive correlation property does not hold in the (NSC) case is more general. To illustrate this, we turn to an oligopoly setting next.

3.2.3. Oligopoly

This subsection uses a tractable duopoly model on the supply side. It serves as an illustration that the results from the monopoly setting also carry over to imperfect competition settings.

We assume that there are two profit maximizing insurers located at the end points 0 and 1 of a Hotelling line. Agents of both types are uniformly distributed over the $[0, 1]$ interval. The share of high risk types in the population is denoted by ϕ which is assumed to be independent from location $x \in [0, 1]$. An agent at position $x \in [0, 1]$ incurs transportation cost xt ($(1 - x)t$) when buying from insurer a (b) with $t > 0$. The agent maximizes the expected utility from the insurance contract minus the transportation costs. Each insurer offers a menu of contracts $\{(q^h, p^h), (q^l, p^l), (0, 0)\}$ where the first contract is intended for the θ^h type, the second for the θ^l type and the third "contract" denotes the agent's outside option of not buying insurance at all (which will not be used in equilibrium).[13] Insurers simultaneously offer menus and consumers choose their preferred contract afterwards.

For the following result, we need the fairly standard assumption $u_{pp} \leq 0$, i.e. the higher the price the higher is the utility loss from a marginal price increase. Put differently, there is a decreasing marginal utility from other goods.

Proposition 3 *Assume $u_{pp}(1, p, \theta^l) \leq 0$ for $p \in (0, p(1, \bar{u}^l, \theta^l))$. If $p(1, \bar{u}^l, \theta^l) > p(1, \bar{u}^h, \theta^h)$, then there exist parameter values $\phi > 0$ and $t > 0$ such that type θ^l obtains a full coverage*

[13]This means that "transportation costs" are not relevant for the participation decision. This ensures that firms compete also for high values of t. Hence, we rule out the case of (local) monopoly which was analyzed above.

contract in a separating equilibrium.

This proposition is similar to the result obtained for the monopoly setup. If the low risk type has a higher willingness to pay for full coverage, there are separating equilibria where he obtains full coverage. As noted above, since $p(0, \bar{u}^l, \theta^l) = p(0, \bar{u}^h, \theta^h)$, (SC) implies that $p(1, \bar{u}^l, \theta^l) < p(1, \bar{u}^h, \theta^h)$. But with a violation of single crossing it can indeed be the case that $p(1, \bar{u}^l, \theta^l) > p(1, \bar{u}^h, \theta^h)$. This is illustrated with two numerical examples in section 4.

4. Examples

In this section, we give two numerical examples in which our assumptions are satisfied, the violation of single crossing occurs and $q^h < q^l = 1$ can happen in monopoly and duopoly.[14]

Example 2 *Consider the following mean-variance utility set up. There are two states of the world: an agent either falls ill or stays healthy. The probability of falling ill is denoted by F^h ($F^l < F^h$) for type θ^h (θ^l). We choose $F^h = 0.07 > 0.05 = F^l$. Once an agent falls ill, the set of possible treatments is denoted by $\Gamma = \{\underline{\tau}, \bar{\tau}\}$. The utility of an agent of type $i = h, l$ with treatment choice $\tau \in \{\underline{\tau}, \bar{\tau}\}$ is written as:*

$$\begin{aligned} u(q, p, \theta^i) = & F^i(v(\tau, \theta^i) - (1 - q)\tau) + (1 - F^i)v(1, \theta^i) - p \\ & - \tfrac{1}{2} r^i F^i(1 - F^i)(v(1, \theta^i) - v(\tau, \theta^i) + (1 - q)\tau)^2 \end{aligned} \tag{13}$$

where $v(\tau, \theta^i)$ denotes the utility for type $i = h, l$ of having health τ and $r^i > 0$ denotes the degree of risk aversion. Hence an agent's utility is given by the expected utility minus $\frac{1}{2}r^i$ times the variance in the agent's utility. This is a simple way to capture that the agent is risk averse.[15]

Along an indifference curve where u is fixed, we find the following slope:

$$\frac{dp}{dq} = F^i \tau(q, \theta^i) + r^i F^i(1 - F^i)(v(1, \theta^i) - v(\tau(q, \theta^i), \theta^i) + (1 - q)\tau(q, \theta^i))\tau(q, \theta^i) \tag{14}$$

[14] For the Python code used to generate the examples, see: `http://sites.google.com/site/janboonehomepage/home/webappendices`. This webappendix also verifies that (EI) is satisfied for these two examples.

[15] When an agent of type i buys a product at price p that gives utility v, there are two ways to capture the marginal utility of income for agent i. First, overall utility can be written as $v - \alpha^i p$ where v is the same for each type i and α^i can differ. Low income types are then modeled to have high α^i; high marginal utility of income. Alternatively, one can write $v^i - \alpha p$ where α is the same for all types. Then low income types have low v^i. We have chosen the latter formalization with $\alpha = 1$. The assumption that treatment is

where $\tau(q, \theta^i)$ is the solution for τ solving

$$\max_{\tau \in \{\underline{\tau}, \bar{\tau}\}} v(\tau, \theta^i) - (1-q)\tau.$$

In words, once an agent falls ill, she decides which treatment to choose based on the benefit $v(\tau, \theta^i)$ and the out-of-pocket expenses $(1-q)\tau$.

We assume $\bar{\tau} = 0.6, \underline{\tau} = 0.2$ and the associated utilities for the θ^h type equal $v(1, \theta^h) = 0.9, v(\bar{\tau}, \theta^h) = 0.7, v(\underline{\tau}, \theta^h) = 0.45$ and similarly for the θ^l type: $v(1, \theta^l) = 1.1, v(\bar{\tau}, \theta^l) = 0.9, v(\underline{\tau}, \theta^l) = 0.5$. Hence, having high health is more important for the θ^l type compared to the θ^h type. We set $r^h = 1.1, r^l = 1.5$ which implies that

$$r^h F^h (1 - F^h)(v(1, \theta^h) - v(\bar{\tau}, \theta^h))\bar{\tau} = r^l F^l (1 - F^l)(v(1, \theta^l) - v(\bar{\tau}, \theta^l))\bar{\tau}. \tag{15}$$

In words, at $q = 1$ (where both types choose the highest treatment $\bar{\tau}$) the variance terms in the slope dp/dq (equation (14)) are equalized. Hence, assumption (C1) is satisfied because $F^h\bar{\tau} > F^l\bar{\tau}$ in equation (14).[16]

Figure 2 shows two indifference curves for the θ^l type (in red) and one for the θ^h type (in blue). It is clear that (SC) is violated. Indeed, there is $\tilde{q}$ such that for $q < \tilde{q}$ the indifference curve for the θ^l type is steeper than for the θ^h type. This is due to the fact that the θ^l type buys the expensive treatment $\bar{\tau}$ while the θ^h type buys $\underline{\tau}$. The kink in the indifference curve for the θ^h type happens at the value $\tilde{q}$ where the θ^h type switches from the cheap to the more expensive treatment. Hence, small increases in q for $q > \tilde{q}$ are worth more to the θ^h type than small increases in $q < \tilde{q}$. In fact, the figure shows that for $q > \tilde{q}$, the indifference curve for the θ^h type is steeper than the one for the θ^l type. This is the violation in single crossing.

In a simple mean-variance utility framework, it is therefore straightforward and intuitive to generate a violation of (SC).

The l-type has a higher willingness to pay for full coverage compared to the h-type. This can be seen in figure 2: the solid lines are the indifference curves of both types that go

a normal good is then implemented by assuming that

$$v(\bar{\tau}, \theta^h) - v(\underline{\tau}, \theta^h) < v(\bar{\tau}, \theta^l) - v(\underline{\tau}, \theta^l).$$

[16] Note also that the numerical values below are chosen such that the variance term in the utility function is also equal at full coverage (and weakly higher for θ^h if $q < 1$). Therefore, the violation of single crossing in our example is due to the different utilization of health insurance and not to differences in risk aversion that were the driving force in other papers on the violation of single crossing.

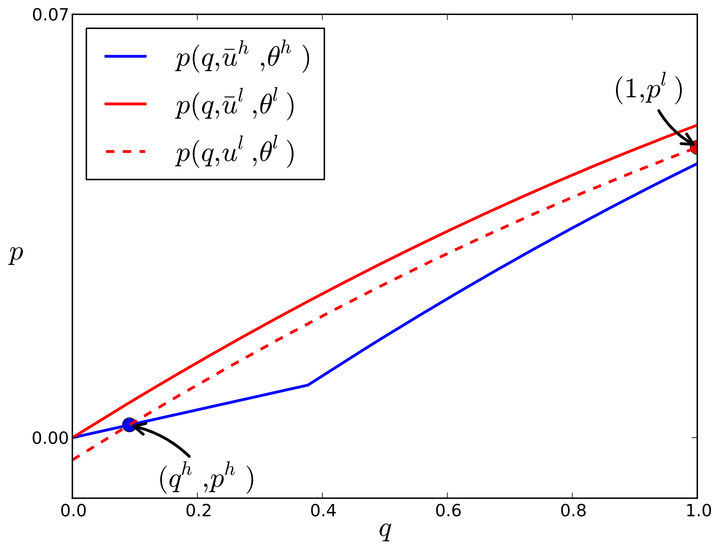

Figure 2: Indifference curves with parameter values in example 2 and $t = 0.018$

through the empty contract $(0, 0)$*, i.e. the indifference curves corresponding to individual rationality. Willingness to pay for full coverage is given by the value at* $q = 1$ *where the low risk type's indifference curve is above the high risk type's indifference curve.*

For this example, proposition 1 and lemma 2 imply that with perfect competition (in the RS or MW sense) we find that $q^h = 1$ in equilibrium. However, proposition 2 implies that in a monopoly framework $q^l = 1$. If the share of low risk type's is high enough, the optimal monopoly menu separates the two types and the positive correlation property will be violated.

Finally, we use the example to illustrate the logic behind proposition 3. We will show that it is straightforward to find examples where $q^l = 1$ and $q^h < 1$ in duopoly. The easiest way to do this is to find parameter values such that the individual rationality (IR) curve (that is, the indifference curve $p(q, \bar{u}, \theta)$) for the θ^l type lies everywhere above the IR curve for the θ^h type. As shown in figure 2, this is the case for the parameter values of our numerical example. Clearly, the Hotelling equilibrium contracts have to lie on or below the relevant IR curves.

First, assume that $\phi = 0$. In words, there are only θ^l types. Then it is routine to verify that $q^l = 1$ (because of assumption 3) and the Hotelling equilibrium price on the θ^l-market equals $p^l = F^l\bar{\tau} + t$.[17] This contract is denoted $(1, p^l)$ in figure 2 for the parameter values given above and $t = 0.018$. As this contract lies below θ^l's IR curve, it is, indeed, the equilibrium outcome. Let $u^l_{hotel.}$ denote θ^l's utility level associated with the $(1, p^l)$ contract: $u^l_{hotel.} = u(1, p^l, \theta^l)$. Contract (q^h, p^h) (although not bought by anyone as $\phi = 0$) is defined by the intersection of indifference curve $p(q, u^l_{hotel.}, \theta^l)$ (dashed curve in the figure) and θ^h's IR curve. This is the best contract on θ^h's IR curve that satisfies θ^l's incentive compatibility constraint.

Now increase ϕ slightly to $\phi > 0$ (but small). We claim that this results in an equilibrium with $q^l = 1 > q^h$. For this to be an equilibrium, we need that the indifference curve for the θ^l type at $q = 1$ lies above the indifference curve for the θ^h type at $q = 1$. Note that the equilibrium indifference curve for the θ^h type $(p(q, u^h_{hotel}, \theta^h))$ cannot lie above θ^h's IR curve. Hence, a sufficient condition for an equilibrium with $q^l = 1 > q^h$ is that θ^l's indifference curve $p(q, u^l_{hotel.}, \theta^l)$ at the new Hotelling equilibrium lies above θ^h's IR curve at $q = 1$. This is formally shown in the proof of proposition 3 and is intuitively clear: small changes in ϕ will lead to small changes in the indifference curve $p(q, u^l_{hotel.}, \theta^l)$. As this curve is above θ^h's IR curve at $q = 1$ in case $\phi = 0$, it will be above θ^h's IR curve for small positive values of ϕ.

Hence, a straightforward way to generate equilibria where the positive correlation property fails, is to find examples where the IR constraint for the θ^l type lies above the IR constraint for the θ^h type for each $q \in \langle 0, 1]$. Then there exist $t > 0$ and $\phi > 0$ such that the example has an equilibrium with $q^l > q^h$.

We use this example to illustrate the differences between our model and the Chiappori et al. (2006) framework. First, in their framework, higher risk types file higher expenditure claims (in expectation) with their insurer. In the example above, the high risk type has expected expenditure on treatments equal to $F^h\underline{\tau} = 0.014$ while for the low type we have $F^l\bar{\tau} = 0.030$; the opposite of Chiappori et al. (2006). Second, Chiappori et al. (2006, pp. 789) assume (NIP assumption) that the contract with lower coverage yields a higher profit margin for the insurer. In the example above, we find for the l-type that this margin equals $\pi^l = p^l - F^l\bar{\tau} = t = 0.018$. For the h-type we know that $p^h \leq p(q^h, \bar{u}^h, \theta^h)$ and hence we find that $\pi^h \leq p(q^h, \bar{u}^h, \theta^h) - q^h F^h\underline{\tau} = 0.0008$. In other words, the margin on the generous contract exceeds the margin on the partial coverage contract. This outcome

[17] Recall that in a Hotelling model with constant marginal costs c, the equilibrium price is given by $c + t$. See, for instance, Tirole (1988, pp. 280) or the proof of proposition 3 in the appendix with $u_p = -1$.

violates the NIP assumption in Chiappori et al. (2006).

Example 3 *Assume that a type θ^i agent falls ill with probability θ^i where $i = h, l$. Falling ill reduces her health state from 1 to 0. If she falls ill, the agent can choose a treatment $\tau \in \Gamma = [0, 1]$ which will bring back her health state to h. The costs of this treatment are also τ. Hence, an agent with wealth w and insurance coverage q at price p will have money to spend on other goods equal to $m = w - p - (1 - q)\tau$.*

The agent's utility depends on her health state and the money she has to spend on other goods. We assume that the utility function takes the CARA form in both components and is additively separable in the two components. The utility of an agent with health state τ and money for other goods m is therefore

$$U(m, h) = -e^{-m} - e^{-\tau}.$$

The agent maximizes expected utility with her treatment and contract choice. Hence, conditionally on falling ill, the agent maximizes

$$-e^{-(w-p-(1-q)\tau)} - e^{-\tau}$$

with her treatment choice. Therefore, the optimal treatment choice of the agent can be derived as

$$\tau(w, q, p) = \max\{0, \min\{1, \frac{w - p - ln(1 - q)}{2 - q}\}\} \quad (16)$$

where the max *and* min *expression ensure that the treatment is in the feasible treatment range $[0, 1]$. It is straightforward to verify that $p_q(1, u^i, \theta^i) = \theta^i$ and therefore (C1) holds.*

We want to show that $p(1, \bar{u}^h, \theta^h) < p(1, \bar{u}^l, \theta^l)$ which then implies that $q^l = 1$ in a monopoly setting according to proposition 2 (and by choosing the proportion of high and low risk types it is then straightforward to generate separating equilibria where $q^h < 1$). We choose $w^l = 2$ which implies that low risk types will choose $\tau = 1$ even when being uninsured. With $w^h = 1.2$, $\theta^h = 0.5$ and $\theta^l = 0.45$ we get $p(1, \bar{u}^h, \theta^h) = 0.537 < 0.573 = p(1, \bar{u}^l, \theta^l)$. Hence, the low risk type has the higher willingness to pay for a full coverage insurance contract. The individual rationality constraints are depicted in figure 3. The figure also shows that (SC) is violated in the example: the low risk type's indifference curve is steeper at $q = 0$ but flatter at full coverage. The reason is again that the l-type chooses the treatment $h = 1$ at any coverage level while the h-type will do so only if coverage is high enough. As figure 3 is qualitatively similar to figure 2, a duopoly equilibrium with $q^l = 1 > q^h$ can be constructed in the same way as described above.

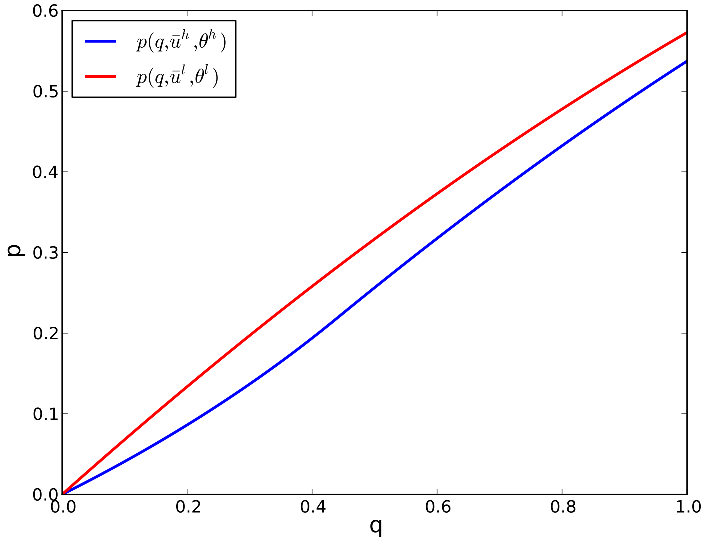

Figure 3: Indifference curves with parameter values as in example 3.

5. Advantageous selection

Adverse selection models predict that high risk types have efficient coverage while low risk types are under-insured. We cited evidence in the introduction showing that in the US health insurance market the problem is the opposite: high risk types are under-insured while low risk types tend to have efficient coverage.

Advantageous selection models can generate the prediction that high risk types have low coverage while low risk types have high coverage. As we mentioned above, we do not find these models convincing for the health insurance market. In this section, we compare our framework to the advantageous selection set-up. We discuss three main differences and explain why our model fits the health insurance market better. We conclude this section by pointing out the differences in policy implications of these two approaches.

Recall from the introduction a typical advantageous selection model, as in Jullien et al. (2007). Suppose that people differ in their degree of risk aversion. Then more risk averse people have a higher valuation of insurance. Further, assume that more risk averse people tend to have lower expected costs. This can be the case because more risk averse people tend to take more precautions in the form of prevention activities.

We follow Chetty and Finkelstein (2012) in defining adverse (advantageous) selection. Take two contracts, denoted H and L where H is more generous (say, features lower co-payments) than L. Assume each contract is priced at the average cost of the people choosing the contract. Then adverse (advantageous) selection is defined as the case where the high risk type—θ^h in our notation—chooses the H (L) contract while θ^l chooses the L (H) contract.

In a model where people differ in risk aversion, it can happen that the high risk averse type with low expected costs chooses the H contract, while the low risk averse type with high expected costs chooses the L contract: advantageous selection.

Although reminiscent of our result above where θ^h buys a contract with less coverage than θ^l, there are three important differences. First, in the definition of Chetty and Finkelstein (2012) (where contracts are priced at the average cost of the people buying the contract) our model is an adverse selection model. As shown in section 3.2.1, perfect competition (price equal to average cost) implies that θ^h gets higher coverage than θ^l. Only when insurers have market power, we get that the H (L) contract is bought by $\theta^l(\theta^h)$.

Second, in our set-up the relation between expected expenditure and type is not one-to-one. At full insurance, θ^h features higher expected claims than θ^l. However, at less than full insurance this is not necessarily the case. As the θ^h agent consumes less health care than θ^l if there are substantial co-payments, it is not clear which type features higher expected expenditures. We find in example 2 that the expected expenditure for θ^h is lower than for θ^l. In empirical research, one has to be careful in interpreting this situation where the L contract is bought by people with lower expected costs (while they are actually the high risk type).

Finally, the most important difference between the advantageous selection model above and our framework is our assumption 3: full insurance is efficient for each type. In the advantageous selection model, people differ in both their degree of risk aversion and their expected costs. It is not hard to think of examples where such a model makes sense. For instance, people who are more risk averse tend to be more careful drivers and hence have lower expected claims in car insurance. However, we do not find this a convincing model of under-insurance in a health care context. Such a framework where people differ in expected costs and degree of risk aversion implies that people that go without health insurance (like currently in the US) decide to do so because they are almost risk neutral and hence do not need much insurance. But recent empirical work by Fang et al. (2008) shows that differences in risk aversion are not a source of advantageous selection in the medigap insurance market. Moreover, the advantageous selection model implies that due to their low degree of risk aversion, it is in fact optimal for these people to remain uninsured. Consequently, there is not much reason for government intervention. In fact, in advantageous selection models policies aimed at raising insurance coverage—like mandatory insurance—tend to reduce welfare (De Meza and Webb, 2001; Einav and Finkelstein, 2011).

In contrast, in our model, the high risk type is under-insured (or even goes without insurance) because she cannot afford the H contract. She needs to spend money on food, housing, transport etc. and hence is left with the L contract. The under-insurance then prevents access to certain valuable treatments as documented by Cohn (2007) and Schoen et al. (2008).

Further, because of assumption 3 government intervention raising coverage for θ^h can be welfare enhancing. This gives a justification for Obama's plans for health insurance reform to reduce the number of people without insurance as laid down in the Patient Protection and Affordable Care Act.

6. Conclusion

Standard insurance models, e.g. Rothschild and Stiglitz (1976) or Stiglitz (1977), predict higher coverage for agents with higher risks. We show that this prediction no longer holds if single crossing is violated and firms have market power.

In the health care sector, agents with higher income have lower risks and more insurance. Put differently, the predictions of the standard insurance model with single crossing are contradicted by the data. We show that the negative correlation between income and risk can cause a violation of single crossing. With a violation of single crossing, the empirical findings in the health literature can be reconciled with a standard insurance model.

From an empirical point of view, our paper casts doubt on the positive correlation test: given our result that separating equilibria exist in which agents with higher risk have less coverage (negative correlation), it is evident that the results of such a test have to be interpreted with care. In particular, such a test cannot be used to test for the presence of asymmetric information when single crossing is violated.

Further, our analysis shows that one should be careful with interpreting actual expenditure as a signal of risk type. Indeed, it is easy to find examples where the high risk type (due to low coverage and low utilization) has lower expected expenditure than the low risk type.

Although our result that high risk types end up under-insured while low risk types have efficient insurance is reminiscent of advantageous selection models, the welfare implications are very different. In our set-up, policy measures that raise insurance coverage for the under-insured can be welfare enhancing. This is in line with the motivation of the ObamaCare reform package.

References

Araujo, A. and H. Moreira (2003). Non-monotone insurance contracts and their empirical counterparts. Discussion Paper #512, Getulio Vargas Foundation.

Araujo, A. and H. Moreira (2010). Adverse selection problems without the Spence-Mirrlees condition. *Journal of Economic Theory 145*(5), 1113–1141.

Bajari, P., H. Hong, and A. Khwaja (2005). A semiparametric analysis of adverse selection and moral hazard in health insurance contracts. Technical report, Preliminary Mimeo, University of Michigan and Duke University.

Cardon, J. H. and I. Hendel (2001). Asymmetric information in health insurance: Evidence from the national medical expenditure survey. *RAND Journal of Economics 32*(3), 408–427.

Chetty, R. and A. Finkelstein (2012, October). Social insurance: Connecting theory to data. Working Paper 18433, National Bureau of Economic Research.

Chiappori, P.-A., B. Jullien, B. Salanié, and F. Salanié (2006). Asymmetric information in insurance: general testable implications. *RAND Journal of Economics 37*(4), 783–798.

Cohn, J. (2007). *Sick: The untold story of America's health care crisis–and the people who pay the price.* Harper Perennial.

Cutler, D. M., A. Finkelstein, and K. McGarry (2008). Preference heterogeneity and insurance markets: explaining a puzzle of insurance. *American Economic Review 98*(2), pp. 157–162.

Dafny, L. (2010). Are health insurance markets competitive? *American Economic Review 100*(4), 1399–1431.

De Meza, D. and D. Webb (2001). Advantageous selection in insurance markets. *RAND Journal of Economics 32*(2), 249–262.

Einav, L. and A. Finkelstein (2011). Selection in insurance markets : theory and empirics in pictures. *Journal of Economic Perspectives 25*(1), 115–138.

Fang, H., M. P. Keane, and D. Silverman (2008). Sources of advantageous selection: evidence from the medigap insurance market. *Journal of Political Economy 116*(2), 303–350.

Finkelstein, A. and K. McGarry (2006). Multiple dimensions of private information: evidence from the long-term care insurance market. *American Economic Review 96*(4), 938–958.

Frijters, P., J. Haisken-DeNew, and M. Shields (2005). The causal effect of income on health: evidence from German reunification. *Journal of Health Economics 24*(5), 997–1017.

Glicksberg, I. L. (1952). A further generalization of the Kakutani fixed point theorem with application to Nash equilibrium points. *Proceedings of the American Mathematical Society 38*, 170–174.

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

Gravelle, H. and M. Sutton (2009). Income, relative income, and self-reported health in Britain 1979-2000. *Health Economics 18*(2), 125–145.

Hemenway, D. (1990). Propitious selection. *Quarterly Journal of Economics 105*(4), 1063–1069.

Jack, W. (2006). Optimal risk adjustment with adverse selection and spatial competition. *Journal of Health Economics 25*(5), 908 – 926.

Jullien, B., B. Salanie, and F. Salanie (2007). Screening risk-averse agents under moral hazard: single-crossing and the CARA case. *Economic Theory 30*(1), 151–169.

Ma, C. and T. McGuire (1997). Optimal health insurance and provider payment. *American Economic Review 87*(4), 685–704.

Miyazaki, H. (1977). The rat race and internal labor markets. *Bell Journal of Economics 8*(2), 394–418.

Munkin, M. and P. Trivedi (2010). Disentangling incentives effects of insurance coverage from adverse selection in the case of drug expenditure: a finite mixture approach. *Health Economics 19*(9), 1093–1108.

Netzer, N. and F. Scheuer (2010). Competitive screening in insurance markets with endogenous wealth heterogeneity. *Economic Theory 44*(2), 187–211.

Netzer, N. and F. Scheuer (2011). A game theoretic foundation of competitive equilibria with adverse selection. *Working paper.*

Olivella, P. and M. Vera-Hernández (2007). Competition among differentiated health plans under adverse selection. *Journal of Health Economics 26*(2), 233 – 250.

Piette, J., M. Heisler, and T. Wagner (2004a). 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.

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

Rothschild, M. and J. Stiglitz (1976). Equilibrium in competitive insurance markets: an essay on the economics of imperfect information. *Quarterly Journal of Economics 90*(4), 629–649.

Schoen, C., S. R. Collins, J. L. Kriss, and M. M. Doty (2008). How many are underinsured? trends among U.S. adults, 2003 and 2007. *Health Affairs (Project Hope) 27*(4), 298–309.

Schoen, C., R. Osborn, D. Squires, M. M. Doty, R. Pierson, and S. Applebaum (2010). How health insurance design affects access to care and costs, by income, in eleven countries. *Health Affairs 29*(12), 1–12.

Schottmüller, C. (2011). Adverse selection without single crossing: the monotone solution. *Center Discussion Paper, no. 2011–123*.

Smart, M. (2000). Competitive insurance markets with two unobservables. *International Economic Review 41*(1), 153–170.

Spence, A. (1978). Product differentiation and performance in insurance markets. *Journal of Public Economics 10*(3), 427–447.

Stiglitz, J. (1977). Monopoly, non-linear pricing and imperfect information: the insurance market. *Review of Economic Studies 44*(3), 407–430.

Tirole, J. (1988). *The Theory of Industrial Organization.* Mit Press.

Wambach, A. (2000). Introducing heterogeneity in the Rothschild-Stiglitz model. *Journal of Risk and Insurance 67*(4), 579–591.

Wilson, C. (1977). A model of insurance markets with incomplete information. *Journal of Economic Theory 16*(2), 167–207.

7. Appendix: Proofs

Proof of lemma 1. We start with the proof of the second statement. Suppose both incentive constraints were binding, i.e. θ^h and θ^l are both indifferent between the two contracts. First, look at the case where $q^h, q^l < 1$. Call the utility levels of the two types under the equilibrium contracts u^l and u^h. Now take the indifference curves corresponding to these utility levels and call them $p(q, u^l, \theta^l)$ and $p(q, u^h, \theta^h)$ and define $\iota = \arg\max_{k\in\{l,h\}} p(1, u^k, \theta^k)$. Changing θ^ι's menu point to $(1, p(1, u^\iota, \theta^\iota))$ will increase profits by assumption 3. By the definition of ι, this change is also incentive compatible.

Second, take the case where $q^k = 1$ and $q^{-k} < 1$ for some $k \in h, l$ and suppose again that both incentive constraints were binding. But according to assumption 3 pooling on the contract of θ^k would lead to higher profits. Hence, at most one incentive constraint is binding.

$q^\iota = 1$ follows from the argument in the first step and therefore at least one type has to have full coverage. *Q.E.D.*

Proof of proposition 1. Suppose to the contrary that $q^h < 1$ in equilibrium. Lemma 1 implies then that $q^l = 1$. Note that θ^h has to prefer her contract strictly to the θ^l contract: otherwise, pooling on the θ^l contract would be a profitable deviation by assumption 3. Given that (IC_h) is not binding, the θ^l contract leads to zero profits: otherwise, marginally decreasing its price (and thereby attracting all demand of θ^l types) would be a profitable deviation.

The contract (q^h, p^h) leads to nonnegative profits; otherwise it would not be offered in equilibrium.[18] Denote by u^h the utility level θ^h derives from (q^h, p^h) and by $p(q, u^h, \theta^h)$ the indifference curve of θ^h associated with her contract. By assumption 3, the contract $(1, p(1, u^h, \theta^h))$ for type θ^h yields higher profits than (q^h, p^h). For $\varepsilon > 0$ small enough, the contract $(1, p(1, u^h, \theta^h) - \varepsilon)$ is strictly preferred by θ^h to (q^h, p^h) and yields higher profits than (q^h, p^h). If the contract $(1, p(1, u^h, \theta^h) - \varepsilon)$ also attracts θ^l types, profits from those θ^l types will be positive as well as those are better risks. This would be an additional gain as it was shown above that the θ^l contract yields zero profits. Therefore, $(1, p(1, u^h, \theta^h) - \varepsilon)$ is a profitable deviation, i.e. a contract with strictly positive profits and demand. Consequently, $q^h < 1$ cannot be an equilibrium. *Q.E.D.*

Proof of lemma 2 If equation (11) binds, the MW contract coincides with the RS

[18] In fact, it has to be a zero profit contract.

contract and the result follows from proposition 1 above.

Consider the case where inequality (11) is slack in the MW solution. Then it must be the case that (8) binds. To see this, suppose—by contradiction—that both (8) and (11) are slack. But then it is possible to raise p^h and reduce p^l in such a way that (9) and (10) remain satisfied. This increases utility of the θ^l type, contradicting that we are considering a solution to (P_{MW}). It follows from this contradiction that (8) is binding.

Now we claim that $q^h = 1$. Again, we prove this by contradiction. Suppose that $q^h < 1$. We consider two cases:

1. $p(1, u^l, \theta^l) \leq p(1, u^h, \theta^h)$: we show that increasing $q^h < 1$ to 1 raises the θ^l type's utility. First, note that increasing q^h to 1 and p^h to $\tilde{p}^h = \min\{p(1, u^h, \theta^h), c(1, u^h, \theta^h\}$ does not violate (8) as utility stays the same or is increased for the θ^h type. Second, it does not violate (11) by construction of $\tilde{p}^h$. Third, (10) is not violated by this change either. To see this, recall that initially (11) was slack: $p^h < c^h$. Now we either have $\tilde{p}^h = c^h$ or $\tilde{p}^h = p(1, u^h, \theta^h)$. In both cases, equation (10) is strictly relaxed by this change (in the latter case because of assumption 3). Finally, consider (9). If this constraint is violated (because $\tilde{p}^h = c(1, u^h, \theta^h) < p^h$ and $\tilde{u}^h > u^h$), the solution $q^l = 1, p^l = \tilde{p}^h, q^h = 1, p^h = \tilde{p}^h$ is feasible (since $c(1, u^l, \theta^l) < c(1, u^h, \theta^h)$). If this constraint is not violated, the change to $q^h = 1$ is feasible. As the change relaxed constraint (10) strictly, it is possible to reduce p^h, p^l slightly such that all constraints remain satisfied and θ^l's utility increases. This contradicts that the solution with $q^h < 1$ solves problem (P_{MW}).

2. $p(1, u^l, \theta^l) > p(1, u^h, \theta^h)$: this implies that $q^l = 1$. Assume by contradiction that $q^l < 1$. Increase q^l to $\tilde{q}^l = 1$ and p^l to $\tilde{p}^l = p(1, u^l, \theta^l) - \varepsilon$ for $\varepsilon > 0$ small. By assumption 3, this change satisfies (10). Further, (8) is satisfied because $p(1, u^l, \theta^l) > p(1, u^h, \theta^h)$. Finally, (9) is relaxed by this change and (11) is unaffected. Hence this change satisfies all constraints and increases utility for the θ^l-type. This contradicts that the contracts with $q^l < 1$ solve (P_{MW}). But then we have $q^l = 1$ and $p(1, u^l, \theta^l) > p(1, u^h, \theta^h)$ which implies that (8) is slack. However, we showed above that (8) is binding if (11) is slack.

In each of these two cases we have a contradiction and hence we conclude that $q^h = 1$. *Q.E.D.*

Proof of proposition 2. Define $\iota = \arg\max_{k \in \{h,l\}} p(1, \bar{u}^k, \theta^k)$. By lemma 1, one type has full coverage. Suppose that $q^\iota < 1$ and therefore $q^\kappa = 1$ with $\kappa \in \{h, l\}$ and $\kappa \neq \iota$. Note that the individual rationality constraint of θ^ι cannot be binding as otherwise θ^ι would misrepresent as θ^κ by the definition of ι. But then the incentive compatibility constraint of θ^ι has to be binding as the monopolist could increase p^ι otherwise. By assumption 3, the monopolist could achieve a higher profit by pooling both types on θ^κ's contract. This contradicts the optimality of $q^\iota < 1$.

If both types are pooled, the optimal contract will be $(q, p) = (1, p(1, \bar{u}^\kappa, \theta^\kappa))$ and the individual rationality constraint of θ^κ will be binding. If the types are separated, the incentive compatibility constraint of θ^κ cannot bind: since $q^\iota = 1$, pooling on θ^ι's contract would lead to higher profits by assumption 3 if the incentive constraint was binding. As increasing p^κ relaxes the incentive compatibility constraint of θ^ι, the individual rationality constraint of θ^κ has to bind: otherwise, increasing p^κ would increase profits.

Last note that increasing p^ι would be feasible and increase profits if neither the incentive compatibility nor the individual rationality constraint of θ^ι was binding. *Q.E.D.*

Proof of proposition 3. The first step is to analyze the game where $\phi = 0$, i.e. a standard Hotelling game where only low risk type exist. From assumption 3, $q^l = 1$ in this setting and firms only compete in prices. By assumption 2, costs do then not depend on price and can be denoted by $\bar{c}$. A firm maximizes $(p^l - \bar{c})(\frac{1}{2} + \frac{u(1,p,\theta^l) - u^b}{2t})$ where u^b is the utility offered by the other firm. Because of the assumption $u_{pp}(1, p, \theta^l) \leq 0$, the objective is concave and the best response is defined by the first order condition

$$t + u(1, p, \theta^l) - u^b + (p - \bar{c})u_p(1, p, \theta^l) = 0.$$

Note that there is a symmetric equilibrium defined by the equation $(p - \bar{c})u_p(1, p, \theta^l) = -t$. The left hand side of this equation is decreasing in p and therefore there is only one symmetric equilibrium. We will now argue that there are also no asymmetric equilibria, i.e. the game has a unique equilibrium. The argument is that the slope of the best response function is less than one whenever crossing the 45° line where $p = p^b$: by the implicit function theorem, $p'(p^b) = \frac{-u_p(1,p^b,\theta^l)}{-2u_p(1,p,\theta^l) + (p - \bar{c})u_{pp}(1,p,\theta^l)}$. Consequently, $0 < p'(p^b) < 1$ whenever $p = p^b$. Given that there is a symmetric equilibrium where $p = p^b$, this implies that the best response functions can only intersect once, i.e. there is a unique equilibrium.

The second step is to see that by choosing t appropriately the game with $\phi = 0$ leads to an equilibrium price $p^* \in (p(1, \bar{u}^h, \theta^h), p(1, \bar{u}^l, \theta^l))$. This simply follows from the fact that t "shifts" the first order condition above. For the rest of the proof let t take such a value.

The third step is to show that for ϕ small enough $q^l = 1$ and $q^h < 1$.[19] By lemma 1, at least one type has to have full coverage. If $q^h = 1$, then $p^h \leq p(1, \bar{u}^h, \theta^h)$ to satisfy individual rationality. Suppose, both types were pooled. Recall that $p^l < p(1, \bar{u}^h, \theta^h)$ is not an equilibrium when $\phi = 0$, i.e. there exists a profitable deviation for at least one insurer. For ϕ small enough, this deviation is again profitable as profit functions are continuous in all variables and parameters. Hence, there cannot be pooling for small but positive ϕ. Next suppose there were separating equilibria with $q^h = 1$ and $q^l < 1$ for all $\phi > 0$. By assumption 3, q^l has to converge to 1 as ϕ decreases (otherwise setting $q^l = 1$ and adjusting the price to keep u^l fixed is a profitable deviation for small enough ϕ). But then the same argument as in the pooling case shows that there is a profitable deviation for ϕ small enough. It follows that $q^l = 1$ and $q^h < 1$ for small enough $\phi > 0$. *Q.E.D.*

[19] Existence of equilibrium follows from Glicksberg (1952).