Competitive Screening in Insurance Markets with Endogenous Wealth Heterogeneity

Nick Netzer* Florian Scheuer†‡

Second Revision, June 2009

Abstract

We examine equilibria in competitive insurance markets with adverse selection when wealth differences arise endogenously from unobservable savings or labor supply decisions. The endogeneity of wealth implies that high risk individuals may ceteris paribus exhibit the lower marginal willingness to pay for insurance than low risks, a phenomenon that we refer to as irregular-crossing preferences. In our model, both risk and patience (or productivity) are privately observable. In contrast to the models in the existing literature, where wealth heterogeneity is exogenously assumed, equilibria in our model no longer exhibit a monotone relation between risk and coverage. Individuals who purchase larger coverage are no longer higher risks, a phenomenon frequently observed in empirical studies.

JEL-classification: D82, G22, J22.

Keywords: Insurance Markets, Adverse Selection, Multidimensional Screening.

---

*University of Zurich, Socioeconomic Institute, Bluemlisalpstrasse 10, CH-8006 Zurich, Switzerland. Email: nick.netzer@soi.uzh.ch.

†Massachusetts Institute of Technology, Department of Economics, E52-391, Cambridge, MA 02139, USA. Email: scheuer@mit.edu.

‡We are grateful to Daron Acemoglu, Abhijit Banerjee, Felix Bierbrauer, Friedrich Breyer, Wolfgang Buchholz, Peter Diamond, Glenn Ellison, Oliver Fabel, Amy Finkelstein, Jon Gruber, Mathias Kifmann, Jim Poterba, Casey Rothschild, Harris Schlesinger, Arthur Snow, two anonymous referees and seminar participants at the Universities of Augsburg, Konstanz, MIT, Zurich and the 2007 Annual Meetings of the Risk Theory Society, the American Risk and Insurance Association, and the European Economic Association for valuable comments on earlier versions of this paper. The usual disclaimer applies.

1 Introduction

In the standard screening model going back to Rothschild and Stiglitz (1976), individuals differ only in a single dimension, namely their risk of incurring a loss. In this simple framework, equilibrium contracts are such that risk types are fully separated, and high risk individuals obtain more insurance coverage than low risks. As shown by Chiappori, Jullien, Salanié, and Salanié (2006), this positive correlation between risk and coverage is a robust property of equilibria in a much larger class of competitive screening models, including extensions to heterogeneous preferences, moral hazard and non-expected utility. It has therefore been the basis for much of the empirical research trying to identify adverse selection in insurance markets. The evidence to support this prediction, however, has been mixed, with many studies not being able to reject the null hypothesis of zero correlation.[1] Nevertheless, this cannot be interpreted as indicating that there is no adverse selection in these markets. For instance, Finkelstein and McGarry (2006) show that preference based selection in the US long-term care insurance market may offset risk based selection so that, in aggregate, those with more insurance are not higher risks.

In this paper, we develop a novel theoretical explanation for the empirical results based on *endogenous wealth accumulation*. We extend the Rothschild-Stiglitz screening model to a dynamic setting in which wealth levels are the result of optimal savings or labor supply decisions. We show that, with endogenous wealth accumulation, high risk individuals may ceteris paribus exhibit the *lower* marginal willingness to pay for insurance than low risk individuals, a phenomenon that we refer to as *irregular-crossing* indifference curves. The reason is that, facing the same insurance contract, high risk individuals ceteris paribus choose to supply more labor and to save more, which results in a higher wealth level and thus lower risk aversion. If this effect is sufficiently strong, it overcompensates the direct risk effect.

It turns out that the existence of irregular-crossing preferences, which cannot arise in previously considered screening models, is crucial for characterizing the possible equilibria in competitive insurance markets. Most importantly, for an economy with two-dimensional heterogeneity in both risk and patience (or risk and productivity if wealth differences arise from labor supply rather than savings), our central result shows that equilibria can emerge in which those with more insurance coverage are not nec-

[1] See Cawley and Philipson (1999) for the US life insurance market, Chiappori and Salanié (2000) and Chiappori, Jullien, Salanié, and Salanié (2006) for the French automobile insurance market, and Cardon and Hendel (2001) for the US health insurance market.

essarily higher risks, which is consistent with much of the empirical evidence. The intuition is that patient (or productive) individuals are willing to save (or work) more in response to the risk they face and thus demand less insurance in equilibrium due to their higher wealth and thus lower risk aversion.

Our paper contributes to a recent theoretical literature that attempts to derive and explain empirically testable correlations in insurance markets. An important example is the work by Chiappori and Salanié (2000) and Chiappori, Jullien, Salanié, and Salanié (2006) mentioned above, which provides general conditions for a positive correlation between insurance coverage and ex post risk. The assumption used to derive this property in their model is that profits do not increase with coverage in the equilibrium set of contracts (see Chiappori, Jullien, Salanié, and Salanié (2006), p. 789). Although this assumption may be satisfied in many competitive settings (especially when competition drives down the profits of all contracts to zero in equilibrium), we show that it does not necessarily hold in our model with multiple dimensions of private information and endogenous wealth heterogeneity. Despite perfect competition, binding incentive constraints may prevent profitable contracts from being undercut, thus allowing for positive profits in equilibrium.[2] The violation of the non-increasing profits assumption is the key reason why the logic of Chiappori, Jullien, Salanié, and Salanié (2006) fails in our model. More importantly, we are able to construct insurance market equilibria in which there is no longer a perfect ordering between risk and coverage, without assuming any of the ingredients considered as necessary by Chiappori, Jullien, Salanié, and Salanié (2006), such as imperfect competition or fixed costs, biased beliefs of customers, or a perfect correlation between risk and unobservable risk aversion.

Our model shares a common goal with other theoretical work that tries to explain why insurance markets with adverse selection may not exhibit a positive correlation between risk and insurance coverage. Most of this literature has focused on combining adverse selection and moral hazard in insurance markets. In these models, individuals can reduce their damage probability by an unobserved effort decision, which gives rise to moral hazard. To introduce adverse selection, De Meza and Webb (2001) and Jullien, Salanié, and Salanié (2007) assume that individuals differ in their privately known risk attitude, which affects their effort decision.[3] These models indeed generate equilibria

[2] In the empirical part of their paper, Chiappori, Jullien, Salanié, and Salanié (2006) find strong evidence against the validity of the non-increasing profits property in the data and interpret it as indicating some sort of market power. Our model demonstrates that deviations from perfect competition are not required to explain a failure of this property.

[3] This idea was first put forward informally by Hemenway (1990) and Hemenway (1992). Another approach within this class of models, chosen by Stewart (1994) and Chassagnon and Chiappori (1997),

where those with more insurance coverage do not have a higher ex post risk. They do so, however, by making a number of additional assumptions. In particular, they depart from the assumption of competitive markets. Jullien, Salanié, and Salanié (2007) consider a monopolistic insurer and De Meza and Webb (2001) introduce costs that also drive a wedge between premiums and expected claims. In contrast, our model generates the result in a perfectly competitive framework without introducing an additional moral hazard problem and assumptions on insurers' cost functions. More importantly, the previous models stick to a framework with one-dimensional heterogeneity between agents, where ex post risk and risk attitude are perfectly correlated.[4] It remains an open question whether their equilibria continue to exist in settings that allow for a less restrictive structure of heterogeneity. Our results do not rely on a perfect correlation between risk and risk aversion.

Our model is also related to theoretical contributions that extend the basic framework of Rothschild and Stiglitz (1976) to two-dimensional heterogeneity. Such models have been developed by Smart (2000), Wambach (2000) and Villeneuve (2003). They assume that insurance customers differ in wealth and hence risk aversion in addition to risk, whereby the correlation between risk and risk aversion is not assumed to be perfect. Countervailing incentives and thus deviations from usual risk separation emerge if individuals differ in both characteristics so that the resulting effects work in opposite directions. It is a central result of all these models, however, that risk and coverage are perfectly ordered in equilibrium: the largest coverage contract purchased by any low risk type still has a (weakly) smaller coverage than the smallest coverage contract purchased by any high risk type. Our model differs conceptually from this literature as we do not assume differences in wealth to be exogenous, but explicitly model their emergence in a dynamic model. Our main result shows that the ordering of types obtained in the models with exogenous wealth differences does not generalize to a more realistic setting where heterogeneity in wealth and hence risk aversion arises endogenously from unobservable savings or labor supply decisions, thus allowing us to generate equilibria that are consistent with the empirical evidence.

is to assume that agents differ in their effort cost, which is also private information. While these models yield some deviations from the standard Rothschild-Stiglitz model, they have in common that the correlation between ex post risk and insurance coverage will always be positive in equilibrium.

[4] De Meza and Webb (2001) assume that some individuals are risk-neutral and hence neither purchase insurance nor take preventive actions. Their expected damage is therefore larger than that of the individuals who purchase partial insurance and take preventive measures due to their higher risk aversion. This generates a negative relationship between individuals' risk and their insurance coverage. Jullien, Salanié, and Salanié (2007) also consider a two-type model only. They are concerned with the question how risk aversion affects the power of incentives provided by the optimal contract.

Finally, at a technical level, our paper extends the work by Rochet and Chone (1998) and Armstrong and Rochet (1999) on multidimensional screening to a situation with competitive principals and non-quasilinear preferences of the agents. The interaction of imperfect insurance markets with self-insurance in the form of saving and labor supply leads to a situation with countervailing incentives in the sense of Lewis and Sappington (1989). This is because the agents' marginal willingness to pay for insurance is influenced not only by their risk, but also their wealth level. Since savings and labor supply and hence wealth levels react to uncertainty and hence depend on the insurance market outcome, countervailing incentives arise endogenously in our model. We demonstrate how the resulting interactions between labor, capital and insurance markets affect insurers' ability to screen their customers.

The paper is structured as follows. In section 2, we show that the endogeneity of wealth differences may give rise to irregular-crossing preferences. In section 3, we introduce our model of the insurance market with two-dimensional heterogeneity and show that it generates an empirically appealing equilibrium. Section 4 concludes. Several proofs are relegated to the appendix.

2 Preferences for Insurance with Endogenous Wealth

In this section, we develop a model to capture both a competitive insurance market with asymmetric information and self-insurance in the form of wealth accumulation. There are two major ways to do this. First, we may assume that individuals differ in both their damage risk and their labor productivity, and choose their labor supply endogenously. Insurance companies can neither observe individual risks, productivities, nor hours of work. An alternative interpretation is to assume that individuals live for two periods, save in the first period and face different damage risks in the second period. A further dimension of heterogeneity is introduced by assuming that individuals differ in their rate of time preference. Again, risks, savings and time preferences are private information. Since both interpretations are completely equivalent, we choose the second modelling strategy and provide comments on how to interpret our results in terms of the first approach as we go along.

A second difference between our model and most of the screening literature starting from Rothschild and Stiglitz (1976) is that we conduct our analysis in the contract space as opposed to the output (wealth) space. As we will demonstrate below, with endogenous wealth accumulation, individual endowment points in the output space depend on the chosen insurance contract. This is because the income effects and the level of uncertainty implied by insurance contracts affect optimal labor supply and savings. Characterizing equilibria in the contract space circumvents this endogeneity of endowment points and is therefore technically more convenient.

2.1 The Setup

To understand how individuals' wealth accumulation decisions and their demand for insurance interact, we consider a simple dynamic model with two periods $t = 1, 2$. There is a continuum of individuals who are born at the beginning of period 1 with identical endowments W_1. Individuals consume in both periods with intertemporal preferences given by $u_1(c_1) + \delta_i u_2(c_2)$, where $\delta_i \in (0, 1]$ is the discount factor, and we assume $u_t(.)$, $t = 1, 2$, to be continuous and increasing, and $u_2(.)$ to be weakly concave. Note that this general formulation allows for different sub-utility functions for each period, including $u_1(.) = u_2(.)$ as a special case. There is a linear savings technology that allows agents to transfer wealth from period 1 to period 2 at some exogenous rate of return R, which we normalize to $R = 1$. We consider an economy of individuals who differ in their patience δ_i, $i = L, H$, and their probability p_j, $j = L, H$, of incurring a damage of size D in the second period $t = 2$, with the conventions $\delta_L < \delta_H$ and $p_L < p_H$. Let n_{ij} denote the share of individuals with patience δ_i and risk p_j. These individuals will be referred to as $ij-$individuals.

2.2 Savings and Risk

In period $t = 1$, each type ij chooses an amount of savings $s_{ij} = W_1 - c_{1,ij}$, which generates second period wealth $W_{2,ij} = s_{ij}$. In the first period, individuals also purchase insurance contracts that specify the share $\beta \in [0, 1]$ of the damage that is covered, and a premium $d \in \mathbb{R}^+$, both paid in period $t = 2$. Throughout the paper, the notation $A > B$ implies that insurance contract A has a larger coverage and a larger premium than contract B. Given a contract $C = (\beta, d)$ from the contract space $\mathscr{C} = [0, 1] \times \mathbb{R}^+$, type ij's optimal savings s^*_{ij} solve the following program

$$\max_{s_{ij} \in [0, W_1]} u_1(W_1 - s_{ij}) + \delta_i[p_j u_2(s_{ij} - (1 - \beta)D - d) + (1 - p_j)u_2(s_{ij} - d)]. \quad (1)$$

Let the solution be denoted by $s^*_{ij}(\beta, d)$ or, for simplicity, $s^*_{ij}(C)$ defined as

$$\max\left\{\arg\max_{s_{ij}\in[0,W_1]} u_1(W_1 - s_{ij}) + \delta_i[p_j u_2(s_{ij} - (1-\beta)D - d) + (1-p_j)u_2(s_{ij} - d)]\right\}.$$

This assumes the largest savings level to be chosen in case of indifference. Our first proposition provides conditions for this solution to exist, and characterizes the comparative statics of optimal savings with respect to the risk type p_j.

Proposition 1. *$s^*_{ij}(C)$ as defined above exists and, at any given contract $C \in \mathscr{C}$, weakly increases in p_j.*

Proof. For notational simplicity, we suppress subscripts and define

$$F(s,p) \equiv u_1(W_1 - s) + \delta\left(pu_2(s - \alpha_1) + (1-p)u_2(s - \alpha_2)\right)$$

with $\alpha_1 \equiv (1-\beta)D + d$ and $\alpha_2 \equiv d \leq \alpha_1$. Then (1) and its solution can be written as

$$\max_{s\in S} F(s,p)$$

and

$$s^*(p) \equiv \max\{\arg\max_{s\in S} F(s,p)\}$$

with $S \equiv [0, W_1]$ and $p \in P \equiv [0,1]$. To prove the proposition, we make use of the following theorem due to Topkis (1998):

Lemma 1 (Topkis' Theorem)**.** *Let $s^*(p)$ be as defined above. Let $S \subset \mathbb{R}$ be compact and P be a partially ordered set. If $F : S \times P \to \mathbb{R}$ is continuous in s and supermodular, then $s^*(p)$ exists and is weakly increasing.*

Clearly, in our case, $S = [0, W_1]$ is compact, $P = [0,1]$ is a partially ordered set, and $F(s,p)$ is continuous by the assumption that u_1 and u_2 are continuous. Thus, to prove the proposition, we only need to verify that $F(s,p)$ is supermodular. Indeed, $\forall s', s''$ with $s' < s''$,

$$\begin{aligned} F(s'',p) - F(s',p) \quad = \quad & u_1(W_1 - s'') - u_1(W_1 - s') + \delta[u_2(s'' - \alpha_2) - u_2(s' - \alpha_2) \\ & + p\left(u_2(s' - \alpha_2) - u_2(s' - \alpha_1) - (u_2(s'' - \alpha_2) - u_2(s'' - \alpha_1))\right)]. \end{aligned}$$

Since u_2 is increasing and weakly concave by assumption, $\alpha_1 \geq \alpha_2$, and $s'' > s'$,

$$u_2(s' - \alpha_2) - u_2(s' - \alpha_1) - (u_2(s'' - \alpha_2) - u_2(s'' - \alpha_1)) \geq 0,$$

and thus $F(s'',p) - F(s',p)$ is indeed weakly increasing in p. $F(s,p)$ is therefore supermodular, completing the proof. □

Proposition 1 shows that, due to risk aversion in the second period, high risk individuals save more than low risks given the same insurance contract. The proof makes use of monotone comparative static methods based on the supermodularity of expected utility (as given by equation (1)) in savings s and the damage risk p, which allow us to establish the result under our very general conditions (notably, no particular assumptions on the form of the first period utility function u_1, nor differentiability conditions are required).[5]

This positive effect of risk on savings (and hence accumulated wealth), holding fixed an insurance contract, is well documented empirically. The most direct evidence is provided in Engen and Gruber (1995), who use exogenous variation in unemployment insurance across states and time in the US to study the effect of insurance coverage and unemployment risk on financial wealth. While their main interest is in the precautionary savings effect (which predicts that individuals with less insurance coverage save more, controlling for their risk type), they also do an extension exercise that is of interest for our purpose. To test whether the size of the precautionary savings effect increases with risk, they estimate the regression equation

$$\text{wealth} = \alpha + \beta_1 \text{insurance} + \beta_2 \text{risk} + \beta_3 \text{insurance} \times \text{risk} + \gamma X + \varepsilon,$$

where X includes various controls and risk is a location-specific measure of the probability of being unemployed (equation 5 in Engen and Gruber (1995)). While they are concerned primarily about the precautionary savings effect β_1 and the interaction coefficient β_3, the effect characterized in Proposition 1, which describes how risk affects wealth, controlling for the insurance contract, is captured by β_2. Indeed, table 7 in Engen and Gruber (1995) shows that local unemployment risk has a significantly positive effect on net assets, and a positive but insignificant effect on gross assets.[6]

2.3 Irregular-Crossing Preferences

We now proceed to show how the endogeneity of wealth demonstrated in the preceding subsection affects an individual's willingness to pay for insurance. Substituting the optimal savings $s^*_{ij}(\beta, d)$ into the expected utility function yields the indirect utility function $V_{ij}(\beta, d)$ or $V_{ij}(C)$, which is continuous by Berge's maximum theorem, in-

[5] Using the same methods as in the proof of Proposition 1, it could be immediately shown that $s^*_{ij}(C)$ is also weakly increasing in the patience δ_i, but this result is less important for our following analysis.

[6] Similar evidence is available in other settings, for instance annuity markets, where individuals with a higher life-expectancy (which corresponds to a higher risk type) tend to accumulate more wealth, controlling for the degree of annuitization (see e.g. Brown (2001)). Brown and Finkelstein (2009) provide related evidence for long-term care insurance markets.

creasing in β and decreasing in d. Thus indifference curves in $\mathscr{C}$ and marginal rates of substitution can be obtained from V_{ij}, where it is important to notice that optimal savings will adjust as we move along an individual's indifference curve. Consider the following example.

Example 1 (Differentiable Utility). Assume that both u_1 and u_2 are continuously differentiable and strictly concave. Then, the necessary and sufficient condition for expected utility maximization (assuming that (1) has an interior solution) becomes

$$u_1'(W_1 - s_{ij}^*) = \delta_i \left[p_j u_2'(s_{ij}^* - (1-\beta)D - d) + (1-p_j)u_2'(s_{ij}^* - d)\right]. \tag{2}$$

(2) is a standard Euler equation stating that savings $s_{ij}^*(\beta, d)$ are determined so as to equalize expected discounted marginal utility across periods. Denote second period consumption in case of loss by $c_{2,ij}^0 = s_{ij}^* - (1-\beta)D - d$ and $c_{2,ij}^1 = s_{ij}^* - d$ otherwise. Then, the slope of an indifference curve of an individual with patience δ_i and risk p_j in contract (β, d) becomes,

$$\mathrm{MRS}_{ij}(\beta, d) = \left.\frac{\mathrm{d}d}{\mathrm{d}\beta}\right|_{V_{ij}=\bar{V}} = \frac{D p_j u_2'(c_{2,ij}^0)}{p_j u_2'(c_{2,ij}^0) + (1-p_j)u_2'(c_{2,ij}^1)} > 0. \tag{3}$$

For the differentiable model, we can use (3) to examine the crossing properties of different individuals' indifference curves at a given insurance contract, which are crucial for equilibrium outcomes. Let us first ignore differences in patience and consider individuals that only differ in their risk. In the standard adverse selection model where wealth is exogenous, high risks have a steeper indifference curve than low risks at any given contract. Put formally, the marginal rate of substitution between coverage and premium given in (3) is increasing in p_j. Clearly, the property immediately follows from (3) if savings are held fixed at the same level for all types. In that case, $c_{2,ij}^0$ and $c_{2,ij}^1$ do not depend on type and (3) is increasing in damage probability. By the following definition, we refer to this as "regular-crossing" of indifference curves.

Definition 1. *The indifference curves of two individuals that differ only in risk exhibit "regular-crossing" at a contract* (β, d) *if* $\mathrm{MRS}_{iL}(\beta, d) < \mathrm{MRS}_{iH}(\beta, d)$. *Otherwise, they exhibit "irregular-crossing".*

Definition 1 introduces a local concept at a given contract. If regular-crossing holds in the whole contract space $\mathscr{C}$, as it does in the Rothschild-Stiglitz model, it implies the global property of single crossing for indifference curves of two individuals that differ only in risk, which is crucial for determining equilibria in standard models. When savings are endogenous, however, this property does not necessarily hold. Savings vary with risk according to Proposition 1, with larger risks saving more. Then, if u_2 exhibits decreasing risk-aversion, high risks are less risk-averse, which ceteris paribus reduces their marginal rate of substitution (3). If this effect is strong enough, a high risk's indifference curve can be flatter than a low risk's, i.e. irregular-crossing can occur.[7] We proceed to explicitly demonstrate this possibility in a second example, based on stepwise linear utility functions. This setup allows us to derive optimal savings, indirect utility and indifference curves in closed form and thus to derive the underlying mechanism most transparently. In section 3, we will then use Example 2 to explicitly construct an equilibrium with empirically relevant properties.[8]

Example 2 (Stepwise Linear Utility). Utility in period 1 is given by

$$u_1(c) = \begin{cases} \alpha_3 c + (\alpha_1 - \alpha_2)\tilde{c}_1 + (\alpha_2 - \alpha_3)\tilde{c}_2 & \textit{if} \;\; \tilde{c}_2 < c \\ \alpha_2 c + (\alpha_1 - \alpha_2)\tilde{c}_1 & \textit{if} \;\; \tilde{c}_1 < c \leq \tilde{c}_2 \\ \alpha_1 c & \textit{if} \;\; c \leq \tilde{c}_1, \end{cases}$$

where $\tilde{c}_1 < \tilde{c}_2$ and $\alpha_i > 0,\ i = 1, 2, 3$. Utility in period 2 is given by

$$u_2(c) = \begin{cases} \gamma_3 c + (\gamma_1 - \gamma_2)\bar{c}_1 + (\gamma_2 - \gamma_3)\bar{c}_2 & \textit{if} \;\; \bar{c}_2 < c \\ \gamma_2 c + (\gamma_1 - \gamma_2)\bar{c}_1 & \textit{if} \;\; \bar{c}_1 < c \leq \bar{c}_2 \\ \gamma_1 c & \textit{if} \;\; c \leq \bar{c}_1, \end{cases}$$

where $\bar{c}_1 < \bar{c}_2$ and $0 < \gamma_3 < \gamma_2 < \gamma_1$. These utility functions satisfy all our previous assumptions. We further assume w.l.o.g. that $\delta_H = 1$. The model is then completely specified by a collection of parameters

$$P = ((\alpha_k, \gamma_k)_{k=1,2,3}, (\bar{c}_k, , \tilde{c}_k)_{k=1,2}, n_{LL}, n_{HL}, n_{LH}, n_{HH}, \delta_L, p_L, p_H, D)\,.$$

[7] In the labor supply model, this would correspond to high risks working more than low risks. The resulting higher labor income reduces risk aversion, which can lead to irregular-crossing.

[8] The possibility of irregular-crossing could also be illustrated within a simpler example with utility functions that consist of only two linear segments and thus one kink. The more general structure of Example 2, where preferences have two kinks, however, is needed for our equilibrium construction in section 3, which is why we introduce it here already. In general, the kinked utility functions are used for analytical convenience only, and should be thought of as an approximation for a differentiable model.

The complete analysis of the model can be found in Appendix A. There, we identify the following set of conditions under which the model is most easily tractable:

Assumption 1. *The parameter collection* P *satisfies*

(i) $\alpha_1 < \alpha_2 < \alpha_3$*,*

(ii) $\bar{c}_2 - \bar{c}_1 > D$*,*

(iii) $p_H\gamma_1 + (1-p_H)\gamma_2 < \alpha_3 < \delta_L\gamma_1$*,*

(iv) $p_H\gamma_2 + (1-p_H)\gamma_3 < \alpha_2 < \delta_L\gamma_2$*, and*

(v) $\gamma_3 < \alpha_1 < p_L\delta_L\gamma_2 + (1-p_L)\delta_L\gamma_3$*.*

Let $\mathscr{P}$ denote the set of parameter collections P that satisfy Assumption 1, which is shown to be nonempty in Appendix A. As we also show there, Assumption 1 implies that, in any contract (β, d) and for all four types, there are only three savings levels that could potentially be optimal: $s^1 = \bar{c}_1 + d$, $s^2 = \bar{c}_2 + d$, and $s^3 = \bar{c}_2 + d + (1-\beta)D$, with $s^1 \le s^2 \le s^3$. Different types might, however, find a different choice among s^k, $k = 1, 2, 3$, optimal in contract (β, d). To determine $s^*_{ij}(\beta, d)$, we only need to compare the premium d to three threshold values

$$d^{12}_{ij}(\beta),\ d^{23}_{ij}(\beta) \text{ and } d^{13}_{ij}(\beta),$$

which are given by equations (11) - (13) in the Appendix. If, for example, $d < d^{12}_{ij}(\beta)$, then type ij prefers s^1 over s^2, while s^2 yields a larger utility than s^1 for ij if $d \ge d^{12}_{ij}(\beta)$. Type ij's preference between the remaining two pairs of potentially optimal savings are computed analogously using $d^{23}_{ij}(\beta)$ and $d^{13}_{ij}(\beta)$, and the globally optimal savings level can easily be determined. Specifically, since $d \ge d^{km}_{ij}(\beta)$, $k \in \{1, 2\}$, $m > k$, always implies that the larger savings s^m are preferred over the smaller s^k, $s^*_{ij}(\beta, d)$ is weakly increasing in the premium d.

As predicted by Proposition 1, we show in the Appendix that the three critical values $d^{12}_{ij}(\beta)$, $d^{23}_{ij}(\beta)$ and $d^{13}_{ij}(\beta)$ are all decreasing in individual damage risk, and strictly so for contracts with less than full coverage ($\beta < 1$). Thus savings are (weakly) increasing in risk: if we fix a contract (β, d) with $\beta < 1$ and increase an individual's damage probability, the critical values will eventually fall below d and savings will jump upwards.

The marginal rate of substitution $\mathrm{MRS}_{ij}(\beta, d)$ then depends on whether s^1, s^2 or s^3 is optimal for ij in (β, d). For example, if $s^*_{ij}(\beta, d) = s^1$, we obtain (see Appendix A for the derivations)

$$\mathrm{MRS}_{ij}(\beta) = \mathrm{MRS}^1_{ij}(\beta) \equiv \frac{\delta_i\gamma_1}{\alpha_3}p_jD. \tag{4}$$

The marginal rates resulting from the other two savings levels are

$$\text{MRS}_{ij}^{2}(\beta, d) = \frac{\delta_i \gamma_2}{\alpha_2} p_j D \tag{5}$$

and

$$\text{MRS}_{ij}^{3}(\beta, d) = \left(1 - (1 - p_j)\frac{\delta_i \gamma_3}{\alpha_1}\right) D. \tag{6}$$

Closer inspection of (4) - (6) reveals our main insight concerning irregular-crossing. First, each of the expressions is strictly increasing in damage probability p. Thus, whenever two types iL and iH that differ only in risk find the same level of savings optimal in a contract (β, d), regular-crossing holds. Translated to a differentiable model, this corresponds to a case where the high risk does not save much more than the low risk, so that the standard risk effect dominates. If the two individuals differ in both risk and patience, any of them can of course have the steeper indifference curve, because the marginal rate is affected by both p_j and δ_i. This corresponds to a standard case of countervailing incentives due to exogenous two-dimensional heterogeneity, as in Smart (2000). Given our arguments from above, however, it is now also possible that iL prefers a smaller level of savings than iH in (β, d). Suppose, for example, that $s_{iL}^{*}(\beta, d) = s^1$ and $s_{iH}^{*}(\beta, d) = s^2$ because $d_{iH}^{12}(\beta) < d < d_{iL}^{12}(\beta)$, which is possible since $d_{iH}^{12}(\beta) < d_{iL}^{12}(\beta)$. We then need to compare (4) to (5) and find irregular-crossing $\text{MRS}_{iH}^{2}(\beta, d) \leq \text{MRS}_{iL}^{1}(\beta, d)$ if

$$p_L \frac{\gamma_1}{\gamma_2} \geq p_H \frac{\alpha_3}{\alpha_2},$$

which can be satisfied in addition to Assumption 1. Figure 1 illustrates this case. It depicts the contract space together with the zero-profit lines for the two risks (dotted lines). The *regime switching lines* $d_{iL}^{12}(\beta)$ and $d_{iH}^{12}(\beta)$ are also depicted, with a contract C that lies in between them and in which irregular-crossing occurs. Translated into a model with differentiable utility, this corresponds to the case where the high risk saves sufficiently more than the low risk to have a lower marginal rate of substitution. Proceeding similarly, conditions (20) and (21) in the Appendix imply irregular-crossing if $s_{iL}^{*}(\beta) = s^1$, $s_{iH}^{*}(\beta) = s^3$ or $s_{iL}^{*}(\beta) = s^2$, $s_{iH}^{*}(\beta) = s^3$, respectively.

3 Equilibrium Analysis

In this section, we aim at exploring the implications of irregular-crossing preferences, which cannot arise in models of adverse selection in insurance markets considered pre-

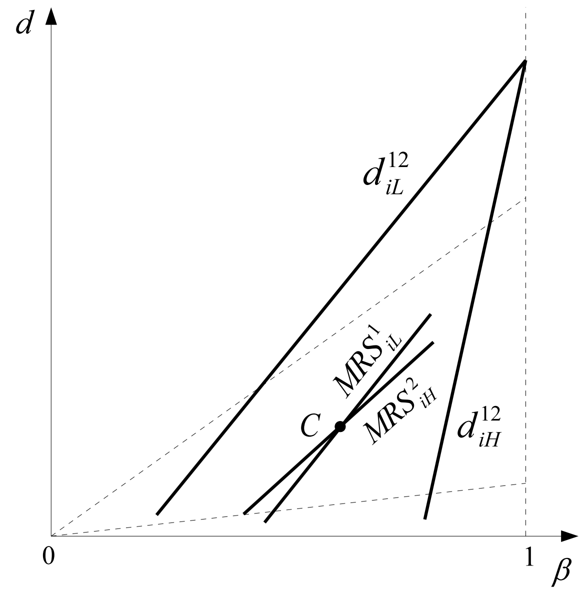

Figure 1: Irregular-Crossing.

viously, for the set of screening equilibria in our model.

3.1 Equilibrium Definition

In line with much of the insurance literature, we use the Rothschild-Stiglitz equilibrium concept, which is directly based on the contracts that are offered by the firms. This equilibrium can be thought of as arising in a game with an additional period $t = 0$ in which risk-neutral insurance companies can each offer one contract $(\beta, d) \in \mathscr{C}$.[9] The expected profit of such a contract if it is subsequently purchased by a of the low-risk and b of the high-risk individuals is given by

$$\pi(\beta, d, a, b) = a[d - p_L \beta D] + b[d - p_H \beta D]. \tag{7}$$

For any finite set $\mathscr{C}' = \{C_1, ..., C_N\} \subset \mathscr{C}$ of contracts, where $C_k = (\beta_k, d_k)$, let $\pi_{\mathscr{C}'}(C_k)$ be the profit of contract C_k if all individuals choose their preferred contract from $\mathscr{C}'$ (together with their simultaneously determined optimal savings). In case of indifference,

[9]See, for example, Netzer and Scheuer (2009) for a new approach to modelling insurance markets with adverse selection game-theoretically. The Rothschild-Stiglitz contracts can be the equilibrium outcome in their setting. Recent work by Dubey and Geanakoplos (2002) and Bisin and Gottardi (2006) has shown that Rothschild-Stiglitz type equilibria can also be generated by competitive equilibrium models rather than the original game-theoretic setup.

we assume the contract with larger coverage to be chosen. We then define an equilibrium as follows:

Definition 2. *The set of contracts $\mathscr{C}'$ is an equilibrium if*
(i) $\pi_{\mathscr{C}'}(C_k) \geq 0 \ \forall C_k \in \mathscr{C}'$, *and*
(ii) $\nexists C \in \mathscr{C} \backslash \mathscr{C}'$ *with* $\pi_{\mathscr{C}' \cup \{C\}}(C) > 0$.

The definition goes back to Rothschild and Stiglitz (1976). In their model, all equilibrium contracts must earn zero profits. Already in the two-dimensional model of Smart (2000), and also in our model with endogenous wealth, equilibrium contracts can earn strictly positive profits. Competition does not eliminate such contracts, because any contract that is slightly more attractive to the consumers would also attract bad risk types and become unprofitable. As a result, the non-increasing profits property of Chiappori, Jullien, Salanié, and Salanié (2006) can be violated in our model, which, as we will demonstrate in the following, allows for equilibria with an imperfect ordering of risk and coverage.[10]

Before we move to the equilibrium construction in the next section, it needs to be emphasized that we actually need two dimensions of unobservable heterogeneity (in risk *and* patience / productivity) to obtain this main result. Even with endogenous wealth accumulation, low risk individuals will never purchase more coverage than high risks in any equilibrium of a model with one-dimensional heterogeneity in risk only. In Appendix B, this claim is stated and proven formally for the very general setup introduced in Section 2. Intuitively, the countervailing incentives arising from endogenous wealth accumulation need to be reinforced by patience differences to reverse the usual ordering of risk and coverage in equilibrium. We therefore move to a richer (and more realistic) model with multidimensional heterogeneity. However, because exogenous heterogeneity in wealth only generates equilibria with a positive correlation of risk and coverage (as shown by Smart 2000, Wambach 2000 and Villeneuve 2003), endogeneity is still the crucial ingredient that drives our result.

3.2 Equilibrium with Imperfect Ordering of Risk and Coverage

In this section, we explicitly construct an equilibrium to illustrate our main point. In the next section, we discuss the plausibility of this equilibrium, in particular with respect

[10] See Smart (2000) for a way to reconcile positive profits with possibly unlimited entry by new firms, based on strictly positive but vanishing entry costs. De Meza and Webb (2001) propose an additional approach to deal with potential problems from positive profits in perfect competition.

to its empirical predictions beyond the non-monotone ordering of risk and coverage. However, we do not attempt a full characterization of all possible equilibria in the two-dimensional framework. Such a task would be associated with substantial complications in our very general model. For example, separating or pooling equilibria as those identified by Smart (2000) could also exist in our model.

Indeed, our construction builds on Smart (2000), who presents a pooling equilibrium in which individuals with high risk and high risk-aversion (corresponding to our type LH with low patience and high risk) obtain an actuarially fair contract with full coverage, and individuals with high risk / low risk-aversion and low risk / high risk-aversion (corresponding to our types HH and LL) are pooled in a contract with partial coverage (p. 162). The individuals with low risk and low risk aversion (our type HL) cannot be part of the pool, because they could always be attracted to a contract offer with smaller coverage and premium, due to regular-crossing. Hence they must obtain a contract with smaller coverage than the pool in equilibrium, which establishes the positive correlation property.

Our equilibrium is similar in that the LH type also obtains the fair full coverage contract and that there will be a pooling contract purchased by two different risk types. Irregular-crossing between the LH and the LL type in this pooling contract, however, now implies that the LL (not HL) type can be attracted away from the pool by a contract with *larger* coverage and premium, because this type saves little and is very risk-averse. In the equilibrium, the LL type obtains more coverage than the pool of the HH and HL types, which breaks the monotone ordering logic. Consider the preferences given in Example 2. The above argument translates into the following proposition:

Proposition 2. *Let $\mathscr{C}' = \{C_1, C_2, C_3\} \subset \mathscr{C}$ be given by*

$$\begin{aligned}
C_1 &= (1, p_H D),\\
C_2 &= \arg\max_{(\beta,d)\in\mathscr{C}} \; V_{HL}(\beta, d) \quad s.t. \quad (i)\; V_{LH}(C_1) = V_{LH}(\beta, d),\\
&\qquad\qquad\qquad\qquad\qquad\qquad\quad (ii)\; \pi(\beta, d, n_{HL}, n_{HH}) \geq 0,\\
C_3 &= \arg\max_{(\beta,d)\in\mathscr{C}} \; V_{LL}(\beta, d) \quad s.t. \quad (i)\; V_{LH}(C_1) = V_{LH}(\beta, d).
\end{aligned}$$

Then, there exists a nonempty open subset $\mathscr{P}' \subset \mathscr{P}$ such that, for all $P \in \mathscr{P}'$, it holds that $C_2 < C_3$ and $\mathscr{C}'$ is an equilibrium.

Proof. See Appendix C (Section 5.2). □

The equilibrium is depicted in Figure 2. Type LH prefers contract C_1, types HL and HH prefer C_2, and LL prefers C_3. It can be seen from the figure that (i) the con-

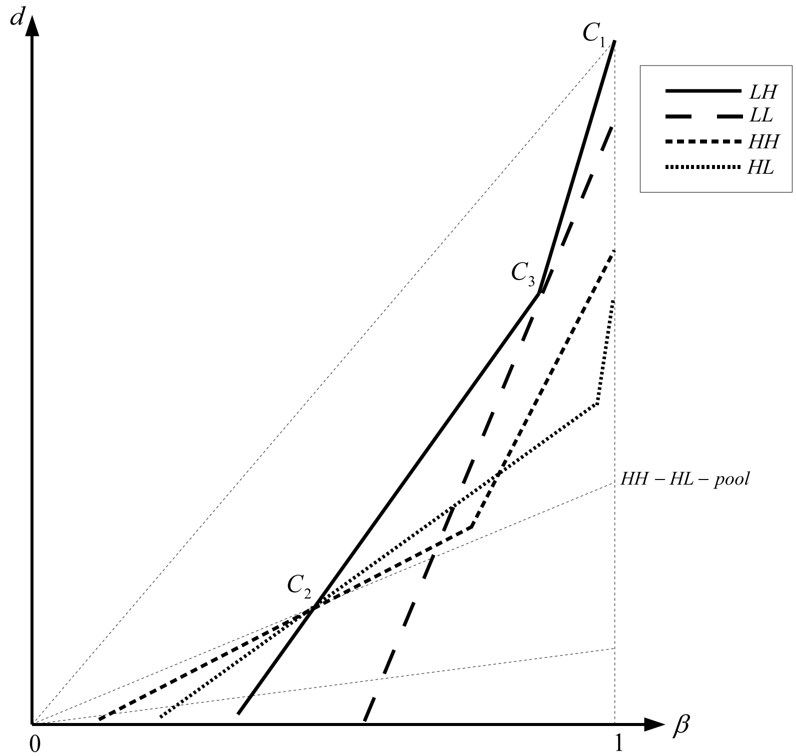

Figure 2: Equilibrium.

tracts C_1, C_2 and C_3 all earn non-negative profits (specifically, C_3 earns strictly positive profits), and (ii) there is no contract outside of $\{C_1, C_2, C_3\}$ that would attract low risks only and therefore earn positive profits for sure. Clearly, contracts offered in addition to $\{C_1, C_2, C_3\}$ would attract different subgroups of the population, according to their specific location. But these subgroups will necessarily include large risks and no deviation is profitable for an appropriate composition of the population. The assumptions assuring this are a multi-dimensional generalization of the Rothschild-Stiglitz existence condition that requires that there are not too many low risks. See Appendix C for a complete specification of the existence conditions in our framework.

Most important is the fact that C_3, purchased by low risks that are impatient, has a larger coverage than the pooling contract C_2, which is purchased by patient high risks. Hence there is no longer the perfect ordering of risk and coverage that the models with exogenous heterogeneity predict. As Figure 2 makes clear, the non-increasing profits assumption by Chiappori, Jullien, Salanié, and Salanié (2006) is not satisfied in the equilibrium of Proposition 2. In fact, C_3 will always make more profits per capita than C_2 although $C_3 > C_2$, since it is only bought by low risks. Thus, the non-increasing profits property cannot be considered as a general characteristic of equilibrium in competitive insurance markets as soon as multidimensional heterogeneity and unobserved actions such as labor supply or savings are accounted for. Its violation is the key reason why the argument by Chiappori, Jullien, Salanié, and Salanié (2006) does not apply in our framework.

3.3 Evidence on the Underlying Mechanism

The equilibrium is consistent with the empirical evidence that a monotone relation between risk and coverage is not observed although adverse selection seems to be a relevant phenomenon in insurance markets. Technically, it is the possibility of irregular-crossing preferences which is crucial for generating such equilibria. Notably, deviations from perfect competition are not necessary to obtain the result.

On the other hand, however, our model predicts that there should again be a positive correlation between risk and coverage when conditioning on income/wealth or on more direct measures of the second dimension of private information, i.e. patience and productivity. This pattern of (conditional) correlations is empirically well-supported. Notably, Fang, Keane, and Silverman (2008) consider the US Medigap market (a market for supplemental private health insurance covering spending excluded by Medicare), and first show that there is a significantly negative correlation between risk and coverage on average (table 2, panel A in their paper). The corresponding regression is

$$\text{coverage} = \alpha + \beta_1 \text{risk} + \beta_2 D + \varepsilon,$$

where risk is measured by total medical expenditure and D are variables known to insurers and used to price contracts (equation 6 in their paper). β_1 is indeed significantly negative (Table 6 rows 1 and 9). The authors then investigate the sources of this negative correlation by considering various conditional correlations. This is done by estimating the regression equation

$$\text{coverage} = \alpha + \beta_1 \text{risk} + \beta_2 D + \beta_3 X + \varepsilon,$$

where X are now variables that measure the second dimension of private information (i.e. income/wealth or productivity/patience in our framework). Specifically, they include measures of an individual's risk aversion, income, education and cognitive ability, which relate to productivity, and longevity expectation and financial planning horizon, which capture patience (see equation 7 and table 6 in their paper). The authors add more and more of the above-mentioned variables in X and show that the originally negative (unconditional) correlation β_1 becomes a more and more positive conditional correlation (table 6 rows 2-8 and 10-16).[11]

4 Conclusion

Based on recent empirical findings, the theoretical literature on adverse selection has started to realize that screening in most relevant real-world situations is associated with more than one dimension of privately known heterogeneity, and that the resulting countervailing incentives significantly alter the nature of equilibrium compared to the standard model of Rothschild and Stiglitz (1976). These models typically assume that all dimensions of heterogeneity are given exogenously, and they predict a positive correlation between risk occurrence and insurance coverage in equilibrium. This is contradicted by empirical studies which mostly observe a zero or negative correlation.

In this paper, we asked the question how insurance market equilibrium looks like if heterogeneity in some dimensions is not given exogenously but arises from the individuals' choices. As a natural example of such a situation, we considered a model where individuals not only differ in risk and select an insurance contract, but also choose their savings or labor supply endogenously, which affects their wealth and hence risk attitude. This allows for irregular-crossing of preferences in the sense that, among individuals who exogenously *only* differ in risk, high risk individuals have the *lower* marginal willingness to pay for insurance than low risks since they save more or supply more labor and hence are less risk averse.

We show that this can lead to equilibria with positive profit contracts that violate the non-increasing profits property of Chiappori, Jullien, Salanié, and Salanié (2006), and therefore exhibit a non-monotone relation between risk and coverage. Interestingly, this latter result provides an explanation for the empirical findings without assuming non-competitive insurance markets or imposing restrictions on the structure of heterogeneity. In addition, the model predicts that the correlation between risk and coverage should

[11] Finkelstein and McGarry (2003) also provide related evidence for the market for the long-term care insurance. They also find an overall negative correlation between risk and coverage (their equation 2 and table 4, column 1). This also holds when controlling for variables that insurers use for pricing, notably age (columns 2 and 4). However, when they control for income quartile and asset quartile in addition to age (see p. 17 and table 3 in their paper), it turns out that the originally significantly negative correlation then first vanishes and becomes positive (see column 3 of table 4 in their paper).

turn positive when conditioning on the second dimension of heterogeneity, which is also well-documented empirically.

Our model raises a number of issues for further research. First, our informational assumption that risk, patience and savings are privately known by the individuals may make our model a helpful tool for the analysis of policy questions such as taxation under risk. Models addressing these issues need to combine multidimensional heterogeneity with the endogenous choice of private insurance and wealth.[12] Natural questions to ask are about the effects of taxes or social insurance in this framework, and, more generally, about the efficiency properties of the equilibria that arise in our model. While the efficiency properties of equilibria with one-dimensional private information have received much attention,[13] little is known about the existence of Pareto-improving policies in multidimensional screening models.

Furthermore, as pointed out above, the possibility of irregular-crossing is the driving force behind our novel results. Our model is just one - though certainly natural - example of a situation where irregular-crossing can arise. The results extend, however, to other settings. Generally, irregular-crossing preferences can result from some unobserved decision that does not affect the agent's risk but risk aversion. This may not only be relevant in models of insurance, but also of credit markets, portfolio choice, or labor contracts.

References

Armstrong, M., and J.-C. Rochet (1999): "Multidimensional Screening: A User's Guide," *European Economic Review*, 43, 959–979.

Bisin, A., and P. Gottardi (2006): "Efficient Competitive Equilibria with Adverse Selection," *Journal of Political Economy*, 114, 485–516.

Boadway, R., M. Leite-Monteiro, M. Marchand, and P. Pestieau (2006): "Social Insurance and Redistribution with Moral Hazard and Adverse Selection," *Scandinavian Journal of Economics*, 108, 279–298.

Brown, J. R. (2001): "Private Pensions, Mortality Risk, and the Decision to Annuitize," *Journal of Public Economics*, 82, 29–62.

[12]For recent studies in this direction, but typically with a more restrictive treatment of multidimensional heterogeneity than in the present paper, see Boadway, Leite-Monteiro, Marchand, and Pestieau (2006), Netzer and Scheuer (2007) and Chetty and Saez (2008).

[13]See, for instance, Wilson (1977), Crocker and Snow (1985), Hellwig (1987) and Gale (1996).

Brown, J. R., and A. N. Finkelstein (2009): “The Private Market for Long-Term Care Insurance in the United States: A Review of the Evidence,” *Journal of Risk and Insurance*, 76, 5–29.

Cardon, J., and I. Hendel (2001): “Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey,” *RAND Journal of Economics*, 32, 408–427.

Cawley, J., and T. Philipson (1999): “An Empirical Examination of Information Barriers to Trade in Insurance,” *American Economic Review*, 89, 827–846.

Chassagnon, A., and P. Chiappori (1997): “Insurance under Moral Hazard and Adverse Selection: The Case of Pure Competition,” *Discussion Paper, DELTA, Paris.*

Chetty, R., and E. Saez (2008): “Optimal Taxation and Social Insurance with Endogenous Private Insurance,” *NBER Working Paper No. 14403.*

Chiappori, P., B. Jullien, B. Salanié, and F. Salanié (2006): “Asymmetric Information in Insurance: General Testable Implications,” *RAND Journal of Economics, forthcoming.*

Chiappori, P., and B. Salanié (2000): “Testing for Asymmetric Information in Insurance Markets,” *Journal of Political Economy*, 108, 56–78.

Crocker, K., and A. Snow (1985): “The Efficiency of Competitive Equilibria in Insurance Markets with Asymmetric Information,” *Journal of Public Economics*, 26, 207–219.

De Meza, D., and D. Webb (2001): “Advantageous Selection in Insurance Markets,” *RAND Journal of Economics*, 32, 249–262.

Dubey, P., and J. Geanakoplos (2002): “Competitive Pooling: Rothschild-Stiglitz Reconsidered,” *Quarterly Journal of Economics*, 117, 1529–1570.

Engen, E. M., and J. Gruber (1995): “Unemployment Insurance and Precautionary Saving,” *NBER Working Paper 5252.*

Fang, H., M. P. Keane, and D. Silverman (2008): “Sources of Advantageous Selection: Evidence from the Medigap Insurance Market,” *Journal of Political Economy*, 116, 303–350.

Finkelstein, A., and K. McGarry (2003): "Private Information and its Effect on Market Equilibrium: New Evidence from Long-term Care Insurance," *NBER Working Paper 9957*.

——— (2006): "Multiple Dimensions of Private Information: Evidence from the Long-term Care Insurance Market," *American Economic Review*, 96, 938–958.

Gale, D. (1996): "Equilibria and Pareto Optima of Markets with Adverse Selection," *Economic Theory*, 7, 207–235.

Hellwig, M. (1987): "Some Recent Developments in the Theory of Competition in Markets with Adverse Selection," *European Economic Review*, 31, 319–325.

Hemenway, D. (1990): "Propitious Selection," *Quarterly Journal of Economics*, 105, 1063–1069.

——— (1992): "Propitious Selection in Insurance Markets," *Journal of Risk and Uncertainty*, 5, 247–251.

Jullien, B., B. Salanié, and F. Salanié (2007): "Screening Risk-Averse Agents under Moral Hazard: Single-Crossing and the CARA Case," *Economic Theory*, 30, 151–169.

Lewis, T., and D. Sappington (1989): "Countervailing incentives in agency problems," *Journal of Economic Theory*, 49, 294–313.

Netzer, N., and F. Scheuer (2007): "Taxation, Insurance and Precautionary Labor," *Journal of Public Economics*, 91, 1519–1531.

——— (2009): "Competitive Markets Without Commitment," *SOI Discussion Paper No. 0814*.

Rochet, J.-C., and P. Chone (1998): "Ironing, Sweeping and Multidimensional Screening," *Econometrica*, 66, 783–826.

Rothschild, M., and J. Stiglitz (1976): "Equilibrium in Competitive Insurance Markets: An Essay in the Economics of Incomplete Information," *Quarterly Journal of Economics*, 90, 629–649.

Smart, M. (2000): "Competitive Insurance Markets with Two Unobservables," *International Economic Review*, 41, 153–169.

STEWART, J. (1994): "The Welfare Implications of Moral Hazard and Adverse Selection in Competitive Insurance Markets," *Economic Inquiry*, 32, 193–208.

TOPKIS, D. (1998): *Supermodularity and Complementarity.* Princeton University Press, Princeton, USA.

VILLENEUVE, B. (2003): "Concurrence et antisélection multidimensionelle en assurance," *Annales d'Economie et de Statistique*, 69, 119–142.

WAMBACH, A. (2000): "Introducing Heterogeneity in the Rothschild-Stiglitz Model," *Journal of Risk and Insurance*, 67, 579–592.

WILSON, C. (1977): "A Model of Insurance Markets with Incomplete Information," *Journal of Economic Theory*, 12, 167–207.

5 Appendix

5.1 Appendix A: Proofs for Example 2 (irregular-crossing)

5.1.1 Utility and Savings

Fix a contract $(\beta, d) \in \mathscr{C}$. We are interested in the optimal savings $s^*_{ij}(\beta, d)$ of type ij when preferences are given as in Example 2. First, the maximization problem (1) can be rewritten as

$$\max_{s_{ij}} \delta_i[p_j u_2(s_{ij} - (1-\beta)D - d) + (1-p_j)u_2(s_{ij} - d)] - [-u_1(W_1 - s_{ij})],$$

i.e. as the maximization of the difference between expected discounted utility from period 2 and the negative of utility in period 1. This problem is illustrated in Figure 3, for the two types HL and HH. The lower dashed line (long dashes) represents $-u_1(W_1 - s)$, shifted upwards for graphical convenience, as a function of savings s. The slopes of $-u_1(W_1 - s)$ are indicated in the figure (α_1, α_2 and α_3). The other two lines represent the expected period 2 utility for the low risk (upper dashed line, short dashes) and the high risk type (solid line), respectively. Slopes are again given in the figure.

Several additional assumptions are implicit in Figure 3. First, $\alpha_1 < \alpha_2 < \alpha_3$ holds. Second, it is assumed that $\bar{c}_1 + d + (1-\beta)D < \bar{c}_2 + d$. This holds for every contract (β, d) if $\bar{c}_2 - \bar{c}_1 > D$. Under this assumption, the kinks of the expected period 2 utility functions are always ranked as depicted in the graph.[14] The figure finally reflects the following three assumptions:

(i) $p_H\gamma_1 + (1-p_H)\gamma_2 < \alpha_3 < \gamma_1$,

(ii) $p_H\gamma_2 + (1-p_H)\gamma_3 < \alpha_2 < \gamma_2$, and

[14]Fixing a coverage β, varying the premium d then simply amounts to shifting the period 2 utility functions, or, more easily, the function $-u_1(W_1 - s)$ left or right.

(iii) $\gamma_3 < \alpha_1 < p_L\gamma_2 + (1 - p_L)\gamma_3$.

For example, property (i) implies that the first segment of $-u_1(W_1 - s)$ is flatter (with a slope of α_3) than the initial segment of both types' period 2 expected utility function (with a slope of γ_1), but steeper than their second segment (again for *both* types), which has the slopes $p_L\gamma_1 + (1 - p_L)\gamma_2$ and $p_H\gamma_1 + (1 - p_H)\gamma_2$, respectively. Properties (ii) and (iii) are analogous for the remaining segments.[15] Altogether, they imply that there are only three candidates for optimal savings at a contract (β, d), for both HH and HL: $s^1 = \bar{c}_1 + d$, $s^2 = \bar{c}_2 + d$, and $s^3 = \bar{c}_2 + d + (1 - \beta)D$, where $s^1 \le s^2 \le s^3$.

This argument applies to all four types simultaneously (not only to HH and HL) whenever conditions equivalent to (i)-(iii) are also satisfied for the low patient types. These modified conditions are obtained from (i)-(iii) by replacing any γ_i by $\delta_L\gamma_i$, for $i = 1, 2, 3$. It is actually possible to satisfy the original and the modified conditions simultaneously, because the latter are automatically satisfied if the former are, for sufficiently large $\delta_L < 1$. We now summarize all these assumptions as follows:

Assumption A1.

(i) $\alpha_1 < \alpha_2 < \alpha_3$,

(ii) $\bar{c}_2 - \bar{c}_1 > D$,

(iii) $p_H\gamma_1 + (1 - p_H)\gamma_2 < \alpha_3 < \delta_L\gamma_1$,

(iv) $p_H\gamma_2 + (1 - p_H)\gamma_3 < \alpha_2 < \delta_L\gamma_2$, and

(v) $\gamma_3 < \alpha_1 < p_L\delta_L\gamma_2 + (1 - p_L)\delta_L\gamma_3$.

From Figure 3 it is also possible to see that s^1 can only be optimal if $s^1 = \bar{c}_1 + d < W_1 - \tilde{c}_2$. Otherwise, increasing s from s^1 would increase the distance between expected period 2 utility and $-u_1$. Similarly, s^2 can only be optimal if $W_1 - \tilde{c}_2 < s^2 < W_1 - \tilde{c}_1$, and s^3 can only be optimal if $s^3 > W_1 - \tilde{c}_1$. This now allows us to compute the overall (lifetime) utilities from the three candidate savings as follows:

$$U_{ij}(s^1) = \alpha_3[W_1 - \bar{c}_1 - d] + (\alpha_1 - \alpha_2)\tilde{c}_1 + (\alpha_2 - \alpha_3)\tilde{c}_2 + \delta_i\gamma_1\bar{c}_1 - p_j\delta_i\gamma_1(1 - \beta)D, \tag{8}$$

$$U_{ij}(s^2) = \alpha_2[W_1 - \bar{c}_2 - d] + (\alpha_1 - \alpha_2)\tilde{c}_1 + \delta_i(\gamma_1 - \gamma_2)\bar{c}_1 + \delta_i\gamma_2\bar{c}_2 - p_j\delta_i\gamma_2(1 - \beta)D, \tag{9}$$

$$U_{ij}(s^3) = \alpha_1[W_1 - \bar{c}_2 - d - (1 - \beta)D] + \delta_i(\gamma_1 - \gamma_2)\bar{c}_1 + \delta_i\gamma_2\bar{c}_2 + \delta_i\gamma_3(1 - p_j)(1 - \beta)D. \tag{10}$$

5.1.2 Savings Regimes

Comparing the utility levels (8) - (10) pairwise, we can derive lines that separate the areas of the contract space in which one savings level yields a greater utility level than the other, for a type ij. First, a comparison between $U_{ij}(s^1)$ and $U_{ij}(s^2)$ yields the regime border

$$d_{ij}^{12}(\beta) \equiv W_1 - \tilde{c}_2 - \left(\frac{(\alpha_3 - \delta_i\gamma_2)\bar{c}_1 + (\delta_i\gamma_2 - \alpha_2)\bar{c}_2}{\alpha_3 - \alpha_2}\right) - \delta_i\left(\frac{\gamma_1 - \gamma_2}{\alpha_3 - \alpha_2}\right)p_j(1 - \beta)D, \tag{11}$$

[15] Taking these and the previous assumptions together implies that $\gamma_3 < \alpha_1 < \alpha_2 < \gamma_2 < \alpha_3 < \gamma_1$.

so that $U_{ij}(s^1) > U_{ij}(s^2)$ if $d < d_{ij}^{12}(\beta)$ and vice versa. Analogously, comparing s^2 to s^3 yields

$$d_{ij}^{23}(\beta) \equiv W_1 - \tilde{c}_1 - \bar{c}_2 - \left(\frac{p_j\delta_i\gamma_2 + (1-p_j)\delta_i\gamma_3 - \alpha_1}{\alpha_2 - \alpha_1}\right)(1-\beta)D, \tag{12}$$

where s^2 is preferred over s^3 if $d < d_{ij}^{23}(\beta)$. Finally, we obtain

$$\begin{aligned} d_{ij}^{13}(\beta) \equiv \quad W_1 \quad & - \left(\frac{(\alpha_3 - \delta_i\gamma_2)\bar{c}_1 + (\alpha_2 - \alpha_1)\tilde{c}_1 + (\delta_i\gamma_2 - \alpha_1)\bar{c}_2 + (\alpha_3 - \alpha_2)\tilde{c}_2}{\alpha_3 - \alpha_1}\right) \\ & - \left(\frac{p_j\delta_i\gamma_1 + (1-p_j)\delta_i\gamma_3 - \alpha_1}{\alpha_3 - \alpha_1}\right)(1-\beta)D, \end{aligned} \tag{13}$$

where s^1 is preferred over s^3 when $d < d_{ij}^{13}(\beta)$.

For any given parameter constellation satisfying Assumption A1, the optimal savings of type ij in a contract can now be determined by comparing the location of the contract relative to the (linear) regime borders d_{ij}^{12}, d_{ij}^{23} and d_{ij}^{13}. First, fix the coverage β. Due to transitivity of preferences, it must be true that either $d_{ij}^{12}(\beta) \leq d_{ij}^{13}(\beta) \leq d_{ij}^{23}(\beta)$ or $d_{ij}^{23}(\beta) \leq d_{ij}^{13}(\beta) \leq d_{ij}^{12}(\beta)$. In the first case, we obtain

$$s_{ij}^*(\beta, d) = \begin{cases} s^3 & \text{if} \quad d_{ij}^{23}(\beta) \leq d, \\ s^2 & \text{if} \quad d_{ij}^{12}(\beta) \leq d < d_{ij}^{23}(\beta), \\ s^1 & \text{if} \quad d < d_{ij}^{12}(\beta), \end{cases} \tag{14}$$

i.e. the regime border d_{ij}^{13} can be ignored when determining optimal savings.[16] In the second case we have

$$s_{ij}^*(\beta, d) = \begin{cases} s^3 & \text{if} \quad d_{ij}^{13}(\beta) \leq d, \\ s^1 & \text{if} \quad d < d_{ij}^{13}(\beta), \end{cases} \tag{15}$$

i.e. only the regime border d_{ij}^{13} is relevant and s^2 is never optimal. Observe that, in any case, s_{ij}^* is weakly increasing in the premium d.

5.1.3 Marginal Rates of Substitution

We now derive the marginal rates of substitution between coverage and premium for type ij. These rates exist except for contracts on one of the regime borders, where the indifference curves are not differentiable but kinked. We have to distinguish three cases, according to the optimal savings level in a contract (β, d). For $s_{ij}^*(\beta, d) = s^1$ we obtain from the above formulation (8) of $U_{ij}(s^1)$ that

$$\text{MRS}_{ij}^1(\beta, d) = \frac{\delta_i\gamma_1}{\alpha_3} p_j D. \tag{16}$$

Analogously, we have

$$\text{MRS}_{ij}^2(\beta, d) = \frac{\delta_i\gamma_2}{\alpha_2} p_j D \tag{17}$$

and

$$\text{MRS}_{ij}^3(\beta, d) = \left(1 - (1-p_j)\frac{\delta_i\gamma_3}{\alpha_1}\right) D. \tag{18}$$

[16] Specifically, this implies that d_{ij}^{13} can always be ignored whenever d_{ij}^{12} lies below d_{ij}^{23} in the whole contract space.

When comparing the marginal rates of substitution at some contract (β, d) of two individuals that differ only in risk, two possible cases can occur. First, both individuals might find the same savings regime optimal and hence save the same amount s^1, s^2 or s^3. Then, because the three expressions (16) - (18) are each strictly increasing in damage probability, the larger risk has the steeper indifference curve. This is the standard regular-crossing result for the case where different risk types save the same amount and exhibit the same risk-aversion.

It is also possible, however, that two individuals that differ only in risk find different savings levels optimal in the same contract. To see this, examine the regime switching borders (11) - (13) in greater detail. For a fixed coverage $\beta < 1$, and under Assumption A1, the values $d_{ij}^{12}(\beta)$, $d_{ij}^{23}(\beta)$ and $d_{ij}^{13}(\beta)$ are all strictly decreasing in damage probability. This implies that a large risk individual will jump to a larger level of savings already for smaller premiums d, according to (14) and (15). There are thus contracts in which a high risk saves more than a low risk but not reversely (which is exactly what Proposition 1 demonstrates at a greater level of generality).

In fact, an increase in damage probability amounts to a counter-clockwise rotation of the regime switching borders around their value for $\beta = 1$, which is independent of risk. This is exemplified in Figure 1 for d_{iH}^{12} and d_{iL}^{12}. Then, if $C = (\beta, d)$ is such that $d_{iH}^{12}(\beta) < d < d_{iL}^{12}(\beta)$, we need to compare $\text{MRS}_{iL}^1(\beta, d)$ to $\text{MRS}_{iH}^2(\beta, d)$. Irregular-crossing ($\text{MRS}_{iL}^1 \geq \text{MRS}_{iH}^2$) occurs if

$$p_L \frac{\gamma_1}{\gamma_2} \geq p_H \frac{\alpha_3}{\alpha_2}. \tag{19}$$

Analogously, irregular-crossing can occur if type iL saves s^1 but type iH saves s^3 in the same contract. The corresponding condition becomes

$$\delta_i \left(p_L \frac{\gamma_1}{\alpha_3} + (1 - p_H) \frac{\gamma_3}{\alpha_1} \right) \geq 1. \tag{20}$$

Finally, we have $\text{MRS}_{iL}^2 \geq \text{MRS}_{iH}^3$ in some contract if

$$\delta_i \left(p_L \frac{\gamma_2}{\alpha_2} + (1 - p_H) \frac{\gamma_3}{\alpha_1} \right) \geq 1. \tag{21}$$

5.2 Appendix B: One-Dimensional Heterogeneity

In this appendix, we prove that, even with endogenous wealth accumulation, a model with one-dimensional heterogeneity in risk only cannot generate equilibria with a negative correlation between risk and coverage. One-dimensional heterogeneity in risk can be captured within the general framework of Section 2 either by setting $n_{iL} = n_{iH} = 0$ for one patience type i or by imposing $\delta_L = \delta_H$. In any case, we can omit the patience index i in all expressions and simply refer to individuals as low or high risks.

Proposition 3. *Consider any equilibrium $\mathscr{C}'$ under one-dimensional heterogeneity in risk only, and denote by $C_j = (\beta_j, d_j)$ the equilibrium contract chosen by risk type $j = L, H$. Then, $\beta_L \leq \beta_H$ holds.*

Proof. We prove the claim by showing that if $\beta_L \neq \beta_H$, $\beta_L < \beta_H$ must hold. We proceed in two steps.

Step 1. We first show that $V_H(C_H) = V_H(1, p_H D)$ and $d_H = p_H \beta_H D$. By way of contradiction, suppose first that $V_H(C_H) < V_H(1, p_H D)$. Then by continuity of V_H there exists a sufficiently small

$\epsilon > 0$ such that contract $C = (1, p_H D + \epsilon) \in \mathscr{C} \backslash \mathscr{C}'$ still satisfies $V_H(C_H) < V_H(C)$. But then $\pi_{\mathscr{C}' \cup \{C\}}(C) = x[(p_H - p_L)D + \epsilon] + n_H \epsilon > 0$, where $x = n_L$ if the low risks prefer C over C_L and $x = 0$ otherwise. This contradicts part *(ii)* of Definition 2.

Suppose next that $V_H(C_H) > V_H(1, p_H D)$. Since $\beta_H \neq \beta_L$, part (i) of Definition 2 implies that $\pi_{\mathscr{C}'}(C_H) = n_H[d_H - p_H \beta_H D] \geq 0$ must hold. Then, since $V_H(\beta, d)$ is decreasing in d, there exist an $\epsilon \geq 0$ such that contract $C = (\beta_H, d_H - \epsilon)$ satisfies $d_H - \epsilon = p_H \beta_H D$ and $V_H(C_H) \leq V_H(C)$, i.e. C is actuarially fair and makes high risks no worse off than C_H. Denote by $\bar{V}_H$ a high risk's utility from contract $(1, p_H D)$ when (non-optimal) savings $s_H^*(\beta_H, d_H - \epsilon)$ are chosen. We then have $V_H(C_H) \leq V_H(C) \leq \bar{V}_H \leq V_H(1, p_H D)$, where the second inequality holds because concavity of u_2 implies that an increase in coverage from β_H to 1 together with a fair adjustment of the premium (from $d_H - \epsilon$ to $p_H D$) weakly increases period 2 utility given fixed savings. The third inequality follows from switching to optimal savings in $(1, p_H D)$. But this contradicts the assumption $V_H(C_H) > V_H(1, p_H D)$.

Thus $V_H(C_H) = V_H(1, p_H D)$ must hold. In addition, $d_H = p_H \beta_H D$ must be true, since otherwise a contract $C = (\beta_H, d_H - \epsilon)$ for some small $\epsilon > 0$ would satisfy $V_H(C_H) < V_H(C)$ and $\pi_{\mathscr{C}' \cup \{C\}}(C) > 0$.

Step 2. We now show that $\beta_L \leq \beta_H$. To obtain a contradiction, assume that $\beta_H < \beta_L \leq 1$. We first claim that $V_H(C_H) = V_H(C)$ must then hold for any fair contract $C = (\beta, d)$ with $\beta_H < \beta < 1$ and $d = p_H \beta D$. Let $\bar{V}_H'$ denote a high risk's utility from such a contract C but with (non-optimal) savings $s_H^*(\beta_H, d_H)$, and $\bar{V}_H''$ the high risk's utility from $(1, p_H D)$ and savings $s_H^*(C)$. As above, concavity of u_2 then implies $V_H(C_H) \leq \bar{V}_H' \leq V_H(C) \leq \bar{V}_H'' \leq V_H(1, p_H D)$, which in turn implies $V_H(C_H) = V_H(C) = V_H(1, p_H D)$ and establishes the claim. Second, we claim that $C_L = (\beta_L, d_L)$ must actually be such a contract, i.e. satisfy $d_L = p_H \beta_L D$. Incentive compatibility $V_H(C_L) \leq V_H(C_H)$ and $\beta_H < \beta_L$ imply that $d_L \geq p_H \beta_L D$, because of our first claim and because V_H is decreasing in d. But $d_L > p_H \beta_L D$ is impossible due to the otherwise existence of a profitable deviation contract $(\beta_L, d_L - \epsilon)$, which establishes the second claim. Hence $V_H(C_H) = V_H(C_L)$, which contradicts $\beta_H < \beta_L$ because larger coverage is chosen in case of indifference. Therefore, we must have $\beta_L \leq \beta_H$, completing the proof. □

5.3 Appendix C: Proofs for the Equilibrium Construction

We prove the proposition in two steps. First, we provide a parameter collection $P \in \mathscr{P}$ (hence satisfying Assumption 1) for which indifference curves exhibit the shape depicted in Figure 2. This implies that $\mathscr{C}'$ as defined in the proposition indeed satisfies $C_2 < C_3$ for the given P (and all contracts in $\mathscr{C}'$ earn nonnegative profits, which is part (i) in Definition 2). Second, we show that there is no contract outside of $\mathscr{C}'$ that would earn strictly positive profits if offered in addition (part (ii) in Definition 2). The fact that there is an open subset $\mathscr{P}' \subset \mathscr{P}$ (around P) for which this argument is applicable then follows immediately.

Step 1. A collection $P \in \mathscr{P}$ for which indifference curves exhibit the shape as depicted in Figure 2 is given as follows:

Example Parameters $P \in \mathscr{P}$

Slopes		Kinks		Risk & Patience		Population	
α_1	1	$\bar{c}_1$	10	p_L	0.6	n_{HH}	0.1
α_2	3	$\bar{c}_2$	25	p_H	0.65	n_{LH}	0.6
α_3	15	$\tilde{c}_1$	35.5	D	11	n_{HL}	0.2
γ_1	20	$\tilde{c}_2$	50	W	67.6	n_{LL}	0.1
γ_2	3.5			δ_L	0.952		
γ_3	0.9						

Table 1.

Given the parameters from Table 1, the relevant indifference curves can be derived analytically or numerically. We made use of a computer program that first derives the indifference curve of type LH through contract C_1, then calculates the optimal incentive compatible contract C_3 for type LL, and finally contract C_3. The resulting slope- and crossing-patterns are as depicted in Figure 2.

Step 2. Figure 4 is a schematic depiction of the equilibrium candidate constructed above (and illustrated in Figure 2). Specifically, the four types' indifference curves through their equilibrium contracts span several regions, which are labelled with numbers from 1 to 10 in Figure 4. To guarantee that $\mathscr{C}'$ is an equilibrium, it needs to hold that no contract in any of these regions attracts a profitable subset of the population away from $\mathscr{C}'$. Contracts in regions not labelled with a number can never be profitable. Either they do not attract any type, or they earn losses irrespective of which individuals purchase them. In Figure 4, these latter contracts lie below the low risks' zero profit line (dotted, lower right corner).

Any contract in region 1, 4 or 8 is preferred only by high risks over their optimal contract from $\mathscr{C}'$ and cannot be a profitable deviation. In the graphical illustration in Figure 4, any such contract lies below the high-risk zero-profit line (dotted, upper left corner). Any contract in region 9 attracts the individuals who preferred C_2 in $\mathscr{C}'$ and must also earn losses. Graphically, region 9 is below the zero-profit line for the HL-HH-pool (dotted, in the center). The following table then summarizes for the remaining regions the set of types that can be attracted away from $\mathscr{C}'$ by contracts within that region.

Deviation Contracts in $\mathscr{C} \backslash \mathscr{C}'$

Region	2	3	5	6	7	10
Types	LH, LL	LH, LL, HH	LH, LL, HL	LH, HL	LH, HH, LL, HL	LH, HL, HH

Table 2.

Whether or not there is a profitable deviation can now be checked. For example, no contract (β, d) in region 2 is allowed to satisfy $\pi(\beta, d, n_{LL}, n_{LH}) > 0$. This condition can be verified by deriving the zero-profit line for the pool consisting of the LL and the LH types, i.e the function

$$d^0_{LL,LH}(\beta) = \beta \left(\frac{n_{LL} p_L + n_{LH} p_H}{n_{LL} + n_{LH}} \right) D.$$

Indeed it holds that $d < d^0_{LL,LH}(\beta)$ for every contract (β, d) in region 2, which we again verified computer-based. The conditions for the remaining regions are analogous, and all of them are satisfied.

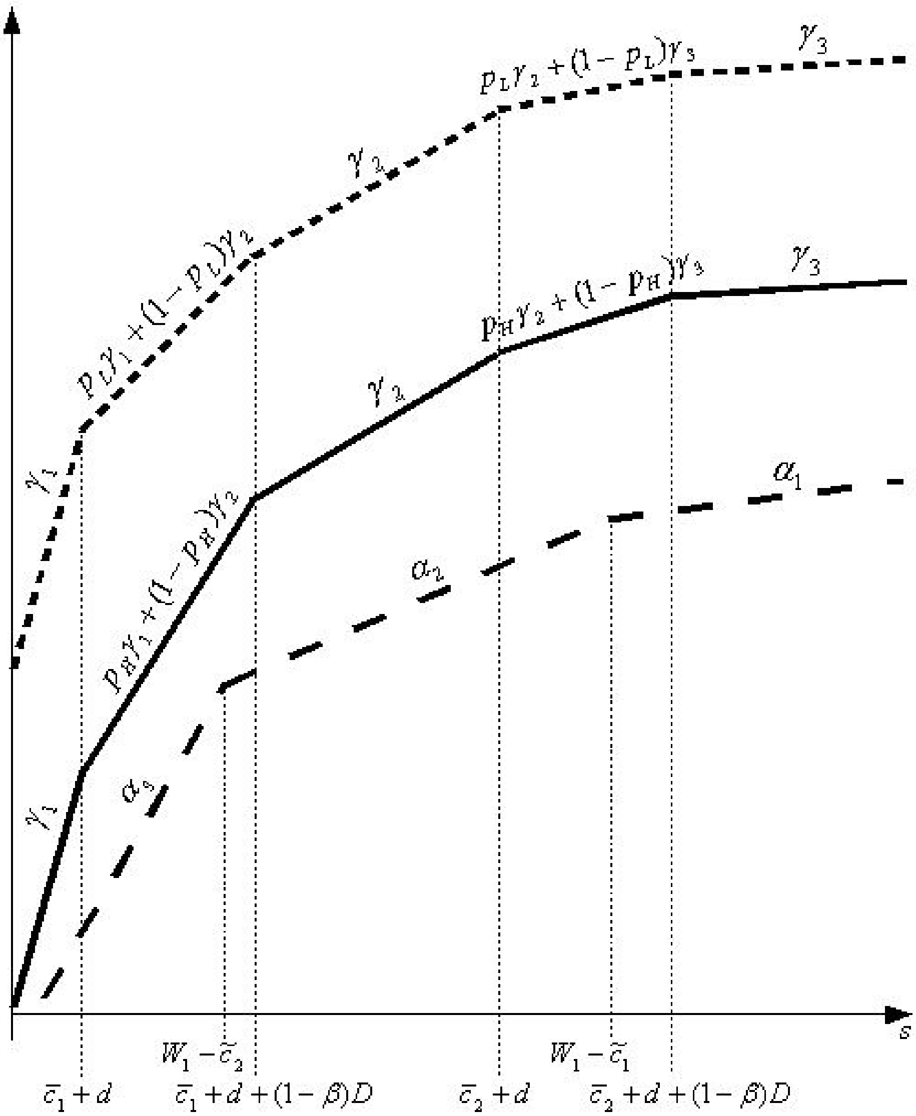

Figure 3: Optimal Savings of Patient Individuals in Contract (β, d).

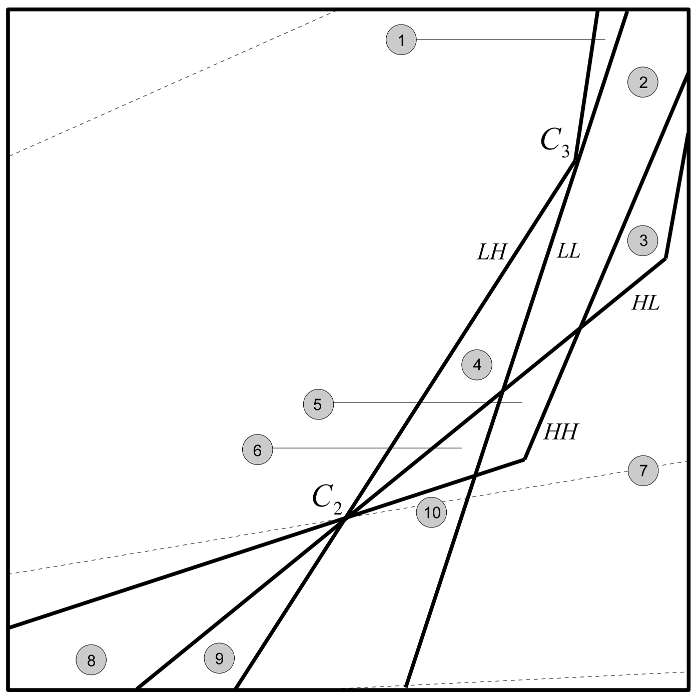

Figure 4: Existence Conditions.