Incomplete information in the samaritan's dilemma: the dilemma (almost) vanishes

Lagerlöf, Johan

discussion papers

FS IV 99 - 12

Incompleteate Information in the Samaritan's Dilemma: The Dilemma (Almost) Vanishes

Johan Lagerlöf

September 1999

ISSN Nr. 0722 - 6748

Forschungsschwerpunkt Marktprozeß und Unternehmensentwicklung

Research Area Market Processes and Corporate Development

Zitierweise/Citation:

Johan Lagerlöf, **Incomplete Information in the Samaritan's Dilemma: The Dilemma (Almost) Vanishes,** Discussion Paper FS IV 99 - 12, Wissenschaftszentrum Berlin, 1999.

Wissenschaftszentrum Berlin für Sozialforschung gGmbH,
Reichpietschufer 50, 10785 Berlin, Tel. (030) 2 54 91 - 0

ABSTRACT

Incomplete Information in the Samaritan's Dilemma: The Dilemma (Almost) Vanishes

by Johan Lagerlöf *

Suppose an altruistic person, *A*, is willing to transfer resources to a second person, *B*, if *B* comes upon hard times. If *B* anticipates that *A* will act in this manner, *B* will save too little from both agents' point of view. This is the Samaritan's dilemma. The mechanism in the dilemma has been employed in an extensive literature, addressing a wide range of both normative and positive issues. However, this paper shows that the undersaving result is not robust to the assumption that information is complete: by adding a slight amount of uncertainty one can sustain an equilibrium outcome that is arbitrarily close to ex post incentive efficiency. One may also sustain outcomes with oversaving.

ZUSAMMENFASSUNG

Unvollständige Information im Fall des Samariter-Dilemmas: Das Dilemma verschwindet (meist)

Angenommen eine altruistische Person *A* ist bereit, Ressourcen einer zweiten Person *B* zu übertragen, wenn für *B* schwierige Zeiten anbrechen. Wenn *B* antizipiert, daß *A* in dieser Weise handeln wird, dann wird *B* wenig sparen und zwar aus der Sicht beider Agenten. Dieses Phänomen wird als Samariter-Dilemma bezeichnet. Die Mechanismen, die dabei wirken, wurden in der Literatur sowohl auf einer normativen als auch empirischen Ebene ausgiebig diskutiert. In diesem Beitrag wird jedoch gezeigt, daß das Ergebnis des „Untersparens“ im Hinblick auf die Annahme unvollständiger Information nicht robust ist. Bei Berücksichtigung eines geringen Maßes an Unsicherheit stellt sich ein Gleichgewichtsergebnis ein, das nahe der Ex-post-Anreizeffizienz liegt. Man kann aber auch Ergebnisse mit dem „Übersparen“ erhalten.

---

* I thank J. Björnerstedt, D. Coen-Pirani, Y. Kang, N.-P. Lagerlöf, B. Persson, D. Soskice, J. Stennek, K. Wärneryd, J. Weibull, and seminar participants at the Stockholm School of Economics, the WZB, ESEM -98, EEA -98, and the Universities of Rochester (the Wallis Institute), Copenhagen and Alicante, as well as Free University of Berlin for helpful comments. Part of this research was conducted while I was visiting the Wallis Institute; I am grateful to them for their hospitality. Financial support from Jan Wallander and Tom Hedelius' Research Foundation *Svenska Institutet. Finanspolitiska Forskningsinstitutet*, and the European Commission (under the TMR Programme; contract No. ERBFMRXCT980203) is gratefully acknowledged.

1. Introduction

Suppose an altruistic person, hereafter called A, is willing to transfer resources to a second person, B, if B comes upon hard times. Then, if B today is to decide how much to save for tomorrow, and if B is well aware of A's altruistic concern for him, B will typically save too little as compared to what is socially optimal. This is what Buchanan (1975) has called the "Samaritan's dilemma." The dilemma arises because A is unable to commit not to help B out. Moreover, A's willingness to bail B out if he undersaves serves as an implicit tax on B's savings. For if B saves an extra dollar, then A will transfer, say, ten cents less to B than otherwise. This implicit tax distorts B's saving incentives. Hence, given the equilibrium level of A's support, B would be better off if he consumed less today and more tomorrow. And, since A has altruistic concerns for the welfare of B, this would make also A better off.[1]

The Samaritan's dilemma effect has been employed in a large number of papers, addressing a wide range of both normative and positive issues. For instance, the inefficiency result has been used to justify and/or explain the existence of compulsory social insurance systems (Thompson, 1980; Veall, 1986; Kotlikoff, 1987; Lindbeck and Weibull, 1988; Hansson and Stuart, 1989). The argument is that a government can force people to save and insure more than they would do voluntarily, thereby making free riding and the Samaritan's-dilemma-type inefficiency impossible. As another example, Bruce and Waldman (1991) and Coate (1995) argue that the Samaritan's dilemma provides an efficiency rationale for in-kind govermental transfers.[2] In those models the government provides a transfer on two occasions over time. Since a cash transfer would be used in an inefficient manner, all parties can benefit if the government gives the first transfer in a tied fashion, such as in the form of an illiquid investment.

Yet another example is from the macroeconomics literature. O'Connell and Zeldes (1993) study an infinite-horizon OLG-model with altruism from children

[1] Formal analyses of the Samaritan's dilemma are found in, e.g., Bernheim and Stark (1988) and Lindbeck and Weibull (1988). In the latter paper it is shown that the dilemma may also arise in the case where both agents are altruistic toward each other.

[2] In the context of intra-family transfers, Becker and Murphy (1988, p. 7) also make this argument. They do not, however, provide a formal model.

toward their parents. If, as is assumed in the standard literature, parental saving is non-strategic, this kind of model is characterized by dynamic inefficiency, i.e., the growth rate of the population exceeds the (endogenous) real interest rate. However, O'Connell and Zeldes demonstrate that the strategic undersaving effect will make the economy dynamically efficient. The reason for this is that less saving leads to a smaller capital stock, which in turn implies a larger marginal product of capital and thus a larger interest rate.

The Samaritan's dilemma also has bearing on the so-called rotten-kid theorem (Becker, 1974). This result concerns a situation where a selfish child can take an action that affects the income of the whole family. The theorem states that if the child's parent is sufficiently altruistic toward the child to transfer resources to it, then the child will choose an action that maximizes the income of the whole family. Hence, the presence of parental altruism induces the child to internalize the externality, and the resource allocation in the family is efficient. One of the conditions needed for this result to hold is that the transfer from parent to child indeed is positive.[3] However, Bruce and Waldman (1990) consider a two-period setting where the child in the first period takes an action affecting the income of the whole family *and* makes a saving decision. They show that if the parent makes an operative transfer (i.e., if the constraint that the transfer is non-negative is not binding) in the second period, then the child indeed chooses the action that maximizes family income; but, because of the Samaritan's dilemma, in this case the child also saves an amount that is too low relative to the efficient level. As a consequence, "rotten kids actually act rotten in at least one dimension, with the result being that the family unit does not achieve the Pareto frontier" (Bruce and Waldman, 1990, p. 157).[4]

Most of the existing literature on the Samaritan's dilemma assumes a setting with complete information: the agents know with certainty their own payoffs and the other agents' payoffs. This is typically an unrealistic assumption. Indeed, in many of the real world situations that are meant to be captured by the models in this literature, there is reason to believe that there is a substantial degree

[3] There are also other conditions, which are left implicit by Becker. For example, Bergstrom (1989) shows that utility must be transferable for the result to hold.

[4] Other papers that discuss and/or apply the Samaritan's dilemma include Kaufman (1995) and Lord and Rangazas (1995).

of incomplete information. It is at any rate hard to believe that there is no incomplete information at all. At the very least, one would therefore like the inefficiency result of the models in the existing literature to be *robust* to incomplete information. That is, it should not be possible to find a model with incomplete information that is close to the original model but which yields a very different prediction concerning the presence or the amount of inefficiences.

However, there is a strong a priori reason to believe that the undersaving result in the standard Samaritan's-dilemma model might *not* be robust to incomplete information. To see this, consider a situation like the one described in the introductory paragraph. Suppose, however, that B (i.e., the recipient of the transfer) has private information about some characteristic of himself which is relevant for his payoff. Moreover, suppose that the characteristic can be represented by a parameter x and that x is such that the larger its magnitude, the lower is B's marginal utility of consumption tomorrow. For example, x could be a measure of an exogenous income that B will receive tomorrow.

Since A cares about the welfare of B, A would be willing to make a larger transfer to B if A believed that x were small. The assumptions about x also imply that the smaller is x, the more B wants to save (everything else being equal). B thus has an incentive to make A believe that x is small, and B may try to do so by using his savings as a signaling device (à la Spence, 1973, 1974). In particular, B has an incentive to save *more* than in the standard setting with complete information. One should thus expect this mechanism to counteract the incentives to undersave in the standard Samaritan's dilemma model. Moreover, a standard result from the signaling literature is that even a very small amount of uncertainty can to a substantial degree distort the behavior in an incomplete information model as compared to the model with complete information. Hence it is quite conceivable that by adding a slight amount of uncertainty to a standard Samaritan's dilemma model, one can sustain outcomes that do not involve undersaving and which are not inefficient.

The purpose of this paper is to investigate whether the above intuition holds true in a formal analysis and, if so, to see how far the counteracting mechanism can take us. Section 2 of the paper starts out by presenting an example which in the simplest possible way shows how the mechanism works. In this example,

B can only make the binary choice whether to "save" or to "squander"; A then chooses one of three different transfer levels. In Sections 3 and 4 a somewhat less stylized but still simple model is considered where A and B can choose among a continuum of transfer and saving levels, respectively. Section 3 formulates the analytical benchmark: a model of the Samaritan's dilemma characterized by complete information. It turns out that in this model the inefficiency result is obtained if the degree of altruism is neither too low nor too high.

Subsection 4.1 extends the model in Section 3 by assuming that B has private information about a parameter in his payoff function. This parameter can take one of two distinct values; that is, B can be one of two *types*. In Subsection 4.2 the set of separating and pooling equilibrium outcomes of this incomplete information model is characterized. The analysis is restricted to values of the degree of altruism such that, for both types of B, the undersaving result would obtain if B's type were common knowledge. Subsection 4.3 then considers the question whether these equilibrium outcomes are efficient. The focus is on the special case where A's uncertainty about B's type is small. This means that the incomplete information model is close to the benchmark model of Section 3.

It turns out that one cannot sustain equilibrium outcomes that are on the Pareto frontier. However, in a subset of the parameter space one can sustain equilibrium outcomes that are arbitrarily close to the Pareto frontier. A sufficient condition for this is that the difference between the two types is large enough. To some extent this result relies on the fact that there are no restrictions on A's beliefs off the equilibrium path. In Subsection 4.4 an equilibrium refinement that eliminates all outcomes of pooling equilibria and all outcomes of separating equilibria except for one is imposed. Then, similarly to the preceding subsection, the question whether this unique equilibrium outcome is arbitrarily close to the Pareto frontier is considered. It turns out that for a subset of the parameter space this is indeed the case, although this subset is smaller than the one where the (almost) efficiency result could be sustained if not imposing the refinement. Interestingly, it can also happen that the unique equilibrium outcome involves *oversaving*. Moreover, regardless of whether the outcome that survives the equilibrium refinement involves undersaving or oversaving or if this outcome is (almost) efficient, it is qualitatively different from the one of the benchmark model with complete information. In particular, the saving level is higher than in the benchmark model.

Section 5 concludes with a discussion of the results. Most of the proofs are found in the Appendix.

2. An Example

Consider the following simple game played between two individuals, A and B (see also Figure 1). B can be of two types: either he is "poor" or "rich." While B knows his type from the outset of the game, A does *not* know B's type. A and B make one decision each. First B chooses whether to "save" or to "squander." Second, A observes whether B has saved or squandered and then chooses whether to give B a "big support" (abbreviated "bs" in the figure), a "small support" (ss), or "no support" (ns). The payoffs are such that if B has chosen to save, A wants to give B no support regardless of whether B is poor or rich. If B has squandered, then A wants to give B a big support if he is poor and a small support if he is rich. B, on the other hand, prefers a big support to a small support and a small support to no support, regardless of his type and regardless of whether he has saved or squandered.

Before analyzing the game with incomplete information depicted in Figure 1, let us make the observation that if A knew with certainty whether B is poor or rich (and if A's knowing this were common knowledge), then we would have an example of the Samaritan's dilemma: B would squander and then get a support from A, an outcome which is not (Pareto) efficient. To see this, start with the case where it is common knowledge that B is poor. It is straightforward to verify that then there is a unique (subgame perfect) equilibrium outcome, namely the one where B squanders and A gives him a big support. This outcome is dominated, however, by the outcomes where B saves and then gets either a small or a big support. Similarly with the case where it is common knowledge that B is rich. Then there is again a unique equilibrium outcome, namely the one where B squanders and gets a small support. This outcome is also inefficient, since it is dominated by the outcomes where B saves and then gets either a small or a big support.

Let us now consider the equilibria of the game where A does *not* know whether

B is poor or rich. As indicated in Figure 1, A puts the prior probability μ on B's being poor and the complementary probability $(1 - \mu)$ on B's being rich. We will be particularly interested in the case where A is almost certain that B is poor, so let us assume that $\mu \in (.5, 1)$. The solution concept that I employ is that of perfect Bayesian equilibrium.[5] First notice that in any such equilibrium, if B has saved, then A will choose "no support." This is because doing so is A's best action regardless of B's type. Similarly, if B has squandered, A will never choose "no support" in an equilibrium. This means that, if being rich, B will squander in any equilibrium. Having established this, let us first look for an equilibrium where the poor type saves and the rich type squanders; this is the only possible separating equilibrium.[6] In this kind of equilibrium, A will be able to infer B's true type perfectly. Thus, A's best response is to give B a small support if B has squandered and no support if B has saved. Given this behavior of A, it is indeed optimal for B to save if being poor and squander if being rich. Hence, this is an equilibrium.

There are a few things we should notice about this equilibrium. First, although the outcome for the rich type is the same as in the equilibrium of the corresponding complete information game, this is not true for the poor type. For the poor type we can in the incomplete information game sustain the outcome where B saves and A does not give B any support. This outcome is efficient among the outcomes that are relevant for the poor type. The reason why the poor type chooses to save is that if he squandered, then he would be perceived as the rich type and get only a small support; and this outcome would give him a lower payoff than he gets in the equilibrium. In other words, the possibility that B is rich exerts an externality on the poor type. This externality is bad for B, since the poor type now gets only the payoff 2 instead of the payoff 3 as in the complete information model. However, the externality is good for A, since she gets the payoff 4 instead of 2 when B is poor.

[5] This solution concept requires that, after having observed that B has either saved or squandered, A forms some beliefs about B's type. These beliefs must be consistent with Bayes' rule and B's equilibrium strategy, whenever Bayes' rule is defined. Whether B has saved or squandered, A makes a decision that, given her beliefs, maximizes her payoff. B is also required to make a decision that maximizes his payoff given A's behavior.

[6] A separating equilibrium is an equilibrium where the types behave differently.

A second thing that we should notice about the separating equilibrium is that it can be sustained for any small prior probability that B is rich, as long as the probability is strictly positive. This might seem counter-intuitive but it is a standard feature of signaling models. Hence, when the probability that B is poor, μ, is arbitrarily close to unity, the equilibrium outcome is, from an ex ante point of view, arbitrarily close to efficiency. Moreover, by letting the probability that B is rich be arbitrarily small, we also get arbitrarily close to the complete information model. As we concluded above, that model has a unique equilibrium where both types squander and which is inefficient. Thus, if the players coordinate on the separating equilibrium, B's saving behavior will make a discrete jump as we go from the environment with complete information to an environment where there is a small probability that B is rich: from having saved too little, B starts saving an amount that is very close to the efficient amount.

There also exists a pooling equilibrium outcome where both types of B squander, and A gives B a big support if B has squandered and no support in the out-of-equilibrium event that B has saved. Finally there is a semi-pooling equilibrium where the poor type squanders with probability $(1-\mu)/\mu$ and saves otherwise, and the high type always squanders. If B has squandered, A gives B a big or a small support with equal probability; if B has saved, then A gives no support. The semi-pooling equilibrium outcome has the same property as the separating one: as μ tends to unity, we get arbirtrarily close to the complete information model and arbitrarily close to efficiency. However, the pooling equilibrium outcome still gives rise to the Samaritan's dilemma for both types and for all values of μ.

It is not clear which one of the three equilibrium outcomes is the most reasonable prediction of the game. But if we believe in the assumption that A is a little bit uncertain about B's type and if we are not certain that the players will behave according to the pooling equilibrium outcome, then it is tempting to conclude that the theoretical case for undersaving is weaker than what one might think if one only looked at the standard formulation of the Samaritan's dilemma. However, one might object that the example we have analyzed is rather special. For instance, it is not clear how restrictive the specification of the players' payoffs is. Moreover, standard equlbrium refinements that put restrictions on the players' beliefs about out-of-equilbrium events do not have any bite in this simple example, wheras one should expect this to be the case in many alternative settings. I will, therefore, in the following two sections further explore the signaling effect, but in a model that is somewhat less stylized than the above example.

3. The Benchmark: Complete Information

The model in this section is characterized by symmetric and complete information, and in many respects it is similar to existing models in the literature. Accordingly the results and insights that shall be derived from the model are not novel. However, the model will serve as a useful a benchmark when, later in Section 4, a model with incomplete information is considered.

3.1. The Model

There are two individuals, A and B, and two time periods, 1 and 2. A lives only in period 2 while B lives in both periods. At the beginning of the first period, B is endowed with exogenous income $\omega > 0$. B's decision concerns how much of this income to save for period 2, $s \in [0, \omega]$. The residual amount, $c_{1B} = \omega - s$, constitutes B's first-period consumption. A's endowment also equals ω. In the second period, after having observed s, A chooses how much of her endowment to transfer to B, $t \in [0, \omega]$.[7] A consumes the residual amount, $c_A = \omega - t$, herself. B's second-period consumption consists of his savings plus the transfer from A: $c_{2B} = s + t$.

B has preferences over his own consumption in period 1 and 2, described by the following utility function:

$$\begin{aligned} U_B(s,t) &= \log(c_{1B}) + \beta \log(c_{2B}) \\ &= \log(\omega - s) + \beta \log(s + t), \end{aligned} \tag{3.1}$$

where $\beta \in (0, 1)$ is a fixed parameter.[8] A is altruistic in the sense that she

[7] Hence a non-negativity constraint is imposed on the transfer, $t \geq 0$: A cannot *take* income from B. This assumption seems natural and it is common in the literature. In Section 4 we will see that the non-negativity constraint is crucial for some of the results concerning efficiency in the model with incomplete information.

[8] The assumption that β is smaller than unity will be made also in the model with incomplete information considered in Section 4. The assumption is not important for my argument but it simplifies the analysis.

has preferences over both her own consumption and B's utility level U_B. These preferences are described by the following utility function:

$$\begin{aligned} U_A(s,t) &= \log(c_A) + \alpha U_B(s,t) \\ &= \log(\omega - t) + \alpha \log(\omega - s) + \alpha\beta \log(s+t). \end{aligned} \tag{3.2}$$

Here $\alpha > 0$ is a fixed parameter that represents the altruistic concern of A for the welfare of B. The structure of the model and in particular the individuals' preferences are common knowledge.

3.2. Analysis

The model described in the preceding subsection constitutes an extensive form game. Denote this game by Γ_C (where C stands for complete information). I will solve for the subgame perfect equilibria of Γ_C through backward induction.

Let us thus begin by considering A's problem in period 2. A then maximizes $U_A(s,t)$ as given in equation (3.2) with respect to t subject to the constraint $t \in [0,\omega]$. Denote the solution to this problem by $\widehat{t}$. It is easy to verify that

$$\widehat{t} = \begin{cases} \frac{\alpha\beta\omega - s}{1+\alpha\beta} & \text{for } s \leq \alpha\beta\omega \\ 0 & \text{otherwise.} \end{cases} \tag{3.3}$$

Note that for any $s < \alpha\beta\omega$, $\widehat{t}$ is decreasing in s. That is, if B increases his savings, A will make a smaller transfer to him. One may think of this effect as an implicit tax on savings. In the analysis that follows we shall see that the implicit tax distorts B's saving incentives and typically makes him consume too much in period 1, as compared to what is socially optimal.

Now consider period 1. Anticipating $\widehat{t}$, B chooses s. B's indirect utility is given by

$$U_B\left(s,\widehat{t}\right) = \log(\omega - s) + \begin{cases} \beta \log(\omega + s) + \beta \log\left(\frac{\alpha\beta}{1+\alpha\beta}\right) & \text{for } s \leq \alpha\beta\omega \\ \beta \log(s) & \text{otherwise.} \end{cases} \tag{3.4}$$

There are two cases to investigate: (i) $\alpha\beta \geq 1$ and (ii) $\alpha\beta < 1$. The analysis of case (ii) is rather cumbersome and is therefore deferred to the Appendix. Case (i)—the one where A's concern for B is relatively great—is very straightforward. Since in this case the non-negativity constraint on t is not binding for any $s \in [0, \omega]$, B solves

$$\max_{s \in [0,\omega]} \log(\omega - s) + \beta \log(\omega + s) + \beta \log\left(\frac{\alpha\beta}{1 + \alpha\beta}\right). \tag{3.5}$$

By assumption $\beta < 1$; hence this problem has the solution $s^* = 0$. That is, B saves nothing and relies fully on the anticipated transfer from A. Substituting $s^* = 0$ into equation (3.3) yields the equilibrium outcome of t, $t^* = \alpha\beta\omega / (1 + \alpha\beta)$.

Proposition 1 summarizes the results of the analysis above as well as the results for the case $\alpha\beta < 1$ derived in the Appendix. Before considering the proposition, however, we must introduce some more notation. Let the function $\varphi : (0, 1) \to \Re$ be defined by

$$\varphi(\beta) = \left[(1 + \beta)^{\frac{1+\beta}{\beta}} - \beta\right]^{-1}. \tag{3.6}$$

The expression in (3.6) is a critical value of α that will be used when describing what the equilibrium outcomes are in different subsets of the parameter space. In the Appendix (Lemma A1), it is proven that the function φ is decreasing, $\varphi' < 0$. One may also show that $\lim_{\beta \to 0} \varphi(\beta) = 1/e$ (where $e^{-1} \approx 0.37$) and it is easy to see that $\lim_{\beta \to 1} \varphi(\beta) = 1/3$. Thus, for any $\beta \in (0, 1)$, $\varphi(\beta)$ approximately equals one third.

Proposition 1. *For any $\alpha \neq \varphi(\beta)$ there exists a unique subgame perfect equilibrium of Γ_C, and the outcome of this equilibrium is*

$$(s^*, t^*) = \begin{cases} \left(\frac{\beta}{1+\beta}\omega, 0\right) & \text{for } \alpha < \varphi(\beta) \\ \left(0, \frac{\alpha\beta}{1+\alpha\beta}\omega\right) & \text{for } \alpha > \varphi(\beta). \end{cases} \tag{3.7}$$

For $\alpha = \varphi(\beta)$ there exists a continuum of subgame perfect equilibria of Γ_C. However, the outcome of any such equilibrium is either $(s^*, t^*) = (\beta\omega / (1 + \beta), 0)$ or $(s^*, t^*) = (0, \alpha\beta\omega / (1 + \alpha\beta))$.

Figure 2 illustrates the results stated in the proposition. The critical value of α defined in equation (3.6), φ, is depicted in the diagram as a function of β. For values of α below this critical value, B saves the fraction $\beta / (1 + \beta)$ of his income and A does not make a transfer. However, for values of α above the critical value, the behavior of A and B is quite different: B saves nothing and A transfers the fraction $\alpha\beta / (1 + \alpha\beta)$ of her income to B. For values of α exactly at the critical value, $\alpha = \varphi(\beta)$, B is indifferent between saving nothing and saving the fraction $\beta/(1+\beta)$ of his income, and any randomization between these two choices may be sustained as part of an equilibrium. For any outcome of such a randomization, A will make a transfer to B according to equation (3.3), i.e., either not make any transfer or transfer the fraction $\alpha\beta/(1+\alpha\beta)$ of her income.

3.3. Efficiency

In the model, resources can be allocated in two dimensions. First, given a value of t, one may reallocate resources intertemporally by varying s. Second, given a value of s, one may reallocate resources inter-individually by varying t. In this subsection I will consider the question whether the resource allocation induced by a subgame perfect equilibrium of Γ_C is Pareto efficient. First, however, consider some formal definitions.

An *allocation* (c_A, c_{1B}, c_{2B}) is a specification of a consumption level $c_A \in \Re_+$ for A and a consumption vector $(c_{1B}, c_{2B}) \in \Re_+^2$ for B. Since there is a one-to-one relationship between a consumption vector (c_A, c_{1B}, c_{2B}) and a saving-transfer vector (s,t), one may also refer to the vector (s,t) as an allocation, and I shall do so throughout this section. An allocation (s,t) is said to be *feasible* if it belongs to the set $[0,\omega]^2$. An allocation (s',t') is said to *dominate* an allocation (s,t) if

$$U_i\left(s',t'\right) \geq U_i(s,t), \qquad \forall i \in \{A,B\} \tag{3.8}$$

with at least one strict inequality. An allocation (s,t) is *Pareto efficient* if there exists no other feasible allocation (s',t') that dominates (s,t).

The following result is proven in the Appendix.

Proposition 2. *Suppose that $\alpha \neq \varphi(\beta)$. Then the allocation induced by the unique subgame perfect equilibrium of Γ_C is Pareto efficient if and only if either $\alpha < \varphi(\beta)$ or $\alpha \geq (1-\beta)^{-1}$. The allocation (s^*,t^*) induced by a subgame perfect equilibrium of Γ_C when $\alpha = \varphi(\beta)$ is Pareto efficient if and only if $(s^*,t^*) = (\beta\omega/(1+\beta), 0)$.*

That is, if A's degree of altruism takes on any value $\alpha \in \left(\varphi(\beta), (1-\beta)^{-1}\right)$, then the equilibrium outcome is not Pareto efficient (cf. Figure 2). Recall that for these values of α, B saves nothing but receives a transfer $t^* = \alpha\beta\omega/(1+\alpha\beta)$ from A. However, if B made a ceteris paribus increase in his savings, he would be better off. Moreover, since A has altruistic concerns for the welfare of B, this would make also A better off. The reason why B saves too little is the implicit tax on his savings: if B saved more, A would have an incentive to make the transfer smaller. Hence, crucial for the inefficiency result is that A can observe how much B has saved and that she cannot precommit to any transfer level.

If the degree of altruism is either sufficiently low ($\alpha < \varphi(\beta)$) or sufficiently high ($\alpha \geq (1-\beta)^{-1}$), then the equilibrium outcome is Pareto efficient. The intuition for this is straightforward. For $\alpha < \varphi(\beta)$, B does not receive any transfer and must rely only on his own savings. Since A's degree of altruism is so small it is, given the level of B's savings, indeed optimal for A not to transfer any income to B. B anticipates this so his saving choice is not distorted. Similarly for the case $\alpha > (1-\beta)^{-1}$. Here, since A cares so much about the welfare of B, it is in *both* A's and B's interest that B does not save anything himself.

4. Incomplete Information and Signaling

In this section I will consider an adaptation of the model described and analyzed in the previous section. Subsection 4.1 describes this model and introduces some new notation. Subsection 4.2 characterizes the equilibrium outcomes of the model. In Subsection 4.3 the question is posed whether the allocations induced by the equilibrium outcomes are efficient. Finally, Subsection 4.4 asks the same question after having imposed an equilibrium refinement that rules out all equilibrium outcomes except for one.

4.1. The Model

Relative to the model considered in the previous section, there is only one change: B is now assumed to have private information about the exact magnitude of the parameter β, and he learns about this in the beginning of the game.[9] The

[9] Hence, the private information does not concern B's second-period income, as was suggested in the Introduction (indeed, in the models described in this section and in Section 3, B does not have any second-period income). This assumption is made for the sake of tractability: it simplifies the analysis considerably if having uncertainty about a parameter multiplied with—

parameter β may be either "low" or "high": $\beta \in \{\beta_L, \beta_H\}$, where $0 < \beta_L < \beta_H < 1$. If $\beta = \beta_L$, then B will be referred to as the "low type"; and if $\beta = \beta_H$, then B will be referred to as the "high type." A places the prior probability $\mu \in (0,1)$ on B's being the high type and the prior probability $(1-\mu)$ on B's being the low type. The magnitude of the parameter μ is common knowledge.

All other model features are identical to the model described in Subsection 3.1. In particular there are two time periods. B lives in both of them while A lives only in period 2. A and B are each endowed with exogenous income $\omega > 0$. In period 1, B first learns his type and then chooses how much of his income to save for period 2, $s_i \in [0,\omega]$ where $i = L, H$; s_L is the amount of savings chosen by the low type and s_H is the amount of savings chosen by the high type. A does not know B's type but observes his actual savings, denoted by s. In period 2, A chooses how much of her income to transfer to B, $t(s) \in [0,\omega]$. Denote A's posterior beliefs that B is the high type, on having observed s, by $\widetilde{\mu}(s)$. Also, denote A's and B's utility functions, given B's type, by $U_A(s,t \mid \beta_i)$ and $U_B(s,t \mid \beta_i)$:

$$U_A(s,t \mid \beta_i) = \log(\omega - t) + \alpha \log(\omega - s) + \alpha\beta_i \log(s+t), \qquad i = L, H, \quad (4.1)$$

$$U_B(s,t \mid \beta_i) = \log(\omega - s) + \beta_i \log(s+t), \qquad i = L, H. \quad (4.2)$$

Again, $\alpha > 0$ is a fixed parameter.

Let us denote this game by Γ_I (where I stands for incomplete information). The equilibrium concept that will be employed in the analysis of Γ_I is that of perfect Bayesian equilibrium. A perfect Bayesian equilibrium in this game is a list of saving and transfer levels $(s_L^*, s_H^*, t^*(s))$ and beliefs $\widetilde{\mu}(s)$ such that (i) s_L^* and s_H^* maximize B's utility, given A's equilibrium transfer $t^*(s)$:

$$s_i^* \in \arg\max_{s_i \in [0,\omega]} U_B(s_i, t^*(s_i) \mid \beta_i), \qquad \forall i \in \{L, H\}; \quad (4.3)$$

(ii) for any s, $t^*(s)$ maximizes A's expected utility, given beliefs $\widetilde{\mu}(s)$:

$$t^*(s) \in \arg\max_{t(s)\in[0,\omega]} (1-\widetilde{\mu}(s)) U_A(s, t(s) \mid \beta_L) + \widetilde{\mu}(s) U_A(s, t(s) \mid \beta_H), \qquad \forall s \in [0,\omega]; \quad (4.4)$$

and not in the argument of—the utility function. However, one should expect the qualitative results of the analysis to be similar if one instead assumed that B has private information about an exogenous second-period income. Section 5 contains a discussion of what results one should expect if one assumed private information about other parameters than β.

and (iii) A's beliefs are consistent with B's observed behavior in the sense that Bayes' rule determines $\widetilde{\mu}(s)$ whenever the probability that B saves s in equilibrium is positive, which implies that

$$\widetilde{\mu}(s_H^*) = 1 \text{ and } \widetilde{\mu}(s_L^*) = 0 \text{ if } s_H^* \neq s_L^*, \tag{4.5}$$

$$\widetilde{\mu}(s_H^*) = \widetilde{\mu}(s_L^*) = \mu \text{ if } s_H^* = s_L^*. \tag{4.6}$$

For any $s \neq s_i^*$, $i = L, H$, perfect Bayesian equilibrium only requires that $\widetilde{\mu}(s) \in [0, 1]$. Henceforth I will write "equilibrium" when I mean perfect Bayesian equilibrium. As my notation indicates I will only consider pure-strategy equilibria. For notational convenience, let us write $t(s_i) = t_i$, and let us denote an outcome of a pure-strategy equilibrium by $(s_L^*, t_L^*, s_H^*, t_H^*)$.

4.2. Equilibrium Behavior

I shall restrict the analysis to the subset of the parameter space satisfying the following assumption.

Assumption 1.

$$\alpha \in \left(\varphi(\beta_L), (1 - \beta_L)^{-1}\right). \tag{4.7}$$

Imposing Assumption 1 means that we only consider the subset of the parameter space where, for both types, the equilibrium outcome is not Pareto efficient in the corresponding complete information model (cf. Proposition 2 and Figure 2). It is important to note that throughout the remainder of the paper, all the results that are reported presuppose that Assumption 1 holds, even where this is not explicitly stated.

4.2.1. Separating Equilibria

To start with, let us characterize the outcomes of separating equilibria. The following lemma (the proof of which is found in the Appendix) states that when $\alpha > \varphi(\beta_L)$ the low type chooses not to save.

Lemma 1. *Suppose that $\alpha > \varphi(\beta_L)$. Then, in any separating equilibrium, $s_L^* = 0$.*

The intuition for this result is straightforward. If possible, and regardless of his true type, B would like to be perceived as the high type in the eyes of A. However, in any separating equilbrium B's type will, by definition, be revealed. Hence, the best thing the low type can do in such an equilibrium is to behave optimally taking into account that A will know B's type when making her transfer decision. We know from the analysis of the benchmark model that, under Assumption 1, this optimal behavior is to save nothing.

We are thus looking for an equilibrium where $s_L^* = 0$ and where s_H^* is positive. The analysis will be facilitated by Figure 3a, which shows the saving-transfer space. The two straight lines in the figure represent A's optimal transfers, as given by equation (3.3); the lower straight line corresponds to A's believing that $\beta = \beta_L$, and the upper one corresponds to A's believing that $\beta = \beta_H$. At values of s larger than $\alpha\beta_L\omega$ respectively $\alpha\beta_H\omega$, the non-negativity constraint on A's transfer is binding, and the optimal transfer is zero. The figure also depicts two indifference curves through the point $(s,t) = (0, \alpha\beta_L\omega/(1+\alpha\beta_L))$, one for the low type and one for the high type. Hence, these are the levels of utility associated with the types' choosing the low type's equilibrium amount of savings and receiving the low type's equilibrium transfer.[10] Notice that each type is made better off if moving northward in the diagram, i.e., if getting a higher t for any given s. A move in the northwest direction is not necessarily making B better off. However, it turns out that if moving northwest along the upper straight line, both types are made better off.

As indicated in the figure, the value of s for which the low type's indifference curve intersects the upper straight line is denoted by s'; and the value of s for which the high type's indifference curve intersects the same line is denoted by s''. These points of intersection, s' and s'', are defined implicitly by the following two

[10] One may show that the two types' indifference curves through the point $(s,t) = \left(0, \frac{\alpha\beta_L\omega}{1+\alpha\beta_L}\right)$ must, as drawn in Figure 3a, be strictly concave functions of s, which tend to infinity as s tends to ω. Moreover, at $s = 0$, the slopes of these functions are negative but still larger than the slope of the lower straight line; and, as s approaches ω, the slopes approach infinity. Finally, for any given s, the low type's indifference curve has a larger slope than the high type's (the single-crossing property).

identities:[11]

$$\log\left(1-\frac{s^{'}}{\omega}\right)+\beta_L\log\left(1+\frac{s^{'}}{\omega}\right)\equiv\beta_L\log\left(\frac{\beta_L\left(1+\alpha\beta_H\right)}{\beta_H\left(1+\alpha\beta_L\right)}\right) \tag{4.8}$$

respectively

$$\log\left(1-\frac{s''}{\omega}\right)+\beta_H\log\left(1+\frac{s''}{\omega}\right)\equiv\beta_H\log\left(\frac{\beta_L\left(1+\alpha\beta_H\right)}{\beta_H\left(1+\alpha\beta_L\right)}\right). \tag{4.9}$$

As the figure is drawn, both $s^{'}$ and s'' are located to the left of $\alpha\beta_H\omega$. In the Appendix (Lemma A2 and A3, respectively) it is shown that we always have $s^{'}<\alpha\beta_H\omega$, and that $s^{''}<\alpha\beta_H\omega$ if and only if either (i) $\alpha\beta_H\geq 1$ or (ii) $\alpha\beta_H<1$ and $\beta_L>f\left(\beta_H,\alpha\right)$, where

$$f\left(\beta_H,\alpha\right)\equiv\left[\frac{1}{\beta_H}\left(1-\alpha\beta_H\right)^{-\frac{1}{\beta_H}}-\alpha\right]^{-1}. \tag{4.10}$$

To start with, suppose that either condition (i) or (ii) holds so that indeed $s^{''}<\alpha\beta_H\omega$, and refer again to Figure 3a. In order to sustain a separating equilibrium there are two necessary conditions. First, the low type must not have an incentive to choose the high type's amount of savings. Second, the high type must not have an incentive to choose the low type's amount of savings. Since A will learn B's type perfectly in any separating equilibrium, A will (after having observed s_L^* or s_H^*) make a transfer according to either one of the two straight lines in the figure. Hence, by mimicking the high type, the low type can get a transfer according to the upper straight line. For the low type not to have an incentive to do this we must have $s_H^*\geq s^{'}$; otherwise the low type could, by saving s_H^*, obtain a saving-transfer pair that gives him a higher utility than he will get if saving only $s_L^*=0$. Similarly, for the high type not to have an incentive to mimic the low type, we must have $s_H^*\leq s''$.

If A's out-of-equilibrium beliefs are chosen appropriately, the two necessary conditions $s_H^*\geq s^{'}$ and $s_H^*\leq s''$ are also sufficient for having $s_L^*=0$ and any

[11] The left-hand side of (4.8) is strictly decreasing in $s^{'}$. Also, the left-hand side of (4.8) tends to $-\infty$ as $s^{'}$ tends to ω; and as $s^{'}$ tends to 0 the left-hand side tends to zero, while the right-hand side of (4.8) is strictly negative. Hence the identity in (4.8) uniquely defines $s^{'}$. Similarly, the identity in (4.9) uniquely defines $s^{''}$.

$s_H^* \in \left[s^{'}, s^{''}\right]$ as part of a separating equilibrium. For instance, one may let

$$\widetilde{\mu}(s) = \begin{cases} 0 & \text{for } s \in [0, s_H^*) \\ 1 & \text{for } s \in [s_H^*, \omega]. \end{cases} \tag{4.11}$$

These posterior beliefs are consistent with the equilibrium requirements and they guarantee that the types do not have an incentive to deviate from s_L^* respectively s_H^*.

Essentially, what is needed for sustaining this kind of equilibrium is that the difference between β_L and β_H is sufficiently small. This can be seen from Figure 4. This figure depicts the (α, β_L)-space for some given β_H. Recall that we have restricted the analysis to the subset of the parameter space satisfying Assumption 1; in the figure, this is the area below the graph of $\alpha = (1 - \beta_L)^{-1}$ and above the graph of $\alpha = \varphi(\beta_L)$. The type of equilibrium outcome that we have just derived can be sustained in the region labeled (3.1); that is, this is the region where either condition (i) or (ii) above holds and where Assumption 1 is met.[12]

It remains to consider the case where $s^{''} \geq \alpha\beta_H\omega$ (or, equivalently, the subset of the parameter space where $\alpha\beta_H < 1$ and $\beta_L \leq f(\beta_H, \alpha)$). This case is illustrated in Figure 3b. Again, in order to sustain a separating equilibrium it must be that the two types do not have an incentive to mimic each other: s_H^* must be between $s^{'}$ and s°, where s° is the value of s at which the high type's indifference point crosses the s-axis from below. It turns out, however, that we cannot sustain all saving levels between $s^{'}$ and s° as part of a separating equilibrium. To see this, notice that regardless of A's beliefs, for saving levels larger than $\alpha\beta_H\omega$ A's transfer will be the same (zero) for both types. This means that if the high type saves some amount greater than $\alpha\beta_H\omega$, then this amount must be equal to $\beta_H\omega/(1+\beta_H)$: the high type's optimal saving level if not expecting any transfer.

We now get two subcases. The first one is obtained for $\beta_H\omega/(1+\beta_H) < \alpha\beta_H\omega$. If this inequality holds we cannot sustain any saving level in the interval $[\alpha\beta_H\omega, s^{\circ}]$ as part of a separating equilibrium. However, by choosing the out-of-equilibrium beliefs appropriately, one can sustain any $s_H^* \in \left[s^{'}, \alpha\beta_H\omega\right)$ as part of a separating equilibrium. For example, one may again let the beliefs be given by equation (4.11). This kind of equilibrium can be sustained in the region in Figure

[12]Some properties of the (inverse of the) function f that are helpful when drawing the figure are proven in the Appendix (Lemma A10).

4 labeled (3.2). That is, α must be large enough and the difference between β_L and β_H must also be sufficiently large.

The other subcase is obtained when $\beta_H\omega/(1+\beta_H) \geq \alpha\beta_H\omega$. If this inequality holds we can indeed sustain some saving level in the interval $[\alpha\beta_H\omega, s^\circ]$ as part of a separating equilibrium, namely $s_H^* = \beta_H\omega/(1+\beta_H)$.[13] We cannot, however, sustain any other saving level in this interval as part of a separating equilibrium, since the high type would then have an incentive to deviate to the saving level $\beta_H\omega/(1+\beta_H)$. Moreover, in order to sustain a saving level in the interval $\left[s', \alpha\beta_H\omega\right)$ as part of a separating equilibrium, the high type must not have an incentive to deviate to the saving level $\beta_H\omega/(1+\beta_H)$. For this condition to be met, the high type's saving level must not be too high; more precisely, it must be lower than or equal to $s^{\#}$, where $s^{\#}$ is defined implicitly by

$$\log\left(1-\frac{s^{\#}}{\omega}\right)+\beta_H\log\left(1+\frac{s^{\#}}{\omega}\right)\equiv\beta_H\log\left(\frac{1+\alpha\beta_H}{\alpha}\right)-(1+\beta_H)\log(1+\beta_H). \tag{4.12}$$

Hence, if $s^{\#} < s'$, then only $s_H^* = \beta_H\omega/(1+\beta_H)$ can be sustained as part of a separating equilibrium; otherwise, if $s^{\#} \geq s'$, we can in addition sustain any saving level in the interval $\left[s', s^{\#}\right]$ as part of a separating equilibrium. (We may again let the out-of-equilibrium beliefs be given by equation (4.11).) The first kind of equilibrium can be sustained if α is lower than a threshold $\psi(\beta_L, \beta_H)$ (the region labeled (3.4) in Figure 4);[14] the latter kind of equilibrium can be sustained if the difference between β_L and β_H is sufficiently large and if α takes on some intermediate value (the region labeled (3.3) in Figure 4). Notice that in the region labeled (3.4), there is a unique separating equilibrium outcome.

The results of the analysis carried out in the previous pharagraphs are summarized in Proposition 3.

Proposition 3. *(Separating equilibria) Separating equilibria exist.* $(s_L^*, t_L^*, s_H^*, t_H^*)$

[13] In the Appendix (Lemma A4) it is shown that $\beta_H\omega/(1+\beta_H)$ cannot be greater than (or equal to) s°.

[14] In the Appendix (Lemma A5), it is shown that the threshold $\psi(\beta_L, \beta_H)$ is well defined and that $\psi(\beta_L, \beta_H) > \varphi(\beta_H)$ for all $\beta_L \in (0, \beta_H)$ and that $\lim_{\beta_L\to 0}\psi(\beta_L, \beta_H) = \lim_{\beta_L\to\beta_H}\psi(\beta_L, \beta_H) = \varphi(\beta_H)$. Moreover, I strongly conjecture that $\psi(\beta_L, \beta_H) < (1+\beta_H)^{-1}$ for all $\beta_L \in (0, \beta_H)$, although I have not shown this analytically.

can be sustained as the outcome of a separating equilibrium if and only if

$$\left(s_L^*, t_L^*, s_H^*, t_H^*\right) \in \{0\} \times \left\{\frac{\alpha\beta_L\omega}{1+\alpha\beta_L}\right\} \times \Omega^{sep} \times \left\{\max\left\{\frac{\alpha\beta_H\omega - s_H^*}{1+\alpha\beta_H}, 0\right\}\right\},$$

where

- *(3.1)* $\Omega^{sep} = \left[s', s''\right]$ *if either (a)* $\alpha\beta_H \geq 1$ *or (b) if* $\alpha\beta_H < 1$ *and* $\beta_L > f\left(\beta_H, \alpha\right)$*;*
- *(3.2)* $\Omega^{sep} = \left[s', \alpha\beta_H\omega\right)$ *if* $\alpha\beta_H < 1$, $\alpha > \left(1+\beta_H\right)^{-1}$*, and* $\beta_L \leq f\left(\beta_H, \alpha\right)$*;*
- *(3.3)* $\Omega^{sep} = \left[s', s^{\#}\right] \cup \left\{\frac{\beta_H\omega}{1+\beta_H}\right\}$ *if* $\alpha \in \left[\psi\left(\beta_L, \beta_H\right), \left(1+\beta_H\right)^{-1}\right]$ *and* $\beta_L \leq f\left(\beta_H, \alpha\right)$*;*
- *(3.4)* $\Omega^{sep} = \left\{\frac{\beta_H\omega}{1+\beta_H}\right\}$ *if* $\alpha < \psi\left(\beta_L, \beta_H\right), \alpha \leq \left(1+\beta_H\right)^{-1}$*, and* $\beta_L \leq f\left(\beta_H, \alpha\right)$*.*

4.2.2. Pooling Equilibria

Let us now consider the existence of pooling equilibria—that is, equilibria where both types make the same saving decision, say $s_L^* = s_H^* = s^*$. After having observed this saving level, A will make a transfer to B according to equation (3.3) where the parameter β has been substituted for the expected value of β, $E\left(\beta\right) \equiv \mu\beta_H + \left(1-\mu\right)\beta_L$. The graph of this optimal transfer function is depicted as the intermediate straight line in Figure 5. Similarly to the case with separating equilibria, there are two conditions that are necessary in order to sustain s^* as part of a pooling equilibrium. First, if $s^* > 0$, then the low type must not have an incentive to deviate to $s = 0$. Second, the high type must not have an incentive to deviate to the saving level $s = \frac{\beta_H\omega}{1+\beta_H}$. This is the optimal saving level for the high type if not expecting any transfer. In the figure, these two conditions are illustrated by indifference curves for the two types through the points $(s,t) = \left(0, \frac{\alpha\beta_L\omega}{1+\alpha\beta_L}\right)$ and $(s,t) = \left(\frac{\beta_H\omega}{1+\beta_H}, 0\right)$, respectively. The two conditions are met for all s smaller than or equal to $\widehat{s}$ respectively $s^{\#\#}$. These critical values of s are defined implicitly by

$$\log\left(1-\frac{\widehat{s}}{\omega}\right) + \beta_L \log\left(1+\frac{\widehat{s}}{\omega}\right) \equiv \beta_L \log\left[\frac{\beta_L\left(1+\alpha E\left(\beta\right)\right)}{E\left(\beta\right)\left(1+\alpha\beta_L\right)}\right] \qquad (4.13)$$

and

$$\log\left(1-\frac{s^{\#\#}}{\omega}\right)+\beta_H\log\left(1+\frac{s^{\#\#}}{\omega}\right)\equiv\beta_H\log\left[\frac{\beta_H\left(1+\alpha E\left(\beta\right)\right)}{\alpha E\left(\beta\right)\left(1+\beta_H\right)}\right]-\log\left(1+\beta_H\right). \tag{4.14}$$

It turns out that we can sustain any $s^* \in \left[0, \min\left\{\widehat{s}, s^{\#\#}\right\}\right]$ as part of a pooling equilibrium. However, it may be that $s^{\#\#} < 0$, in which case pooling equilibria do not exist. These results are summarized in Proposition 4, a formal proof of which is found in the Appendix.

Proposition 4. *(Pooling equilibria) Pooling equilibria exist if and only if* $\alpha \geq \beta_H \varphi\left(\beta_H\right)/E\left(\beta\right)$. *If this inequality is met, then any*

$$\left(s_L^*, t_L^*, s_H^*, t_H^*\right) \in \Omega^{pool} \times \left\{\frac{\alpha E\left(\beta\right)\omega - s^*}{1+\alpha E\left(\beta\right)}\right\} \times \Omega^{pool} \times \left\{\frac{\alpha E\left(\beta\right)\omega - s^*}{1+\alpha E\left(\beta\right)}\right\}, \tag{4.15}$$

where $s_L^* = s_H^* = s^*$ *and* $\Omega^{pool} = \left[0, \min\left\{\widehat{s}, s^{\#\#}\right\}\right]$, *can be sustained as the outcome of a pooling equilibrium.*

When we in the next subsection pose the question whether the allocations induced by the equilibria are efficient, we shall be particularly interested in the case where μ is close to unity. Hence, it will be useful to note a couple of things concerning the set of pooling equilibrium outcomes in this special case. Let us first make the observation that for values of μ close to unity, pooling equilibria exist. This is true since for $\mu = 1$, the condition for existence in Proposition 4 simplifies to $\alpha \geq \varphi\left(\beta_H\right)$, a condition which is guaranteed by Assumption 1. Moreover, as μ approaches unity, $\widehat{s}$ approaches s' and $s^{\#\#}$ approaches $s^{\#}$. These things taken together mean that by choosing μ sufficiently close to unity, we can sustain any $s^* \in \left[0, \min\left\{s', s^{\#}\right\}\right) = \left[0, s'\right)$ as part of a pooling equilibrium.[15] In particular, we can always sustain the outcome where none of the types saves anything and they both get a transfer that is arbitrarily close to $\alpha\beta_H\omega/\left(1+\alpha\beta_H\right)$, i.e., the high type's transfer in the complete information model in Section 3. In other words, if μ is close to unity, there is at least one equilibrium outcome of the incomplete information model that is close to the outcome of the complete information model in Section 3.

[15] We have $\min\left\{s', s^{\#}\right\} = s'$ if $\alpha \geq \psi\left(\beta_L, \beta_H\right)$; see Lemma A5 in the Appendix.

4.3. Efficiency

Let us now investigate whether allocations induced by the equilibria of the game Γ_I are efficient and, if not, whether B saves too little or too much in these equilibria. Similarly to the usage in Section 3, I will refer to a saving-transfer vector (s_L, t_L, s_H, t_H) as an allocation.

In a game with incomplete information the concept of efficiency is not straightforward. In this section of the paper the following definitions will be used. Following Holmström and Myerson (1983) I say that an allocation (s_L, t_L, s_H, t_H) is *incentive feasible* if

$$U_B\left(s_i, t_i \mid \beta_i\right) \geq U_B\left(s_j, t_j \mid \beta_i\right), \qquad \forall i, j \in \{L, H\} \text{ with } i \neq j \tag{4.16}$$

and $(s_L, t_L, s_H, t_H) \in [0, \omega]^4$. An allocation $\left(s_L^{'}, t_L^{'}, s_H^{'}, t_H^{'}\right)$ *ex post dominates* an allocation (s_L, t_L, s_H, t_H) if

$$U_i\left(s_j^{'}, t_j^{'} \mid \beta_j\right) \geq U_i\left(s_j, t_j \mid \beta_j\right), \qquad \forall j \in \{L, H\} \text{ and } \forall i \in \{A, B\}, \tag{4.17}$$

with at least one strict inequality. And an allocation (s_L, t_L, s_H, t_H) is *ex post incentive efficient* if there is no other incentive feasible allocation that ex post dominates this allocation.[16]

It turns out that if the parameters are such that the equilibrium outcome of the corresponding complete information model is not Pareto efficient (i.e., if Assumption 1 holds), then neither an outcome of a separating equilibrium nor an outcome of a pooling equilibrium can be ex post incentive efficient. This is because in a pooling equilibrium the types behave identically, which is not consistent with condition (4.17),[17] and in a separating equilibrium the signaling mechanism does not affect the low type's saving choice (cf. Lemma 1).

[16] Two other possibilities are to make the welfare evaluation ex ante or interim, that is, to evaluate the outcomes when the agents do not know their own types nor the others' types, respectively when the agents *do* know their own types but not the others' types. This gives rise to the alternative notions of ex ante respectively interim incentive efficiency. An allocation that is ex post incentive efficient must be both ex ante and interim incentive efficient; see Holmström and Myerson (1983).

[17] Condition (4.17) requires that if $s^* = 0$, then $t_i^* \in \arg\max_{t \in [0,\omega]} U_A\left(0, t \mid \beta_i\right)$ for both $i = L$ and $i = H$. This in turn requires that $t_L^* \neq t_H^*$, which is not possible in a pooling equilibrium. Condition (4.17) also requires that if $s^* > 0$, then s^* is a solution to *both* $\max_{s \in [0,\omega]} U_B\left(s, t^* \mid \beta_L\right)$ *and* $\max_{s \in [0,\omega]} U_B\left(s, t^* \mid \beta_H\right)$. This is not possible since $\beta_L \neq \beta_H$.

However, in a separating equilibrium the high type's choice is distorted upwards: $s_H^* > 0$; thus, it is conceivable that s_H^* is part of an outcome that is "efficient for the high type." If so, by letting the prior probability that B is the high type, μ, tend to unity we may, from an ex ante point of view, get arbitrarily close to ex post incentive efficiency. Another possibility is that as μ tends to unity, the outcome of a pooling equilibrium approaches an outcome that is "efficient for the high type." If so, we may again get arbitrarily close to ex post incentive efficiency from an ex ante point of view—now in the (slightly weaker) sense that an equlibrium outcome can be sustained that is arbitrarily close to one that is "efficient for the high type," and the probability that we have the high type is arbitrarily close to unity.

By considering the special case where μ tends to unity, we also get arbitrarily close to the benchmark case with complete information. We may hence think of such an analysis as a robustness test of the complete information model. In the following I will carry out this kind of robustness test. That is, I will investigate if and when the high type's saving level in a separating equilibrium is "efficient for the high type," and if and when the saving level in a pooling equilibrium approaches one that is "efficient for the high type" as μ approaches unity.

First consider some formal definitions. An allocation $\left(s_L, t_L, s_H^{'}, t_H^{'}\right)$ *ex post dominates* an allocation (s_L, t_L, s_H, t_H) *for the high type* if

$$U_i\left(s_H^{'}, t_H^{'} \mid \beta_H\right) \geq U_i\left(s_H, t_H \mid \beta_H\right), \qquad \forall i \in \{A, B\}, \tag{4.18}$$

with at least one strict inequality. And an allocation (s_L, t_L, s_H, t_H) is *ex post incentive efficient for the high type* if there is no other incentive feasible allocation that ex post dominates this allocation for the high type.

There are two conditions that are necessary for $(s_L^*, t_L^*, s_H^*, t_H^*)$ to be ex post incentive efficient for the high type as well as the outcome of a separating equilibrium:

$$s_H^* \in \arg\max_{s\in[0,\omega]} U_B\left(s, t_H^* \mid \beta_H\right) \tag{4.19}$$

and

$$t_H^* \in \arg\max_{t\in[0,\omega]} U_A\left(s_H^*, t \mid \beta_H\right). \tag{4.20}$$

The first condition guarantees that the intertemporal allocation is efficient, and the second condition is necessary for $(s_L^*, t_L^*, s_H^*, t_H^*)$ to indeed be the outcome of a separating equilibrium. If an outcome of a pooling equilibrium approaches an outcome that is efficient for the high type as μ approaches unity, then this limit outcome must also meet the two conditions.

It turns out that when $\alpha > (1+\beta_H)^{-1}$, conditions (4.19) and (4.20) are met only for[18]

$$s_H^* = \frac{1-\alpha(1-\beta_H)}{1+\alpha(1+\beta_H)}\omega \equiv \widetilde{s}_I \tag{4.21}$$

and

$$t_H^* = \frac{\alpha(1+\beta_H)-1}{1+\alpha(1+\beta_H)}\omega \equiv \widetilde{t}_I. \tag{4.22}$$

When $\alpha \leq (1+\beta_H)^{-1}$, conditions (4.19) and (4.20) are met only for $s_H^* = \frac{\beta_H}{1+\beta_H}\omega \equiv \widetilde{s}_{II}$ and $t_H^* = 0 \equiv \widetilde{t}_{II}$. One may verify that when $\alpha > (1+\beta_H)^{-1}$, $\left(s_L^*, t_L^*, \widetilde{s}_I, \widetilde{t}_I\right)$ is indeed ex post incentive efficient for the high type; and when $\alpha \leq (1+\beta_H)^{-1}$, $\left(s_L^*, t_L^*, \widetilde{s}_{II}, \widetilde{t}_{II}\right)$ is indeed ex post incentive efficient for the high type.

In order to tell if and when the high type's saving level in a separating equilibrium is efficient for the high type, we need to know, among other things, for what subset of the parameter space $\widetilde{s}_I \in \left[s', s''\right]$. By using the definition of s'', one may show that $\widetilde{s}_I \leq s''$ is equivalent to $\beta_L \leq g(\beta_H, \alpha)$, where

$$g(\beta_H, \alpha) \equiv \left[\frac{1+\alpha(1+\beta_H)}{2\beta_H}\left(\frac{1+\alpha(1+\beta_H)}{2\alpha}\right)^{\frac{1}{\beta_H}} - \alpha\right]^{-1}. \tag{4.23}$$

One may also show (see the proof of Proposition 5) that $\widetilde{s}_I \geq s'$ if and only if $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$, where the function $\widetilde{\alpha}$ is defined implicitly by

$$\log\left[\frac{1+\widetilde{\alpha}(1+\beta_H)}{2\widetilde{\alpha}}\right] + \beta_L \log\left[\frac{\beta_L\left[1+\widetilde{\alpha}(1+\beta_H)\right]}{2\beta_H(1+\widetilde{\alpha}\beta_L)}\right] \equiv 0. \tag{4.24}$$

We are now ready to state Proposition 5, which is proven in the Appendix.

Proposition 5. *Suppose that $\alpha \leq (1+\beta_H)^{-1}$. Then an allocation that is ex post incentive efficient for the high type can be sustained as the outcome of a separating equilibrium if and only if $\beta_L \leq f(\beta_H, \alpha)$. If $\alpha > (1+\beta_H)^{-1}$,*

[18] The assumption $\alpha < (1-\beta_L)^{-1}$ (i.e., Assumtion 1) guarantees that $\widetilde{s}_I > 0$, and the assumption $\alpha > (1+\beta_H)^{-1}$ guarantees that $\widetilde{t}_I > 0$.

then an allocation that is ex post incentive efficient for the high type can be sustained as the outcome of a separating equilibrium if and only if $\beta_L \leq g(\beta_H, \alpha)$ *and* $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$.

Figure 6 illustrates the results stated in the proposition. This figure is similar to Figure 4, but it also shows the graph of the functions $\alpha = g^{-1}(\beta_L; \beta_H)$ and $\alpha = \widetilde{\alpha}(\beta_L, \beta_H)$.[19] To start with, consider the case where $\alpha \leq (1+\beta_H)^{-1}$. For these values of α we want to sustain the saving level $\widetilde{s}_{II}$ as part of the outcome of a separating equilibrium. We know from Proposition 5 that we can do this if and only if $\beta_L \leq f(\beta_H, \alpha)$. That is, we need to be in the region labeled (5.1) in Figure 6.

Now turn to the case where $\alpha > (1+\beta_H)^{-1}$. For these values of α we want to sustain the saving level $\widetilde{s}_I$ as part of the outcome of a separating equilibrium. This is possible in the region labeled (5.2) in Figure 6. However, in the subset of the parameter set satisfying Assumption 1, there are two regions in which we cannot sustain an outcome that is ex post incentive efficient for the high type. The first one of these, region (5.3), is obtained when β_L is sufficiently close to β_H. In this region all outcomes that can be sustained as outcomes of a separating equilibrium involve undersaving. The second region, labeled (5.4) in the figure, is obtained when $\alpha > \widetilde{\alpha}(\beta_L, \beta_H)$. In this region all separating equilibria involve *oversaving* for the high type. However, it turns out that if $\alpha > \widetilde{\alpha}(\beta_L, \beta_H)$, then there is an outcome of a pooling equilibrium that, as μ tends to unity, gets arbitrarily close to the one that is efficient for the high type.

To see the implications of these results, think of the Samaritan's-dilemma model with complete information considered in Section 3. Suppose that the parameter β in that model equals some $\beta = \beta_H \in (0,1)$. We know that if α is neither too small nor too large, then the outcome of the unique subgame perfect equilibrium of this model is inefficient. Now suppose that we add to this model a slight amount of uncertainty, and we do this in the following particular way: with some small probability, the parameter β does not equal β_H but β_L. Furthermore, β_L is sufficiently much smaller than β_H. Since the probability that β equals β_L and not β_H can be made arbitrarily small, this model is very close to the model

[19] Some properties of these functions that are helpful when drawing the figure are proven in the Appendix (Lemma A11 resp. A12).

with complete information. Still, in the new model with incomplete information one can sustain an outcome that is (almost) ex post incentive efficient.

One should bear in mind, however, that the almost efficient outcome is only one of many possible equilibrium outcomes. When an equilibrium outcome is not efficient, it may involve either undersaving or oversaving on the part of B. In particular, as noted in Subsection 4.2, for values of μ close to unity there always exists a pooling equilibrium outcome that is arbitrarily close to the equilibrium outcome of the complete information model. In the next subsection I will consider the question what can be said about efficiency and undersaving when we impose some equilibrium refinements that are often used in signaling models.

4.4. Equilibrium Refinements

In Subsection 4.3 it was shown that if β_L is sufficiently small, then an "almost efficient" level of B's savings can be sustained as part of an equilibrium. However, to some extent this possibility relies on our choosing beliefs off the equilibrium path in a particular manner. To see this, consider Figure 3a and some saving level $s_H^* \in \left(s^{'}, s^{''}\right]$. In order to sustain this as the high type's saving level in a separating equilibrium, A must assign positive probablity to B's being the low type when observing any saving level $s \in \left[s^{'}, s_H^*\right)$; otherwise the high type would have an incentive to deviate to any such saving level s. However, it is clear from the figure that regardless of which posterior beliefs $\widetilde{\mu}(s) \in [0,1]$ A holds, the low type is always strictly better off choosing $s = 0$ than choosing any saving level $s \in \left(s^{'}, s^{''}\right]$. Hence one may argue that any beliefs assigning positive probability to the low type's saving $s \in \left(s^{'}, s^{''}\right]$ are unreasonable. But if we do not allow such beliefs the high type will deviate, and we cannot sustain s_H^* as the high type's saving level in an equilibrium. The same reasoning applies for all $s_H^* \in \left(s^{'}, s^{''}\right]$. In Figure 3a, the only saving level of the high type that survives this refinement is $s_H^* = s^{'}$.

Concerning the set of separating equilibrium outcomes that we derived using Figure 3b, the kind of reasoning described in the previous paragraph works in the same way. Again there is only one surviving outcome. Here this outcome involves the high type saving $s_H^* = s^{'}$ if $\alpha \geq \psi(\beta_L, \beta_H)$ and otherwise $s_H^* = \frac{\beta_H \omega}{1+\beta_H}$. However, this kind of refinement does not have the same bite on the set of pooling equilibria. In order to rule out all pooling equilibria we must invoke some stronger equilibrium refinement. One refinement that will do the job is the so-called intuitive criterion (Cho and Kreps, 1987). Concerning the particular model considered here, this criterion says the following. Fix some equilibrium of the game Γ_I. Suppose that, for any possible best response by A, the equilibrium utility of B when being the low type is strictly greater than the utility the low type would receive if he saved some out-of-equilibrium amount s, regardless of A's beliefs. Moreover, suppose that this is not true for the high type; that is, there are some beliefs $\widetilde{\mu}(s)$ such that, if A makes a best response given these beliefs, the high type would get a higher payoff if choosing s instead of s_H^*. Then, if A observes the saving level s, A should place zero probability on the possibility that B is the low type, $\widetilde{\mu}(s) = 1$. This requirement must hold for all out-of-equilibrium saving levels s. If it does, we say that the equilibrium is *intuitive.* The following proposition is proven in the Appendix.

Proposition 6. *There is a unique intuitive equilibrium outcome of Γ_I. If $\alpha \geq \psi(\beta_L, \beta_H)$, then this outcome is $(s_L^*, t_L^*, s_H^*, t_H^*) = \left(0, \frac{\alpha\beta_L\omega}{1+\alpha\beta_L}, s', \frac{\alpha\beta_H\omega - s'}{1+\alpha\beta_H}\right)$; otherwise this outcome is $(s_L^*, t_L^*, s_H^*, t_H^*) = \left(0, \frac{\alpha\beta_L\omega}{1+\alpha\beta_L}, \frac{\beta_H\omega}{1+\beta_H}, 0\right)$.*

By imposing the intuitive criterion we thus get a unique prediction of the model with incomplete information. This prediction says that, depending on how α relates to the critical level $\psi(\beta_L, \beta_H)$, the high type saves either $s_H^* = s'$ or $s_H^* = \frac{\beta_H\omega}{1+\beta_H}$. Figure 7 illustrates the results. In the region where Assumption 1 holds and $\alpha < \psi(\beta_L, \beta_H)$ (labeled (6.1) in the figure), the high type saves $s_H^* = \frac{\beta_H\omega}{1+\beta_H}$. This amount of savings is efficient for the high type, since $\alpha < \psi(\beta_L, \beta_H)$ implies $\alpha \leq (1+\beta_H)^{-1}$.[20] For $\alpha \geq \psi(\beta_L, \beta_H)$, the high type saves $s_H^* = s'$. This amount of savings is efficient for the high type only if $\alpha = \widetilde{\alpha}$. For α below this critical value, the high type saves too little (the region labeled (6.2)); and for α above this critical value, the high type saves too much (the region labeled (6.3)).

This means that for a subset of the parameter space we can still sustain an (almost) efficient outcome, although this subset is smaller than the corresponding subset without imposing the refinement. Moreover, even where we cannot sustain

[20] However, this argument relies on the presumption that $\psi(\beta_L, \beta_H) < (1+\beta_H)^{-1}$ for all $\beta_L \in (0, \beta_H)$, a result which I have not proven analytically. Cf. note 14.

an almost efficient outcome (i.e., in the regions (6.2) and (6.3)), we now have a unique prediction which is qualitatively different from the one of the complete information model. Interestingly, it can also happen (in region (6.3)) that the high type oversaves.

5. Concluding Discussion

The Samaritan's dilemma—i.e., the idea that, in the presence of altruism, people may choose to save (or work or insure) to a too small extent—is certainly very intuitive. In the alternative formulation of the Samaritan's dilemma considered in this paper there is an additional effect present, which counteracts the undersaving effect. The mechanics of this new force should also—once we have become aware of them—be very intuitive. For the new force to indeed work in the "right" direction (i.e., to counteract the undersaving effect), the following condition must hold. Suppose that B has private information about some parameter x and that he, everything else being equal, has an incentive to save *more* when knowing that x is high (low). Then, believing that x is high (low) will induce A to make the transfer to B *larger*. Since B wants the transfer to be large, he would like A to believe that x is high (low); and he can try to make A believe this by saving more.

The condition is met if, as was suggested in the Introduction, x is a measure of B's second-period income or if, as was assumed in the formal model of Section 4, x is a discount factor or a weight on B's second-period utility. One may wonder whether the presence of the counteracting effect is hinging on the assumption that the incomplete information concerns one of these two particular characteristics of B.[21] What if B had private information about the return on his savings or about his first-period income?

If the parameter x is interpreted as the return on B's savings and if we stick to the log-utility specification in the present paper, then it is clear that there would not be any counteracting force present. This is because with log-utility and with A's transfer t being equal to zero, the optimal saving level is independent of the

[21] However, one should remember that, since this paper is mainly concerned with the *robustness* of the Samaritan's dilemma, it suffices, in principle, to find *one* example of a parameter having the property that the model's prediction makes a discrete jump when we add a small amount of uncertainty about the parameter.

return; and if t is positive, then the optimal saving level is increasing with the return. However, if the intertemporal elasticity of substitution is constant but sufficiently less than one (or, equivalently, if the degree of relative risk aversion is sufficiently greater than one), then B will have an incentive to save more when knowing that the return is low; and A will also have an incentive to make her transfer larger when believing that the return is low. Hence, under this assumption, one would again get counteracting signaling. The assumption about the intertemporal elasticity of substitution seems reasonable: the log-utility assumption in the present paper was made for tractability reasons, and there is empirical evidence that this elasticity is indeed less than one.

Private information about B's first-period income does of course not give rise to any opportunity to signal as long as B's utility function is additively separable over time, since then the size of B's income in the first period does not affect A's incentive to transfer income to him in the second period. However, if B's marginal utility of second-period consumption is increasing with B's first-period consumption, then the condition above is again satisfied. This requirement on the sign of the cross derivative of the utility function is, for instance, met for the following preferences: $U_B(c_{1B}, c_{2B}) = (c_{1B})^a (c_{2B})^b$ for some $a, b > 0$.

Yet another parameter in the model which there could conceivably be uncertainty about is the altruism parameter, α.[22] In the model analyzed in this paper, A's having private information about α would not give rise to any signaling, since A is acting last in the game. However, this is not true for the formulation of the Samaritan's dilemma considered in Lindbeck and Weibull (1988). In that model there are two individuals who are altruistic toward each other. They both, simultaneously, make a saving decision in period one. In period two they observe the other one's saving decision and then, simultaneously, decide how much (if anything) to transfer to each other. If they both are equally wealthy, then, in equilibrium, only the individual who is more altruistic will make a positive transfer. Anticipating this, the less altruistic individual will undersave in the first period. If one to this setting added the assumption that one or both of the individuals have incomplete information about the other one's degree of altruism,

[22] Uncertainty about the degree of altruism has been modeled by Chakrabarti, Lord, and Rangazas (1993) and Lord and Rangazas (1995).

then one should expect the undersaving to be exacerbated, the reason being that both individuals would like to signal that they are less altrusitic than the other one, and a person who indeed has a low altruism parameter does not expect a transfer and should therefore save a lot on his own.

Returning to the setting of Section 4 of the present paper, we saw that the presence of an arbitrarily small amount of incomplete information makes it possible to sustain a large number of different outcomes as an equilibrium, and these outcomes typically involve both undersaving and oversaving. Although full efficiency cannot be obtained, one may in a subset of the parameter space get arbitrarily close to ex post incentive efficiency. Moreover, if the difference between the two types is big enough, this is possible for all values of the altrusim parameter for which we get ineffficiency in the complete information model. There is of course no guarantee that the (almost) efficient outcome will be played. Indeed, among the equilibrium outcomes there is also a pooling one which is close to the unique outcome of the complete information model and which involves undersaving. However, as noted in connection to the example in Section 2, the results show that the theoretical case for undersaving is much weaker than what one might think if only looking at the standard formulation of the Samaritan's dilemma. In this context it is also interesting to recall the school of thought that is usually associated with the University of Chicago and especially with Gary Becker, which suggests that there are strong forces in action making resource allocations efficient, in particular within the family.

Of course there are some caveats. One of them is the fact that we carried out the analyis of the signaling game under the assumption that the magnitude of the altruism parameter is such that we get inefficiency in the corresponding complete information model (Assumption 1). Doing this actually gives rise to a bias in favor of smaller inefficiencies since, if we have a first-best outcome without signaling, incomplete information can only make things worse. Hence, had we investigated also the subset of the parameter space not satisfying Assumption 1, we would have found more of oversaving. Another caveat is the fact that, to some extent, the (almost) efficient equilibrium outcome is sensitive to equilibrium refinements such as the intuitive criterion. Although the choice of a refinement is always controversial,[23] it is interesting to see what the intuitive criterion implies for the present model. What we found in Subsection 4.4 was that the intuitive criterion indeed gives us a unique equilibrium outcome, and this outcome is never close to the outcome of the complete information model. This means that if we increase the amount of uncertainty about the recipient's type, we will get a discontinuity at zero. That is, the unique equilibrium outcome always involves more saving than in the complete information model. Depending on parameter values, we can either get undersaving, (almost) efficiency, or oversaving. The (almost) efficiency result is obtained when the altruism parameter is sufficiently low, but still large enough to give rise to undersaving in the complete information model. The subset of the parameter space where we get this result is the same as for which we get a unique separating equilibrium outcome of the model without any refinement.[24] In Section 4 we saw that the reason why we get this uniqueness is closely related to the fact that there is a non-negativity constraint on the transfer.

As noted in the Introduction, the mechanism in the standard Samaritan's dilemma model and in particular the undersaving result has been employed in an extensive literature, addressing various issues. Although these models are not identical to the benchmark model of the present paper, the basic mechanism is the same. Hence, one should expect the undersaving result also in those other models to be sensitive to the assumption that information is complete. One may object that the literature cited in the Introduction in many cases concerns an efficiency argument; the arguments go through even if the undersaving is mitigated as long as it is not exactly efficient. Still this is not always the case. For example, the argument in O'Connell and Zeldes (1993) (see the Introduction) concerns the actual level of savings and the question how it relates to the real interest rate in the economy. An interesting topic for future research would be to investigate the signaling mechanism in the present paper in a setting that is closer to the one in the existing literature, in order to find out to what extent those results indeed

[23] One reason for this is that the intuitive criterion and other refinements that are based on equilibrium domination are sensitive to the so-called Stiglitz critique. See e.g. Cho and Kreps (1987) or Kreps and Sobel (1994).

[24] Yet another caveat is that I have not proven analytically that the subset of the parameter space where the almost efficiency result is obtained (region 6.1 in Figure 7) never is empty. However, I have not been able find any counter-example when plotting the relevant functions on a computer.

are sensitive to the complete information assumption. Another interesting topic for future research would be to investigate whether the mechanism in this paper might help resolving other problems of time inconsistency.

6. Appendix

Lemma A1. $\varphi^{'}(\beta) < 0$ *for all* $\beta \in (0,1)$.

Proof of Lemma A1: In order to prove Lemma A1 I will first prove two other claims. The first claim will be used when proving the second claim, and the second claim will be used when proving the lemma. Let the function $\kappa : (0,1) \rightarrow \Re$ be defined by $\kappa(\beta) = \log(1+\beta) - \beta(1+\beta)^{-\frac{1}{2}}$. I claim that $\kappa(\beta) < 0$. To prove this, first note that $\lim_{\beta\rightarrow 0}\kappa(\beta) = 0$. It thus suffices to show that $\kappa^{'}(\beta) < 0$. Differentiating $\kappa(\beta)$ yields

$$\kappa^{'}(\beta) = (1+\beta)^{-1} - (1+\beta)^{-\frac{1}{2}} + \frac{\beta}{2}(1+\beta)^{-\frac{3}{2}}. \tag{6.1}$$

Rewriting this expression for $\kappa^{'}(\beta)$ yields

$$\begin{aligned}
\kappa^{'}(\beta) &= \left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]^{-1}\left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]\times \\
&\quad \times\left[(1+\beta)^{-1} - (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right] \\
&= \left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]^{-1}\left[(1+\beta)^{-2} - (1+\beta)^{-1}\left[\frac{2+\beta}{2(1+\beta)}\right]^{2}\right] \\
&= -\left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]^{-1}\frac{\beta^{2}}{4(1+\beta)^{3}}. \qquad (6.2)
\end{aligned}$$

This expression is strictly negative, which proves the claim that $\kappa(\beta) < 0$.

Now let the function $\xi : (0,1) \rightarrow \Re$ be defined by $\xi(\beta) = (1+\beta)^{\frac{1+\beta}{\beta}}$. I claim that $\xi^{'}(\beta) > 1$. In order to prove this, first differentiate ξ:

$$\xi^{'}(\beta) = \frac{\xi}{\beta^{2}}\left[\beta - \log(1+\beta)\right]. \tag{6.3}$$

Note that $\lim_{\beta\to 1}\xi^{'}(\beta) = 4 - \log(16) > 1$. It thus suffices to show that $\xi^{''}(\beta) < 0$. Differentiating ξ once more yields

$$\xi^{''}(\beta) = \frac{\beta^2\xi^{'} - 2\beta\xi}{\beta^4}\left[\beta - \log(1+\beta)\right] + \frac{\xi}{\beta(1+\beta)}. \tag{6.4}$$

Substituting the expression for $\xi^{'}$ in (6.3) into (6.4) and then rewriting yield

$$\xi^{''}(\beta) = \frac{\xi}{\beta^4}\left[\log(1+\beta) + \frac{\beta}{\sqrt{1+\beta}}\right]\left[\log(1+\beta) - \frac{\beta}{\sqrt{1+\beta}}\right]. \tag{6.5}$$

Hence, by the first claim that $\kappa(\beta) < 0$, $\xi^{''}(\beta) < 0$. We have thus proven the claim that $\xi^{'}(\beta) > 1$.

We are now ready to prove Lemma A1. Differentiating the expression in (3.6) yields

$$\varphi^{'}(\beta) = -\frac{\left[\xi^{'}(\beta) - 1\right]}{\left[\xi(\beta) - \beta\right]^2}. \tag{6.6}$$

Thus, by the claim that $\xi^{'}(\beta) > 1$, $\varphi^{'}(\beta) < 0$. □

Proof of Proposition 1: In order to prove Proposition 1 we must investigate the case where $\alpha\beta < 1$, since this was not done in Subsection 3.2. When $\alpha\beta < 1$, B's indirect utility is given by either one of the two lines in equation (3.4), and which one it is depends on the choice of s itself. Let Θ^1 be defined by

$$\Theta^1 = \max_{s\in[0,\alpha\beta\omega)} \log(\omega - s) + \beta\log(\omega + s) + \beta\log\left(\frac{\alpha\beta}{1+\alpha\beta}\right). \tag{6.7}$$

That is, Θ^1 is the highest utility B can obtain if he is constrained to choose some $s \in [0, \alpha\beta\omega)$. Denote the solution to the maximization problem in (6.7) (it does exist) by s^1. Since $\beta < 1$, we have $s^{^1} = 0$. Hence

$$\Theta^1 = (1+\beta)\log\omega + \beta\log\left(\frac{\alpha\beta}{1+\alpha\beta}\right) \tag{6.8}$$

Similarly define Θ^2:

$$\Theta^2 = \max_{s\in[\alpha\beta\omega,\omega]} \log(\omega - s) + \beta\log(s). \tag{6.9}$$

Thus, Θ^2 is the highest utility B can obtain if he is constrained to choose some $s \in [\alpha\beta\omega, \omega]$. Denote the solution to this problem by s^2. We have

$$s^2 = \begin{cases} \frac{\beta}{1+\beta}\omega & \text{for } \alpha \leq \frac{1}{1+\beta} \\ \alpha\beta\omega & \text{otherwise.} \end{cases} \tag{6.10}$$

Thus

$$\Theta^2 = \begin{cases} (1+\beta)\log\left(\frac{\omega}{1+\beta}\right) + \beta\log(\beta) & \text{for } \alpha \leq \frac{1}{1+\beta} \\ (1+\beta)\log(\omega) + \log(1-\alpha\beta) + \beta\log(\alpha\beta) & \text{otherwise.} \end{cases} \tag{6.11}$$

In order to find out what B's optimal choice of s will be, we must compare the magnitudes of Θ^1 and Θ^2. There are two cases: (a) $\alpha \leq (1+\beta)^{-1}$ and (b) $\alpha > (1+\beta)^{-1}$. Consider case (a). Carrying out some algebra yields

$$\Theta^1 \gtreqless \Theta^2 \Leftrightarrow \alpha \gtreqless \varphi(\beta). \tag{6.12}$$

Hence, if $\alpha > \varphi(\beta)$, then $s^* = 0$; and if $\alpha < \varphi(\beta)$, then $s^* = \frac{\beta}{1+\beta}\omega$. If $\alpha = \varphi(\beta)$, then there are two solutions two B's optimization problem, $s^* = 0$ and $s^* = \frac{\beta}{1+\beta}\omega$. From equation (3.3) one obtains the result that when $s^* = 0$, $t^* = \frac{\alpha\beta\omega}{1+\alpha\beta}$; and when $s^* = \frac{\beta}{1+\beta}\omega$, then $t^* = 0$.

Now consider case (b), $\alpha > (1+\beta)^{-1}$. Here we get

$$\begin{aligned} \Theta^1 &> \Theta^2 \Leftrightarrow \log(1-\alpha\beta) + \beta\log(1+\alpha\beta) < 0 \\ &\Leftrightarrow (1-\beta)\log(1-\alpha\beta) + \beta\log\left[1-(\alpha\beta)^2\right] < 0 \end{aligned} \tag{6.13}$$

Inequality (6.13) always holds since $\beta < 1$ and $\alpha\beta < 1$. Hence, we have $s^* = 0$ and $t^* = \frac{\alpha\beta\omega}{1+\alpha\beta}$. Together with the analysis in Subsection 3.2, this completes the proof of Proposition 1. □

Proof of Proposition 2: First assume that $\alpha \leq \varphi(\beta)$. One may then verify that $(s,t) = \left(\frac{\beta}{1+\beta}\omega, 0\right)$ is the unique solution to the following problem:

$$\max_{(t,s)\in[0,\omega]^2} U_A(s,t). \tag{6.14}$$

(This is the solution as long as $\alpha \leq (1+\beta)^{-1}$, which is implied by $\alpha \leq \varphi(\beta)$.) That is, this allocation is A's first-best choice. But for $\alpha < \varphi(\beta)$, $(s,t) =$

$\left(\frac{\beta}{1+\beta}\omega,0\right)$ is also the outcome of the unique subgame-perfect equilibrium; hence, this outcome is Pareto efficient. Similarly, for $\alpha = \varphi(\beta)$, $(s,t) = \left(\frac{\beta}{1+\beta}\omega,0\right)$ is one of two possible equilibrium outcomes and it is Pareto efficient. Now assume that $\alpha \geq (1-\beta)^{-1}$. One may then verify that $(s,t) = \left(0, \frac{\alpha\beta}{1+\alpha\beta}\omega\right)$ is the unique solution to the problem in (6.14). Again, this is also the outcome of the unique subgame perfect equilibrium, so the outcome is Pareto efficient. It remains to show that the allocation $(s,t) = \left(0, \frac{\alpha\beta}{1+\alpha\beta}\omega\right)$ is not Pareto efficient for any $\alpha \in \left[\varphi(\beta), (1-\beta)^{-1}\right)$. Suppose that $\alpha < (1-\beta)^{-1}$. Differentiate U_B with respect to s and evaluate at $(s,t) = \left(0, \frac{\alpha\beta}{1+\alpha\beta}\omega\right)$:

$$\frac{\partial U_B(s,t)}{\partial s}\,|_{(s,t)=\left(0,\frac{\alpha\beta}{1+\alpha\beta}\omega\right)} = \frac{1}{\alpha\omega}\left[1-\alpha(1-\beta)\right] > 0. \tag{6.15}$$

Moreover, since $\frac{\partial U_A(s,t)}{\partial s} = \alpha\frac{\partial U_B(s,t)}{\partial s}$, also A can be made better off if B saved more given the level of A's transfer. Hence the allocation $(s,t) = \left(0, \frac{\alpha\beta}{1+\alpha\beta}\omega\right)$ is not Pareto efficient for any $\alpha \in \left[\varphi(\beta), (1-\beta)^{-1}\right)$. $\square$

Proof of Lemma 1: Suppose that $\alpha > \varphi(\beta_L)$ and that $s_L^* > 0$ in a separating equilibrium. I will show that this leads to a contradiction. To start with, consider the case where $s_L^* \in (0, \alpha\beta_L\omega]$. Then the low type receives a transfer from A according to the first line in equation (3.3) but with β_L substituted for β. The low type's utility is accordingly given by (cf. the first line of equation (3.4)):

$$V' = \log(\omega - s_L^*) + \beta_L \log(\omega + s_L^*) + \beta_L \log\left(\frac{\alpha\beta_L}{1+\alpha\beta_L}\right). \tag{6.16}$$

However, if the low type instead chose $s = 0$, he would receive a transfer of at least $\frac{\alpha\beta_L\omega}{1+\alpha\beta_L}$, which would give him the following utility (cf. again the first line of equation (3.4)):

$$V = (1+\beta_L)\log(\omega) + \beta_L \log\left(\frac{\alpha\beta_L}{1+\alpha\beta_L}\right).$$

V is strictly greater than V' for all $s_L^* \in (0, \alpha\beta_L\omega]$, since V' is strictly decreasing in s_L^*. Now consider the case where $s_L^* \in (\alpha\beta_L\omega, \omega]$. Then the low type receives a transfer from A according to the second line in equation (3.3) but with β_L

substituted for β. However, as long as $\alpha > \varphi(\beta_L)$, the low type is strictly better off from choosing $s = 0$ than from choosing any $s_L^* \in (\alpha\beta_L\omega, \omega]$. This follows from the proof of Proposition 1. We thus have a contradiction, which proves the lemma. □

Lemma A2. $s' < \alpha\beta_H\omega$.

Proof of Lemma A2: It is obvious that the lemma is true for $\alpha\beta_H \geq 1$. Suppose that $\alpha\beta_H < 1$. Then $s' < \alpha\beta_H\omega$ if and only if the left-hand side of (4.8) evaluated at $s' = \alpha\beta_H\omega$ is smaller than the righ-hand side:

$$\log(1-\alpha\beta_H) + \beta_L \log(1+\alpha\beta_H) < \beta_L \log\left(\frac{\beta_L(1+\alpha\beta_H)}{\beta_H(1+\alpha\beta_L)}\right) \tag{6.17}$$

or equivalently $\Upsilon(\alpha, \beta_L, \beta_H) < 0$, where the function $\Upsilon : \Re_{++} \times (0, \beta_H) \times \left(\beta_L, \frac{1}{\alpha}\right) \to \Re$ is defined by

$$\Upsilon(\alpha, \beta_L, \beta_H) = \log(1-\alpha\beta_H) + \beta_L \log\left(\frac{\beta_H(1+\alpha\beta_L)}{\beta_L}\right). \tag{6.18}$$

Note that Υ is strictly concave in its third argument:

$$\frac{\partial^2 \Upsilon(\alpha, \beta_L, \beta_H)}{\partial \beta_H^2} = \frac{-\alpha^2}{(1-\alpha\beta_H)^2} - \frac{\beta_L}{\beta_H^2} < 0. \tag{6.19}$$

Also note that as β_H tends to $\frac{1}{\alpha}$, $\Upsilon(\alpha, \beta_L, \beta_H)$ tends to minus infinity; and as β_H tends to β_L, $\Upsilon(\alpha, \beta_L, \beta_H)$ tends to some number that is also strictly negative:

$$\begin{aligned}\lim_{\beta_H \to \beta_L} \Upsilon(\alpha, \beta_L, \beta_H) &= \log(1-\alpha\beta_L) + \beta_L \log(1+\alpha\beta_L) \\ &= \log\left(1-(\alpha\beta_L)^2\right) - (1-\beta_L)\log(1+\alpha\beta_L) < 0.\end{aligned} \tag{6.20}$$

Hence, if $\Upsilon(\alpha, \beta_L, \beta_H) < 0$ for β_H satisfying

$$\frac{\partial \Upsilon(\alpha, \beta_L, \beta_H)}{\partial \beta_H} = 0 \Leftrightarrow \beta_H = \frac{\beta_L}{\alpha(1+\beta_L)}, \tag{6.21}$$

then $\Upsilon(\alpha, \beta_L, \beta_H) < 0$ for all β_H. Moreover, Υ is strictly decreasing in α. It thus suffices to show that evaluated at $\alpha = \varphi(\beta_L)$ and $\beta_H = \beta_L[\varphi(\beta_L)(1+\beta_L)]^{-1}$,

$\Upsilon(\alpha, \beta_L, \beta_H) \leq 0$. We have:

$$\Upsilon\left(\varphi(\beta_L), \beta_L, \frac{\beta_L}{\varphi(\beta_L)(1+\beta_L)}\right) = \log\left(1 - \frac{\beta_L}{(1+\beta_L)}\right) + \beta_L \log\left(\frac{(1+\beta_L)^{\left(\frac{1+\beta_L}{\beta_L}\right)}}{1+\beta_L}\right) = 0. \tag{6.22}$$

This proves Lemma A2. □

Lemma A3. *$s'' < \alpha\beta_H\omega$ if and only if either (i) $\alpha\beta_H \geq 1$ or (ii) $\alpha\beta_H < 1$ and $\beta_L > f(\beta_H, \alpha)$.*

Proof of Lemma A3: If $\alpha\beta_H \geq 1$, then it is obvious that $s'' < \alpha\beta_H\omega$. Suppose that $\alpha\beta_H < 1$. Then $s'' < \alpha\beta_H\omega$ if and only if the left-hand side of (4.9) evaluated at $s'' = \alpha\beta_H\omega$ is smaller than the righ-hand side of (4.9):

$$\log(1 - \alpha\beta_H) + \beta_H \log(1 + \alpha\beta_H) < \beta_H \log\left(\frac{\beta_L(1+\alpha\beta_H)}{\beta_H(1+\alpha\beta_L)}\right). \tag{6.23}$$

Rewriting (6.23), we get $\beta_L > f(\beta_H, \alpha)$. □

Lemma A4. $s^\circ > \frac{\beta_H\omega}{1+\beta_H}$.

Proof of Lemma A4: Suppose per contra that $s^\circ \leq \frac{\beta_H\omega}{1+\beta_H}$. Then the high type's indifference curve through the point $(s,t) = \left(\frac{\beta_H\omega}{1+\beta_H}, 0\right)$ must be above the s-axis at $s = \frac{\beta_H\omega}{1+\beta_H}$ (cf. Figure 3b). The algebraic expression for this indifference curve is

$$t_H(s) = \exp\left[\frac{K_H - \log(\omega - s)}{\beta_H}\right] - s, \qquad \text{where } K_H \equiv \log(\omega) + \beta_H \log\left(\frac{\alpha\beta_L\omega}{1+\alpha\beta_L}\right). \tag{6.24}$$

Hence we must have

$$t_H\left(\frac{\beta_H\omega}{1+\beta_H}\right) \geq 0 \Leftrightarrow \beta_H \log\left(\frac{\beta_H(1+\alpha\beta_L)}{\beta_L}\right) \leq \beta_H \log(\alpha) + (1+\beta_H)\log(1+\beta_H). \tag{6.25}$$

However, we must also have $\beta_L \leq f(\beta_H, \alpha)$, which is equivalent to

$$\beta_H \log\left(\frac{\beta_H(1+\alpha\beta_L)}{\beta_L}\right) \geq \log\left(\frac{1}{1-\alpha\beta_H}\right). \tag{6.26}$$

Hence, the right-hand side of (6.26) must not be greater than the right-hand side of (6.25), or

$$\beta_H \log(\alpha) + \log(1 - \alpha\beta_H) + (1 + \beta_H)\log(1 + \beta_H) \geq 0. \tag{6.27}$$

This inequality is not consistent with the requirement that $\alpha < (1 + \beta_H)^{-1}$, since for these values of α the left-hand side of (6.27) is strictly increasing in α and the inequality does not hold for $\alpha = (1 + \beta_H)^{-1}$. Thus we have a contradiction and we can conclude that $s^{\circ} > \frac{\beta_H \omega}{1+\beta_H}$. □

Lemma A5. *There is a function ψ such that $s^{\#} \lesseqgtr s^{'}$ as $\alpha \lesseqgtr \psi(\beta_L, \beta_H)$. Moreover, $\psi(\beta_L, \beta_H) > \varphi(\beta_H)$ for all $\beta_L \in (0, \beta_H)$, and*

$$\lim_{\beta_L \to 0} \psi(\beta_L, \beta_H) = \lim_{\beta_L \to \beta_H} \psi(\beta_L, \beta_H) = \varphi(\beta_H). \tag{6.28}$$

Proof of Lemma A5: By carrying out some straightforward calculations, one can show that $\partial s^{\#}/\partial\alpha > 0$ and that $s^{\#} = 0$ if $\alpha = \varphi(\beta_H)$. Also, $\partial s^{'}/\partial\alpha < 0$ and that $s^{'} > 0$ for all $\alpha > 0$. Hence there is exactly one value of α for which $s^{\#} = s^{'}$. For some given β_L and β_H, let $\psi(\beta_L, \beta_H)$ denote this value of α. The function ψ is implicitly defined by

$$\log\left(1 - \frac{s}{\omega}\right) + \beta_L \log\left(1 + \frac{s}{\omega}\right) \equiv \beta_L \log\left(\frac{\beta_L(1 + \psi\beta_H)}{\beta_H(1 + \psi\beta_L)}\right) \tag{6.29}$$

and

$$\log\left(1 - \frac{s}{\omega}\right) + \beta_H \log\left(1 + \frac{s}{\omega}\right) \equiv \beta_H \log\left(\frac{1 + \psi\beta_H}{\psi}\right) - (1 + \beta_H)\log(1 + \beta_H). \tag{6.30}$$

It also follows that $\psi(\beta_L, \beta_H) > \varphi(\beta_H)$ for all $\beta_L \in (0, \beta_H)$. Moreover, since $s^{'}$ tends to zero as $\beta_L \to 0$ and as $\beta_L \to \beta_H$, we get the equalities in (6.28). □

Proof of Proposition 4: I start with proving the claim about existence. Let A's beliefs be given by $\widetilde{\mu}(s) = \mu$ for $s = s^*$ and $\widetilde{\mu}(s) = 0$ for all $s \neq s^*$. These beliefs are clearly the "strongest" possible in that they will generate all the equilibrium outcomes that may exist. For the low type not to have an incentive to deviate from $s = s^*$, we must have

$$U_B\left(0, \frac{\alpha\beta_L\omega}{1 + \alpha\beta_L} \mid \beta_L\right) \leq U_B\left(s^*, \frac{\alpha E(\beta)\omega - s^*}{1 + \alpha E(\beta)} \mid \beta_L\right). \tag{6.31}$$

This necessary condition holds for all $s^* \in [0, \widehat{s}]$, where $\widehat{s} > 0$ is defined by identity (4.13). This identity is obtained by letting inequality (6.31) hold with equality, setting $s^* = \widehat{s}$, and using the explicit functional form given in equation (4.2).

Given the stated beliefs, the most profitable deviation for the high type is to deviate to s^{profit}, where s^{profit} maximizes $H(s)$ with respect to s, subject to $s \in [0, \omega]$ and where

$$H(s) \equiv \log(\omega - s) + \begin{cases} \beta_H \log(\omega + s) + \beta_H \log\left(\frac{\alpha\beta_L}{1+\alpha\beta_L}\right) & \text{for } s \leq \alpha\beta_L\omega \\ \beta_H \log(s) & \text{otherwise.} \end{cases} \tag{6.32}$$

(Here I implicitly assumed that $\widetilde{\mu}(s) = 0$ also for $s = s^*$. This is not a problem, however, since if $s^{profit} = s^*$, then the high type will definitly not have an incentive to deviate.) One can verify that if

$$\beta_H \log\left[\frac{\alpha\beta_L}{1+\alpha\beta_L}\right] \geq \beta_H \log\left[\frac{\beta_H}{1+\beta_H}\right] - \log(1+\beta_H), \tag{6.33}$$

then $s^{profit} = 0$; otherwise $s^{profit} = \frac{\beta_H\omega}{1+\beta_H}$. In a pooling equilibrium, the high type's payoff is

$$U_B\left(s^*, \frac{\alpha E(\beta)\omega - s^*}{1+\alpha E(\beta)} \mid \beta_H\right) = \log(\omega - s^*) + \beta_H \log(\omega + s^*) + \beta_H \log\left(\frac{\alpha E(\beta)}{1+\alpha E(\beta)}\right). \tag{6.34}$$

This payoff obtains its highest value for $s^* = 0$. Hence, pooling eqilibria exist if and only if an equilibrium with $s^* = 0$ exists. If $s^{profit} = 0$, then clearly the high type does not have an incentive to deviate from $s^* = 0$. Thus, a sufficient condition for existence of pooling equilibria is that (6.33) is met. In case that $s^{profit} = \frac{\beta_H\omega}{1+\beta_H}$, then the high type does not have an incentive to deviate from if and only if

$$\beta_H \log\left[\frac{\beta_H}{1+\beta_H}\right] - \log(1+\beta_H) \leq \beta_H \log\left(\frac{\alpha E(\beta)}{1+\alpha E(\beta)}\right). \tag{6.35}$$

Hence, pooling equilibria exist if and only if either (i) (6.33) is met or (ii) (6.33) is not met but (6.35) is. However, (condition (6.35) not met) implies (condition (6.33) not met). Moreover, if condition (6.35) is met, then either (i) or (ii) is met. Thus, pooling equilibria exist if and only if condition (6.35) is met. This condition can be rewritten as $\alpha \geq \beta_H\varphi(\beta_H)/E(\beta)$, which proves the existence claim.

Let us now turn to the characterization claim. It follows from a geometric argument (see Figure 5) that if $s^{profit} = 0$, then the high type does not have an incentive to deviate from any $s^* \in [0, \widehat{s}]$. If $s^{profit} = \frac{\beta_H \omega}{1+\beta_H}$, then the high type does not have an incentive to deviate if and only if $s^* \in \left[0, s^{\#}\right]$. To verify the latter claim, notice that

$$U_B \left(s, \frac{\alpha E(\beta) \omega - s}{1 + \alpha E(\beta)} \mid \beta_H \right) \geq U_B \left(\frac{\beta_H \omega}{1 + \beta_H}, 0 \mid \beta_H \right) \qquad (6.36)$$

holds if and only if $s \leq s^{\#}$. □

Before proving Proposition 5, I will first state and prove four lemmas.

Lemma A6. *Suppose that* $\alpha \leq (1 + \beta_H)^{-1}$ *and* $\beta_L > f(\beta_H, \alpha)$. *Then* $s^{''} < \widetilde{s}_{II}$.

Proof of Lemma A6: From Lemma A3 in this Appendix we know that if $\alpha \leq (1 + \beta_H)^{-1}$ and $\beta_L > f(\beta_H, \alpha)$, then $s^{''} < \alpha \beta_H \omega$. It thus suffices to show that if $\alpha \leq (1 + \beta_H)^{-1}$, then $\widetilde{s}_{II} \geq \alpha \beta_H \omega$. But this follows immediately since, by definition, $\widetilde{s}_{II} = \beta_H \omega (1 + \beta_H)^{-1}$. □

Lemma A7. *Suppose that* $\alpha > (1 + \beta_H)^{-1}$. *Then* $\widetilde{s}_I < \alpha \beta_H \omega$.

Proof of Lemma A7: $\widetilde{s}_I \leq \alpha \beta_H \omega$ is equivalent to

$$1 - \alpha \leq \alpha^2 \beta_H (1 + \beta_H). \qquad (6.37)$$

Note that $\alpha > (1 + \beta_H)^{-1}$ is equivalent to $\beta_H > \frac{1}{\alpha} - 1$ and that the right-hand side of (6.37) is strictly increasing in β_H. Also, evaluated at $\beta_H = \frac{1}{\alpha} - 1$, condition (6.37) holds with equality; hence it holds for all β_H satisfying $\alpha > (1 + \beta_H)^{-1}$. □

Lemma A8. *Suppose that* $\alpha \leq (1 - \beta_H)^{-1}$. *Then there exists a function* $\widetilde{\alpha}(\cdot, \beta_H) : (0, \beta_H) \to \Re$ *defined by (4.24). Furthermore, we have* $\widetilde{s}_I \geq s^{'}$ *if and only if* $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$.

Proof of Lemma A8: By using the definition of s', one can show that $\widetilde{s}_I \geq s^{'}$ is equivalent to

$$\log \left[\frac{1 + \alpha (1 + \beta_H)}{2\alpha} \right] + \beta_L \log \left[\frac{\beta_L [1 + \alpha (1 + \beta_H)]}{2\beta_H (1 + \alpha \beta_L)} \right] \geq 0. \qquad (6.38)$$

Clearly, inequality (6.38) holds if α is sufficiently close to zero. It thus suffices to show that (i) the left-hand side of inequality (6.38) is strictly decreasing in α for all $\alpha \in \left(0, (1-\beta_H)^{-1}\right]$ and that (ii) the inequality does not hold for $\alpha = (1-\beta_H)^{-1}$. To establish (i), differentiate the left-hand side of inequality (6.38) with respect to α; the resulting expression has the same sign as $(\alpha\beta_L(\beta_H - \beta_L) - 1)$.[25] This expression is strictly negative for all $\alpha \leq (1-\beta_H)^{-1}$ if

$$\left(\frac{\beta_L(\beta_H - \beta_L)}{1-\beta_H} - 1\right) < 0 \Leftrightarrow 1 - \beta_H(1-\beta_L) + \beta_L^2 > 0. \tag{6.39}$$

The last inequality is clearly true for all $\beta_H \in (\beta_L, 1)$. To establish (ii), substitute $\alpha = (1-\beta_H)^{-1}$ into the left-hand side of inequality (6.38). This yields

$$\beta_L \log\left[\frac{\beta_L(1-\beta_H)}{\beta_H(1-\beta_H+\beta_L)}\right] \geq 0, \tag{6.40}$$

or $(\beta_L - \beta_H)(1-\beta_H) \geq \beta_L\beta_H$, which clearly does not hold. □

Lemma A9. *Suppose that* $\alpha \in \left(\frac{1}{1+\beta_H}, \frac{1}{\beta_H}\right)$. *Then* $f(\beta_H, \alpha) < g(\beta_H, \alpha)$.

Proof of Lemma A9: Some straightforward algebra shows that $f(\beta_H, \alpha) < g(\beta_H, \alpha)$ is equivalent to $\eta(\beta_H, \alpha) < 1$, where

$$\eta(\beta_H, \alpha) = \frac{1+\alpha(1+\beta_H)}{2}\left(\frac{[1+\alpha(1+\beta_H)](1-\alpha\beta_H)}{2\alpha}\right)^{\frac{1}{\beta_H}}.$$

Evaluating η at $\alpha = (1+\beta_H)^{-1}$ yields $\eta\left(\beta_H, (1+\beta_H)^{-1}\right) = 1$. It thus suffices to show that η is strictly decreasing in α for $\alpha > (1+\beta_H)^{-1}$ or, equivalently, that $\log(\eta)$ is strictly decreasing in α for $\alpha > (1+\beta_H)^{-1}$. We have

$$\frac{\partial \log[\eta(\beta_H, \alpha)]}{\partial\alpha} = \frac{\alpha\beta_H(1+\beta_H)[1-\alpha(1+\beta_H)] - 1}{\alpha\beta_H(1-\alpha\beta_H)[1+\alpha(1+\beta_H)]},$$

which is strictly negative when $\alpha > (1+\beta_H)^{-1}$. □

[25] See equation (6.60).

Proof of Proposition 5: To prove the first part of the proposition we must show that if $\alpha \leq (1+\beta_H)^{-1}$, then $\left(s_L^*, t_L^*, \widetilde{s}_{II}, \widetilde{t}_{II}\right)$ can be sustained as part of a separating equilibrium iff $\beta_L \leq f(\beta_H, \alpha)$. A separating equilibrium must belong to one of the four categories (3.1)-(3.4) listed in Propostion 3. If $\alpha \leq (1+\beta_H)^{-1}$ and $\beta_L \leq f(\beta_H, \alpha)$, then the relevant category is either (3.3) or (3.4). In either case, $\widetilde{s}_{II} = \frac{\beta_H \omega}{1+\beta_H}$ belongs to the set Ω^{sep} of saving levels that can be sustained as part of a separating equilibrium. This shows that the condition $\beta_L \leq f(\beta_H, \alpha)$ is sufficient. To see that this condition also is necessary, suppose that it does not hold, i.e., $\beta_L > f(\beta_H, \alpha)$ (but we still have $\alpha \leq (1+\beta_H)^{-1}$). Then the only possible category is category (3.1), where $\Omega^{sep} = \left[s', s''\right]$. But from Lemma A6 it follows that under the stated conditions, $\widetilde{s}_{II} > s''$; hence, we cannot have $\widetilde{s}_{II} \in \Omega^{sep}$.

To prove the second part of the proposition we must show that if $\alpha > (1+\beta_H)^{-1}$, then $\left(s_L^*, t_L^*, \widetilde{s}_I, \widetilde{t}_I\right)$ can be sustained as part of a separating equilibrium iff $\beta_L \leq g(\beta_H, \alpha)$ and $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$. To start with, suppose that $\alpha \in \left((1+\beta_H)^{-1}, \beta_H^{-1}\right)$. From Lemma A7-A8, the fact that $\widetilde{s}_I \leq s'' \Leftrightarrow \beta_L \leq g(\beta_H, \alpha)$, and Proposition 3 it follows that then $\left(s_L^*, t_L^*, \widetilde{s}_I, \widetilde{t}_I\right)$ can be sustained as part of a separating equilibrium iff one of the following to (mutually exclusive) sets of inequalities holds:[26] $\beta_L > f(\beta_H, \alpha)$ and $\beta_L \leq g(\beta_H, \alpha)$ and $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$; or $\beta_L \leq f(\beta_H, \alpha)$ and $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$. But from Lemma A9 we know that in the subset of the parameter space under consideration, $f(\beta_H, \alpha) < g(\beta_H, \alpha)$. This implies that the latter set of inequalities is equivalent to: $\beta_L \leq f(\beta_H, \alpha)$ and $\beta_L \leq g(\beta_H, \alpha)$ and $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$. It is now clear that one of the set of inequalities will hold true iff: $\beta_L \leq g(\beta_H, \alpha)$ and $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$. It remains to consider the case $\alpha \in \left[\beta_H^{-1}, (1-\beta_H)^{-1}\right)$. One can see that now the only relevant category in Proposition 3 is (3.1). From Lemma A8, the fact that $\widetilde{s}_I \leq s'' \Leftrightarrow \beta_L \leq g(\beta_H, \alpha)$, and Proposition 3 it follows that $\left(s_L^*, t_L^*, \widetilde{s}_I, \widetilde{t}_I\right)$ can be sustained as part of a separating equilibrium iff: $\beta_L \leq g(\beta_H, \alpha)$ and $\alpha \leq \widetilde{\alpha}(\beta_L, \beta_H)$. □

Lemma A10. *$f^{-1}(\beta_L)$ is well-defined and* $\frac{\partial f^{-1}(\beta_L)}{\partial \beta_L} < 0$. Moreover,

$$\lim_{\beta_L \to 0} f^{-1}(\beta_L) = \beta_H^{-1}, \lim_{\beta_L \to \beta_H} f^{-1}(\beta_L) = 0 \text{ and } f^{-1}(\beta_H(1+\beta_H)\varphi(\beta_H)) = (1+\beta_H)^{-1}. \tag{6.41}$$

[26] The first set of inequalities refer to (3.1) in Proposition 3 and the second to (3.2).

Proof of Lemma A10: The function $f^{-1}(\beta_L)$ is well-defined if f is strictly decreasing in α, which it turns out to be. To see this, differentiate the denominator of f (see equation (4.10)) with respect to α:

$$\frac{\partial\left[\frac{1}{\beta_H}(1-\alpha\beta_H)^{-\frac{1}{\beta_H}}-\alpha\right]}{\partial\alpha}=\frac{1}{\beta_H}(1-\alpha\beta_H)^{-\frac{1+\beta_H}{\beta_H}}-1. \tag{6.42}$$

This expression is strictly greater than zero since it is increasing in α and it is strictly greater than zero for $\alpha=0$. Hence, $\frac{\partial f(\beta_H,\alpha)}{\partial\alpha}<0$. To demonstrate the claims in (6.41), it suffices to show that $\lim_{\alpha\to\beta_H^{-1}} f(\beta_H,\alpha)=0$, $\lim_{\alpha\to 0} f(\beta_H,\alpha)=\beta_H$, and $f\left((1+\beta_H)^{-1}\right)=\beta_H(1+\beta_H)\varphi(\beta_H)$. However, it can be easily verified that these three equalities hold true by using the definition of f in (4.10).

Lemma A11. $\widetilde{\alpha}(\beta_L,\beta_H)>1$ *and*

$$\lim_{\beta_L\to 0}\widetilde{\alpha}(\beta_L,\beta_H)=\lim_{\beta_L\to\beta_H}\widetilde{\alpha}(\beta_L,\beta_H)=(1-\beta_H)^{-1}. \tag{6.43}$$

Moreover, for β_L sufficiently close to β_H, $\widetilde{\alpha}(\beta_L,\beta_H)<(1-\beta_L)^{-1}$.

Proof of Lemma A11: It is first shown that $\widetilde{\alpha}(\beta_L,\beta_H)>1$. Substitute $\alpha=1$ into the left-hand side of inequality (6.38). This yields

$$\log\left[\frac{2+\beta_H}{2}\right]+\beta_L\log\left[\frac{\beta_L(2+\beta_H)}{2\beta_H(1+\beta_L)}\right]\equiv z(\beta_L,\beta_H). \tag{6.44}$$

From Lemma A8 it follows that $\widetilde{\alpha}(\beta_L,\beta_H)>1$ if $z(\beta_L,\beta_H)\geq 0$. In order to show that the latter inequality indeed holds true, suppose per contra that, for some (β_L,β_H), $z(\beta_L,\beta_H)<0$. If $z(\beta_L,\beta_H)<0$ for some (β_L,β_H), then this inequality must hold for some β_L at one of the corners with respect to β_H of the function z or at a value of β_H where the partial derivative of z with respect to β_H is zero. That is, at least one of the following three inequalities must hold for some $\beta_L\in(0,1)$:

$$z(\beta_L,\beta_L)=\log\left[\frac{2+\beta_L}{2}\right]+\beta_L\log\left[\frac{2+\beta_L}{2(1+\beta_L)}\right]\equiv\underline{z}(\beta_L)<0, \tag{6.45}$$

$$z(\beta_L,1)=\log\left(\frac{3}{2}\right)+\beta_L\log\left[\frac{3\beta_L}{2(1+\beta_L)}\right]\equiv\overline{z}(\beta_L)<0, \tag{6.46}$$

$$z\left(\beta_L,\widehat{\beta}_H\right)=\log\left[\frac{2+\beta_H}{2}\right]+\beta_L\log\left[\frac{\beta_L\left(2+\beta_H\right)}{2\beta_H\left(1+\beta_L\right)}\right]<0, \tag{6.47}$$

where $\widehat{\beta}_H$ is implicitly defined by $\partial z\left(\beta_L,\widehat{\beta}_H\right)/\partial\beta_H=0$. I will show that none of these three inequalities hold. To start with, consider the first inequality, $\underline{z}\left(\beta_L\right)<0$. One can easily verify that $\underline{z}'\left(0\right)=\frac{1}{2}>0$, $\underline{z}'\left(1\right)=\frac{1}{6}+\log\left(\frac{3}{4}\right)<0$, and $\underline{z}''\left(\beta_L\right)<0$. Moreover, $\underline{z}\left(0\right)=0$ and $\underline{z}\left(1\right)=\log\left(\frac{9}{8}\right)>0$. Hence, $\underline{z}\left(\beta_L\right)>0$ for all $\beta_L\in\left(0,1\right)$.

Next, consider the second inequality, $\overline{z}\left(\beta_L\right)<0$. One can easily verify that $\overline{z}'\left(0\right)=-\infty$, $\overline{z}'\left(1\right)=\frac{1}{2}+\log\left(\frac{3}{4}\right)>0$, and $\overline{z}''\left(\beta_L\right)>0$. Hence, if $\overline{z}\left(\beta_L\right)<0$ for some $\beta_L\in\left(0,1\right)$, then, in particular, it must be that $\overline{z}\left(\beta'\right)<0$ where β' is implicitly defined by

$$\overline{z}'\left(\beta'\right)=0\Leftrightarrow\log\left[\frac{3\beta'}{2\left(1+\beta'\right)}\right]+\frac{1}{1+\beta'}=0. \tag{6.48}$$

Using (6.48), $\overline{z}\left(\beta'\right)$ can be rewritten as follows:

$$\begin{aligned}\overline{z}\left(\beta'\right)&=\overline{z}\left(\beta'\right)-\left(1+\beta'\right)\log\left[\frac{3\beta'}{2\left(1+\beta'\right)}\right]-1\\&=\log\left(\frac{3}{2}\right)+\beta'\log\left[\frac{3\beta'}{2\left(1+\beta'\right)}\right]-\left(1+\beta'\right)\log\left[\frac{3\beta'}{2\left(1+\beta'\right)}\right]-1\\&=\log\left[\frac{\left(1+\beta'\right)}{\beta'}\right]-1.\end{aligned}$$

Hence, $\overline{z}\left(\beta'\right)<0$ is equivalent to:

$$\beta'>\left(e-1\right)^{-1}. \tag{6.49}$$

However, by (6.49), the convexity of $\overline{z}$, and the definition of β', we must have $\overline{z}'\left(\left(e-1\right)^{-1}\right)=\log\left(\frac{3}{2}\right)-\frac{1}{e}<0$, which is not true. Hence, we must have $\overline{z}\left(\beta_L\right)\geq0$ for all $\beta_L\in\left(0,1\right)$.

Finally consider the inequality $z\left(\beta_L,\widehat{\beta}_H\right)<0$. It is readily verified that $\widehat{\beta}_H=2\beta_L$. Hence

$$z\left(\beta_L,\widehat{\beta}_H\right)=z\left(\beta_L,2\beta_L\right)=\log\left(1+\beta_L\right)-\beta_L\log\left(2\right). \tag{6.50}$$

This expression is strictly convex and has a global minimum at $\beta_L = [1 - \log(2)] / \log(2)$. But

$$\log\left(1 + \frac{1 - \log(2)}{\log(2)}\right) - \left[\frac{1 - \log(2)}{\log(2)}\right] \log(2) = -\log(\log(2)) - 1 + \log(2) > 0. \tag{6.51}$$

Hence, we must have $z\left(\beta_L, \widehat{\beta}_H\right) > 0$ for all $\beta_L \in (0, 1)$.

Let us now show that $\lim_{\beta_L \to \beta_H} \widetilde{\alpha}(\beta_L, \beta_H) = (1 - \beta_H)^{-1}$. We know from Lemma A8 that $\widetilde{\alpha}(\beta_L, \beta_H) \leq (1 - \beta_H)^{-1}$ for all $\beta_L \in (0, \beta_H)$; hence, $\lim_{\beta_L \to \beta_H} \widetilde{\alpha}(\beta_L, \beta_H)$ must be finite. Let us use the notation $\lim_{\beta_L \to \beta_H} \widetilde{\alpha}(\beta_L, \beta_H) = \widehat{\alpha}$. First take the left-hand limit of both sides of the identity (4.24):

$$\lim_{\beta_L \to \beta_H} \left\{\log\left[\frac{1 + \alpha(1 + \beta_H)}{2\alpha}\right]\right\} + \lim_{\beta_L \to \beta_H} \left\{\beta_L \log\left[\frac{\beta_L [1 + \alpha(1 + \beta_H)]}{2\beta_H (1 + \alpha\beta_L)}\right]\right\} \equiv 0, \tag{6.52}$$

or

$$\log\left[\frac{1 + \widehat{\alpha}(1 + \beta_H)}{2\widehat{\alpha}}\right] + \beta_H \log\left[\frac{1 + \widehat{\alpha}(1 + \beta_H)}{2(1 + \widehat{\alpha}\beta_H)}\right] = 0. \tag{6.53}$$

It is easily verified that equality (6.53) is met for $\widehat{\alpha} = (1 - \beta_H)^{-1}$. Moreover, the equality cannot be met for any other value of $\widehat{\alpha}$, since the left-hand side of equality (6.53) is strictly decreasing in $\widehat{\alpha}$ (this is left to the reader to verify).

Let us now show that $\lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H) = (1 - \beta_H)^{-1}$. First take the limit of both sides of the identity (4.24):

$$\lim_{\beta_L \to 0} \left\{\log\left[\frac{1 + \alpha(1 + \beta_H)}{2\alpha}\right]\right\} + \lim_{\beta_L \to 0} \left\{\beta_L \log\left[\frac{\beta_L [1 + \alpha(1 + \beta_H)]}{2\beta_H (1 + \alpha\beta_L)}\right]\right\} \equiv 0, \tag{6.54}$$

or

$$\begin{aligned} & \log\left[\frac{1 + (1 + \beta_H) \lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)}{2 \lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)}\right] + \left(\lim_{\beta_L \to 0} \beta_L\right) \log\left[\frac{\lim_{\beta_L \to 0} \beta_L}{2\beta_H}\right] \\ & + \left(\lim_{\beta_L \to 0} \beta_L\right) \log\left[\frac{1 + (1 + \beta_H)\left(\lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)\right)}{1 + \left(\lim_{\beta_L \to 0} \beta_L\right)\left(\lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)\right)}\right] \\ = \; & 0. \end{aligned} \tag{6.55}$$

or, since $\lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)$ is finite,

$$\log\left[\frac{1 + (1 + \beta_H) \lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)}{2 \lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H)}\right] = 0 \tag{6.56}$$

or $\lim_{\beta_L \to 0} \widetilde{\alpha}(\beta_L, \beta_H) = (1-\beta_H)^{-1}$.

Let us finally show that for β_L sufficiently close to β_H, $\widetilde{\alpha}(\beta_L, \beta_H) < (1-\beta_L)^{-1}$. Since $\lim_{\beta_L \to \beta_H} \widetilde{\alpha}(\beta_L, \beta_H) = (1-\beta_H)^{-1}$, it suffices to show that

$$\lim_{\beta_L \to \beta_H} \frac{\partial \widetilde{\alpha}(\beta_L, \beta_H)}{\partial \beta_L} < \lim_{\beta_L \to \beta_H} \frac{\partial (1-\beta_L)^{-1}}{\partial \beta_L} = (1-\beta_H)^{-2}. \tag{6.57}$$

Let us first calculate the derivative on the left-hand side of (6.57):

$$\frac{\partial \widetilde{\alpha}(\beta_L, \beta_H)}{\partial \beta_L} = -\frac{\partial \left[LHS\,(6.38)\right]/\partial \beta_L}{\partial \left[LHS\,(6.38)\right]/\partial \alpha}. \tag{6.58}$$

Straightforward calculations shows that

$$\frac{\partial \left[LHS\,(6.38)\right]}{\partial \beta_L} = \log\left[\frac{\beta_L\left[1+\alpha(1+\beta_H)\right]}{2\beta_H(1+\alpha\beta_L)}\right] + 1 \tag{6.59}$$

and

$$\frac{\partial \left[LHS\,(6.38)\right]}{\partial \alpha} = -\frac{1-\alpha\beta_L(\beta_H-\beta_L)}{\alpha(1+\alpha\beta_L)\left[1+\alpha(1+\beta_H)\right]}. \tag{6.60}$$

Taking the limits of (6.59) and (6.60) yields:

$$\lim_{\beta_L \to \beta_H} \frac{\partial \left[LHS\,(6.38)\right]}{\partial \beta_L} = \log\left[\frac{(2-\beta_H)(1+\beta_H)}{2}\right] + 1 \tag{6.61}$$

$$\lim_{\beta_L \to \beta_H} \frac{\partial \left[LHS\,(6.38)\right]}{\partial \alpha} = \frac{2}{(1-\beta_H)^2}. \tag{6.62}$$

Hence,

$$\lim_{\beta_L \to \beta_H} \frac{\partial \widetilde{\alpha}(\beta_L, \beta_H)}{\partial \beta_L} = \frac{2\log\left[\frac{(2-\beta_H)(1+\beta_H)}{2}\right] + 2}{(1-\beta_H)^2}. \tag{6.63}$$

This means that

$$\lim_{\beta_L \to \beta_H} \frac{\partial \widetilde{\alpha}(\beta_L, \beta_H)}{\partial \beta_L} < (1-\beta_H)^{-2} \Leftrightarrow 2\log\left[\frac{(2-\beta_H)(1+\beta_H)}{2}\right] < 1, \tag{6.64}$$

which always holds since the left-hand side of (6.64) obtains its maximum at $\beta_H = \frac{1}{2}$ and the inequality does not hold for this value of β_H. □

Lemma A12. *$g^{-1}(\beta_L)$ is well-defined and $\frac{\partial g^{-1}(\beta_L)}{\partial \beta_L} > 0$.* Moreover,

$$\lim_{\beta_L \to 0} g^{-1}(\beta_L) = 0, \ \lim_{\beta_L \to \beta_H} g^{-1}(\beta_L) = (1-\beta_H)^{-1} \tag{6.65}$$

and

$$g^{-1}(\beta_H(1+\beta_H)\varphi(\beta_H)) = (1+\beta_H)^{-1}. \tag{6.66}$$

Proof of Lemma A12: The function $g^{-1}(\beta_L)$ is well-defined if g is strictly increasing in α, which it turns out to be. To see this, differentiate the denominator of g (see equation (4.23)) with respect to α:

$$\frac{\partial\left[\frac{1+\alpha(1+\beta_H)}{2\beta_H}\left(\frac{1+\alpha(1+\beta_H)}{2\alpha}\right)^{\frac{1}{\beta_H}}-\alpha\right]}{\partial\alpha}$$
$$= -\frac{1}{2\alpha\beta_H^2}\left(\frac{\frac{1}{\alpha}+1+\beta_H}{2}\right)^{\frac{1}{\beta_H}}\left[1-\alpha\beta_H(1+\beta_H)\right]-1\equiv\widetilde{g}(\beta_H,\alpha)\,. \quad (6.67)$$

The expression on the right-hand side of equation (6.67), $\widetilde{g}(\beta_H,\alpha)$, is clearly increasing in α. Moreover, evaluated at $\alpha=(1-\beta_H)^{-1}$, $\widetilde{g}(\beta_H,\alpha)$ is strictly negative:

$$\widetilde{g}\left(\beta_H,(1-\beta_H)^{-1}\right)=-\frac{(1-\beta_H)}{2\beta_H^2}\left[\frac{1-\beta_H-\beta_H(1+\beta_H)}{1-\beta_H}\right]-1<0\Leftrightarrow\beta_H(2+\beta_H)<1. \quad (6.68)$$

(The last inequality holds for all $\beta_H\in(0,1)$ since $\beta_H(2+\beta_H)$ is strictly concave and reaches its maximum at $\beta_H=1$.) Hence, we must have $\widetilde{g}(\beta_H,\alpha)<0$ for all $\alpha<(1-\beta_H)^{-1}$. It follows that g is strictly increasing in α for all $\alpha<(1-\beta_H)^{-1}$. To demonstrate the claims in (6.65) and (6.66), it suffices to show that $\lim_{\alpha\to0}g(\beta_H,\alpha)=0$, $\lim_{\alpha\to(1-\beta_H)^{-1}}g(\beta_H,\alpha)=\beta_H$, and $g\left((1+\beta_H)^{-1}\right)=\beta_H(1+\beta_H)\varphi(\beta_H)$. However, it can be easily verified that these three equalities hold true by using the definition of g in (4.23).

Proof of Proposition 6: It remains to show that the intuitive criterion rules out all pooling equilibria. Consider a pooling equilibrium where $s^*\in\Omega^{pool}=\left[0,\min\left\{\widehat{s},s^{\#\#}\right\}\right]$. Refer to Figure 5 and imagine an indifference curve for the low type that crosses the intermediate straight line at $s=s^*$; denote this indifference curve by I_L^*. By Lemma A2, I_L^* crosses the upper straight line at some $s\equiv y<\alpha\beta_H\omega$. Now imagine an indifference curve for the high type that also crosses the intermediate straight line at $s=s^*$; denote this by I_H^*. By the results in note 10 (in particular the single-crossing property), I_H^* lies below I_L^* for all $s>s^*$. Moreover, I_H^* crosses the upper straight line at some $s\equiv y'<\alpha\beta_H\omega$; by the single-crossing property, $y'>y$. Clearly, for any beliefs $\mu\left(y''\right)$, the low type prefers his equilibrium payoff to the payoff he would get if deviating to some

$s = y'' \in (y, y')$. The high type, on the other hand, would have an incentive to deviate to y'' if he thereby were perceived as the high type. Hence, s^* can be part of an intuitive equilibrium only if it can be supported by beliefs assigning zero probability to the low type plaing $s = y''$, that is, $\mu(y'') = 1$. However, for these beliefs, the high type will not have an incentive to play $s = s^*$. $\square$

7. References

Becker, G. S., (1974), A theory of social interactions, Journal of Political Economy 82, 1063-1093.

Becker, G. S., and K.M. Murphy, (1988), The family and the state, Journal of Law and Economics 31, 1-18.

Bergstrom, T.C., (1989), A fresh look at the Rotten Kid Theorem—and other household mysteries., Journal of Political Economy 97, 1138-1159.

Bernheim, B. D., and O. Stark, (1988), Altruism within the family reconsidered: Do nice guys finish last?, American Economic Review 78, 1034-1045.

Bruce, N., and M. Waldman (1990), The Rotten-Kid Theorem meets the Samaritan's dilemma, Quarterly Journal of Economics 105, 155-165.

Bruce, N., and M. Waldman (1991), Transfers in kind: Why they can be efficient and nonpaternalistic, American Economic Review 81, 1345-1351.

Buchanan, J. M. (1975), The Samaritan's dilemma, in *Altruism, morality and economic theory*, in E. S. Phelps ed. (New York: Russel Sage Foundation), pp. 71-85.

Chakrabarti, S., W. Lord, and P. Rangazas (1993), Uncertain Altruism and Investment in Children, American Economic Review 83, 994-1002.

Cho, I.-K. and D. Kreps (1987), Signaling games and stable equilibria, Quarterly

Journal of Economics 102, 179-221.

Coate, S. (1995), Altruism, the Samaritan's dilemma, and government transfer policy, American Economic Review 85, 46-57.

Hansson, B., and C. Stuart (1991), Social Security as Trade Among Living Generations, American Economic Review 79, 1182-1195.

Holmström, B. and R. B. Myerson (1983), Efficient and durable decision rules with incomplete information, Econometrica 51, 1799-1819.

Kaufman, D. A., (1995), Consumption Externalities, Preference Changes and the 'Samaritan's Dilemma', Annals of Public and Cooperative Economics 66, 101-125.

Kotlikoff, L. J. (1987), Justifying public provision of social security, Journal of Policy Analysis and Management 6, 674-689.

Kreps, D.M. and J. Sobel (1994), Signalling, in "Handbook of Game Theory, Volume 2", R.J. Aumann and S. Hart, eds. (Amsterdam: North-Holland), pp. 849-867.

Lindbeck, A., and J. W. Weibull (1988), Altruism and time inconsistency: The economics of fait accompli, Journal of Political Economy 96, 1165-1182.

Lord, W., and P. Rangazas, (1995), Uncertainty, Altruism, and Savings: Precautionary Savings Meets the Samaritan's Dilemma, Public Finance / Finances Publiques 50, 404-419.

O'Connell, S.A., and S.P. Zeldes (1993), Dynamic efficiency in the gifts economy, Journal of Monetary Economics 31, 363-379.

Spence, M. (1973), Job market signaling, Quarterly Journal of Economics 87, 355-374.

Spence, M. (1974), *Market signaling: Informational transfers in hiring and related screening processes* (Cambridge: Harvard University Press).

Thompson, E.A. (1980), Charity and nonprofit organizations, in "Economics of Nonproprietory Organizations", K. Clarkson and D. Martin, eds. (Greenwich, CT: JAI Press, Inc.), pp. 125-38.

Veall, M.R. (1986), Public pensions as optimal social contracts, Journal of Public Economics 31, 237-251.

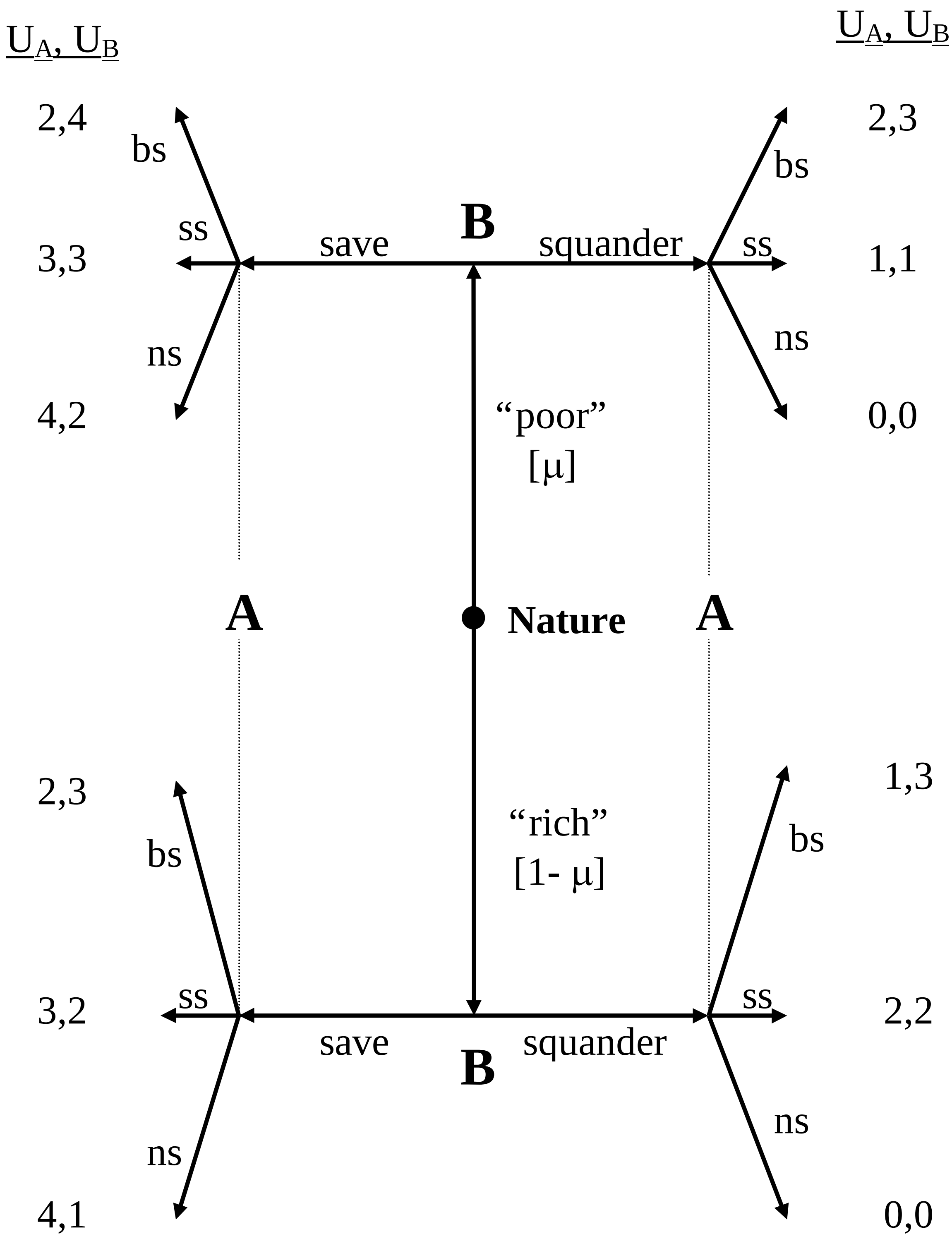

Figure 1 An example.

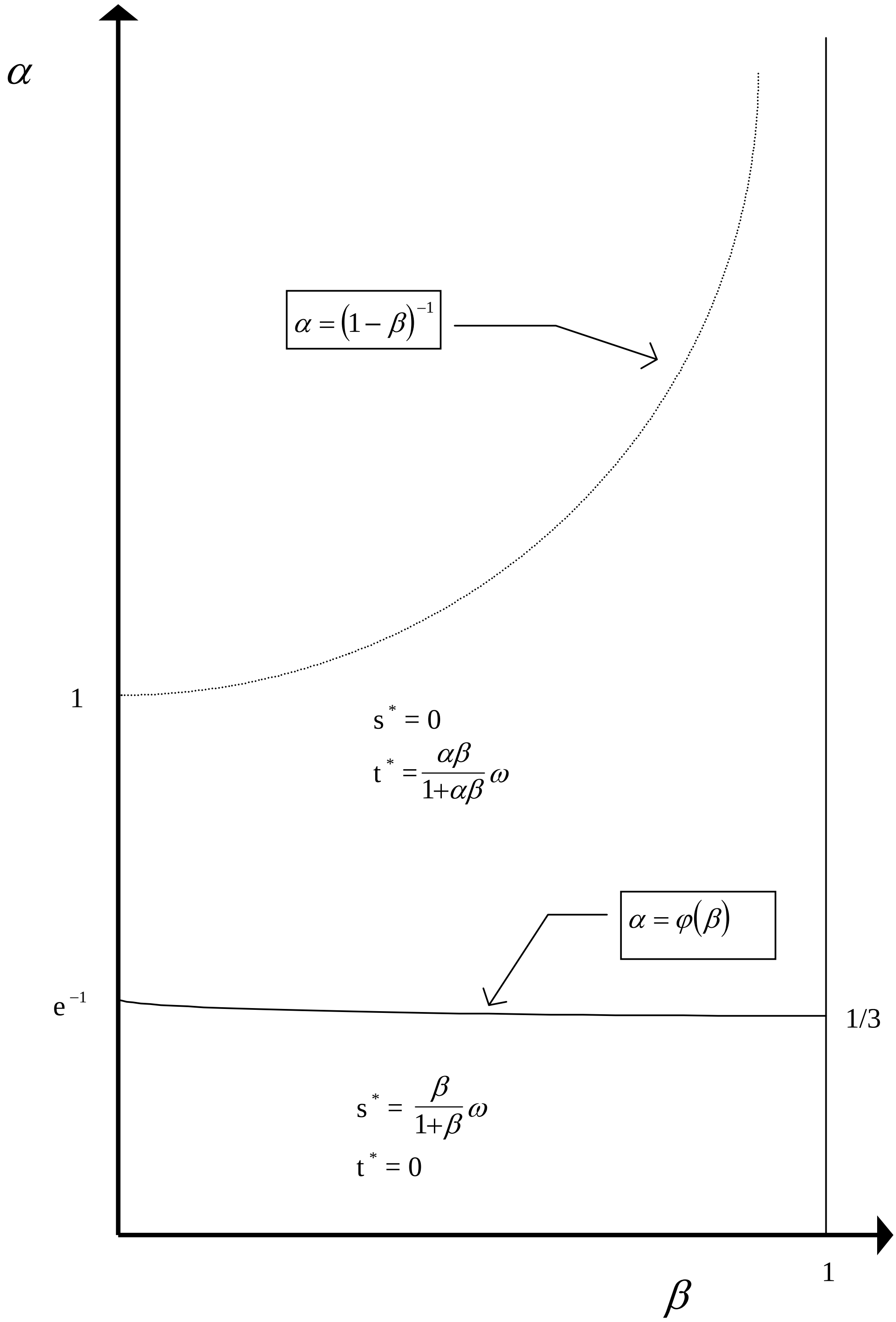

Figure 2. The Benchmark.

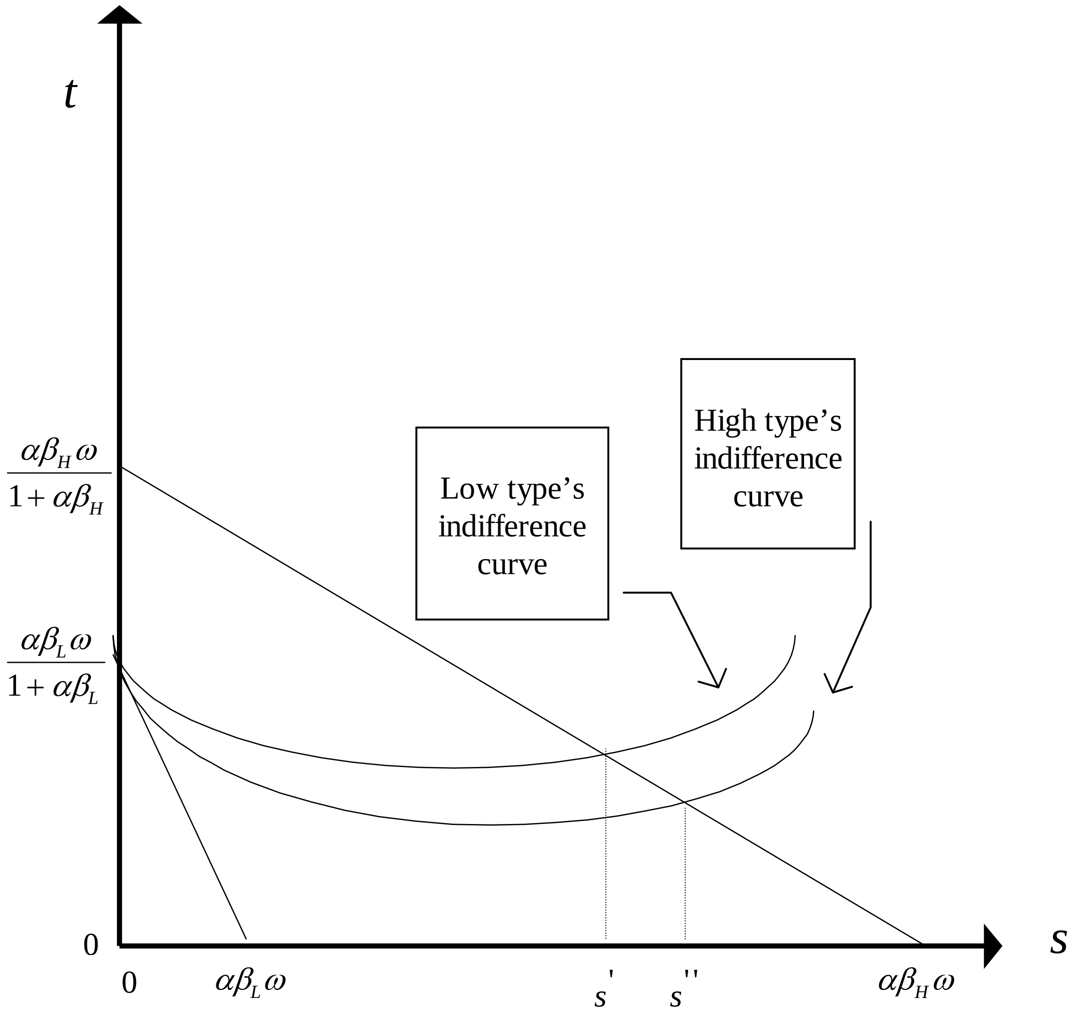

Figure 3a. Separating equilibria.
The case $s'' \leq \alpha\beta\omega$.

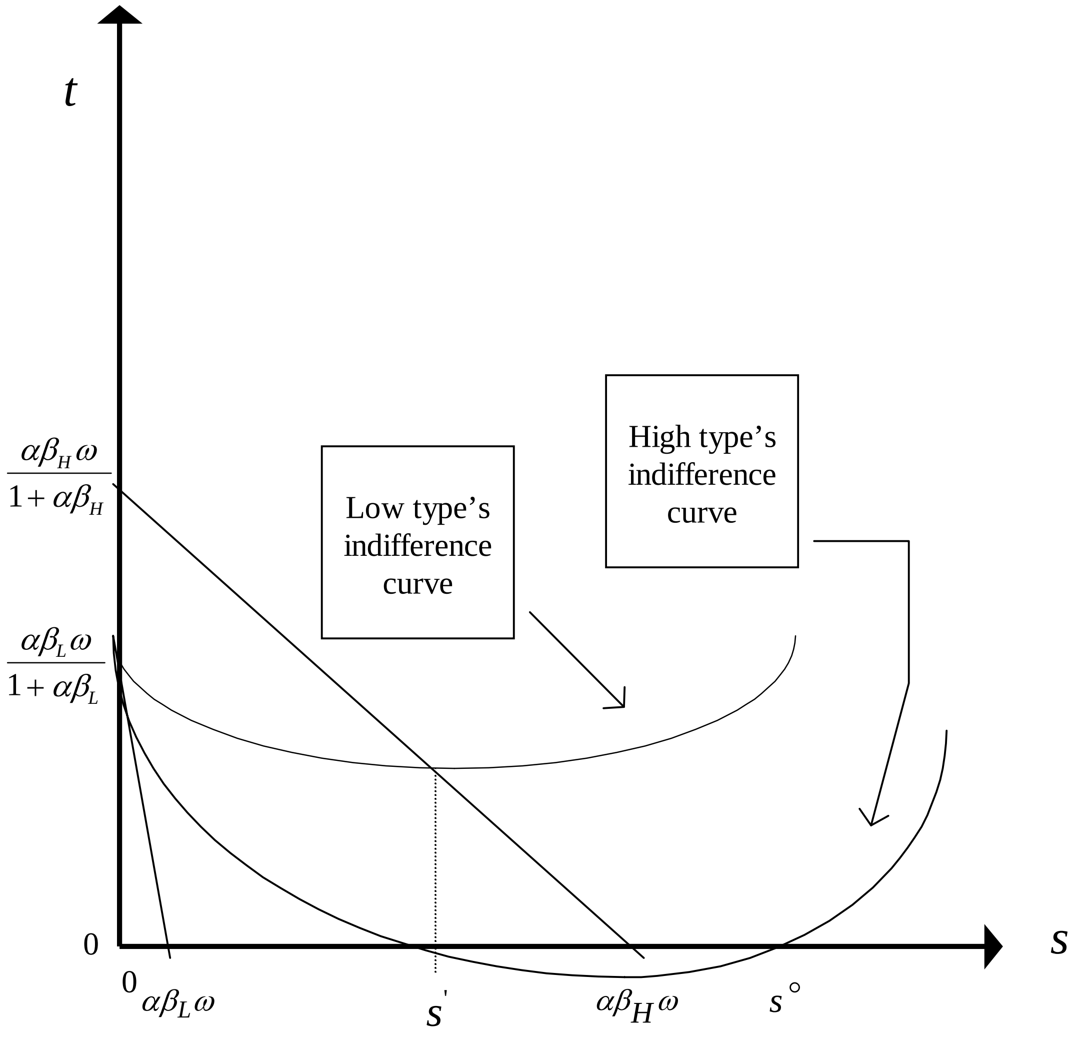

Figure 3b. Separating equilibria.
The case $s'' > \alpha\beta\omega$.

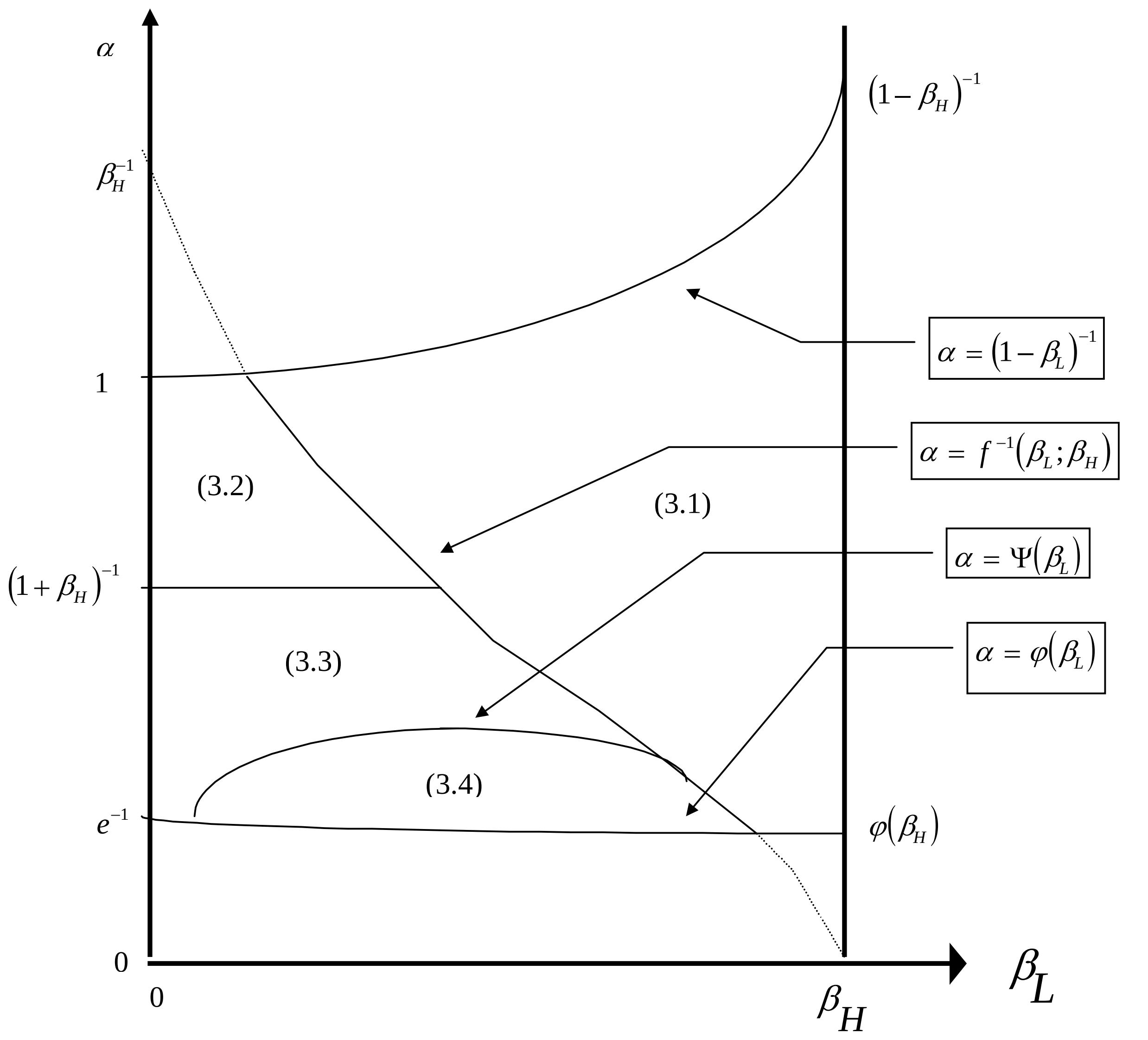

Figure 4 Existence of separating equilibria

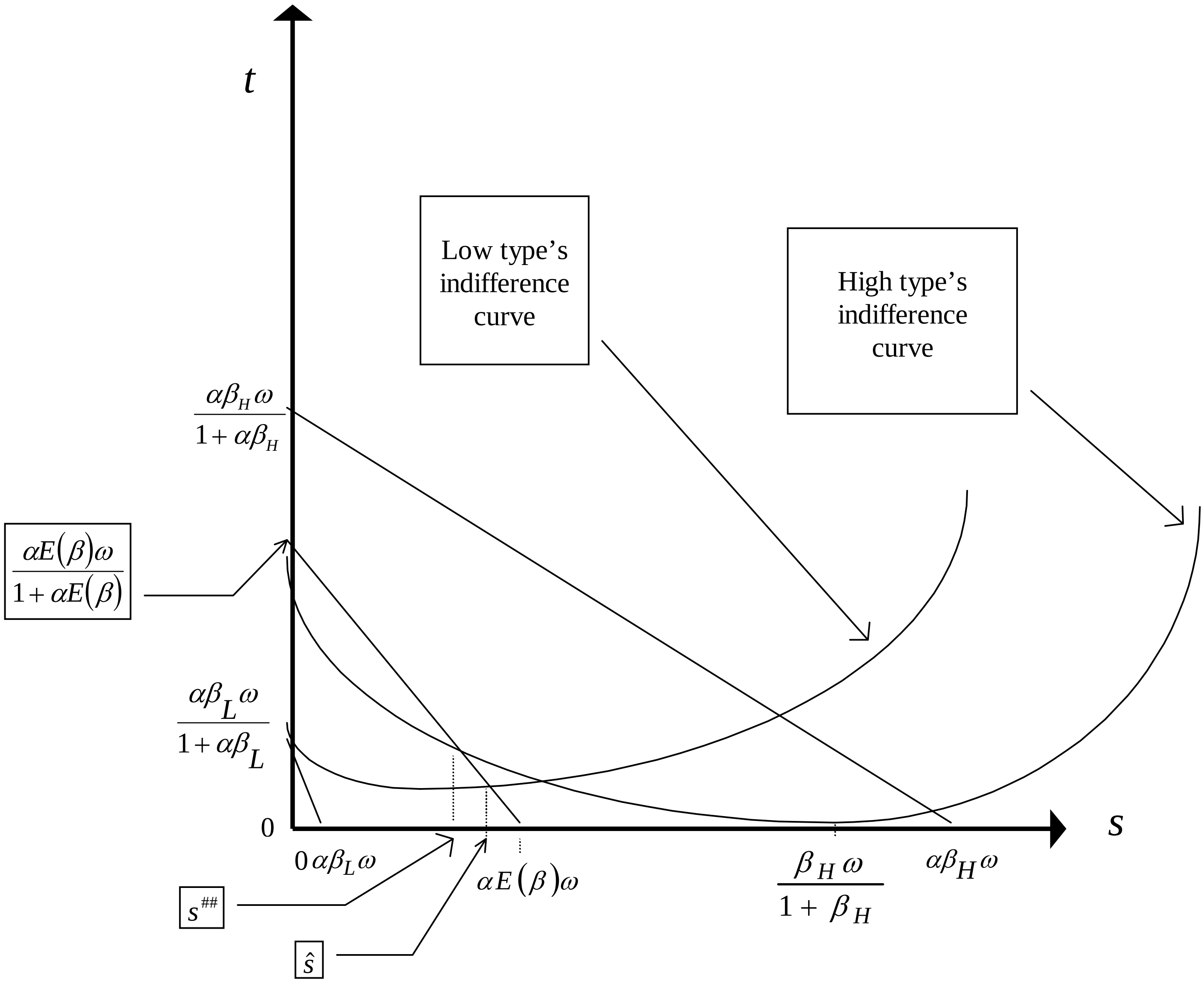

Figure 5. Pooling equilibria.

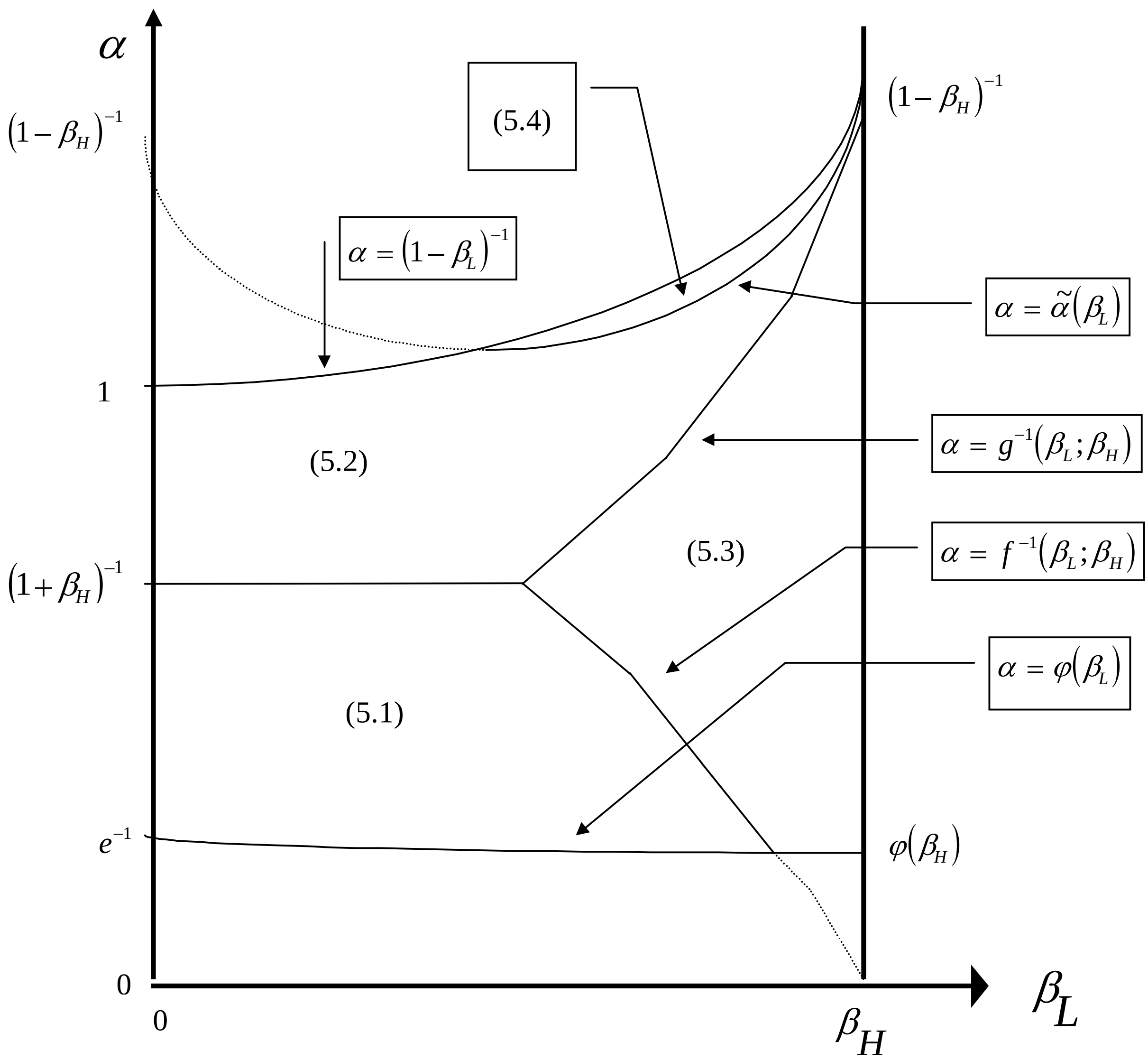

Figure 6. Efficiency.

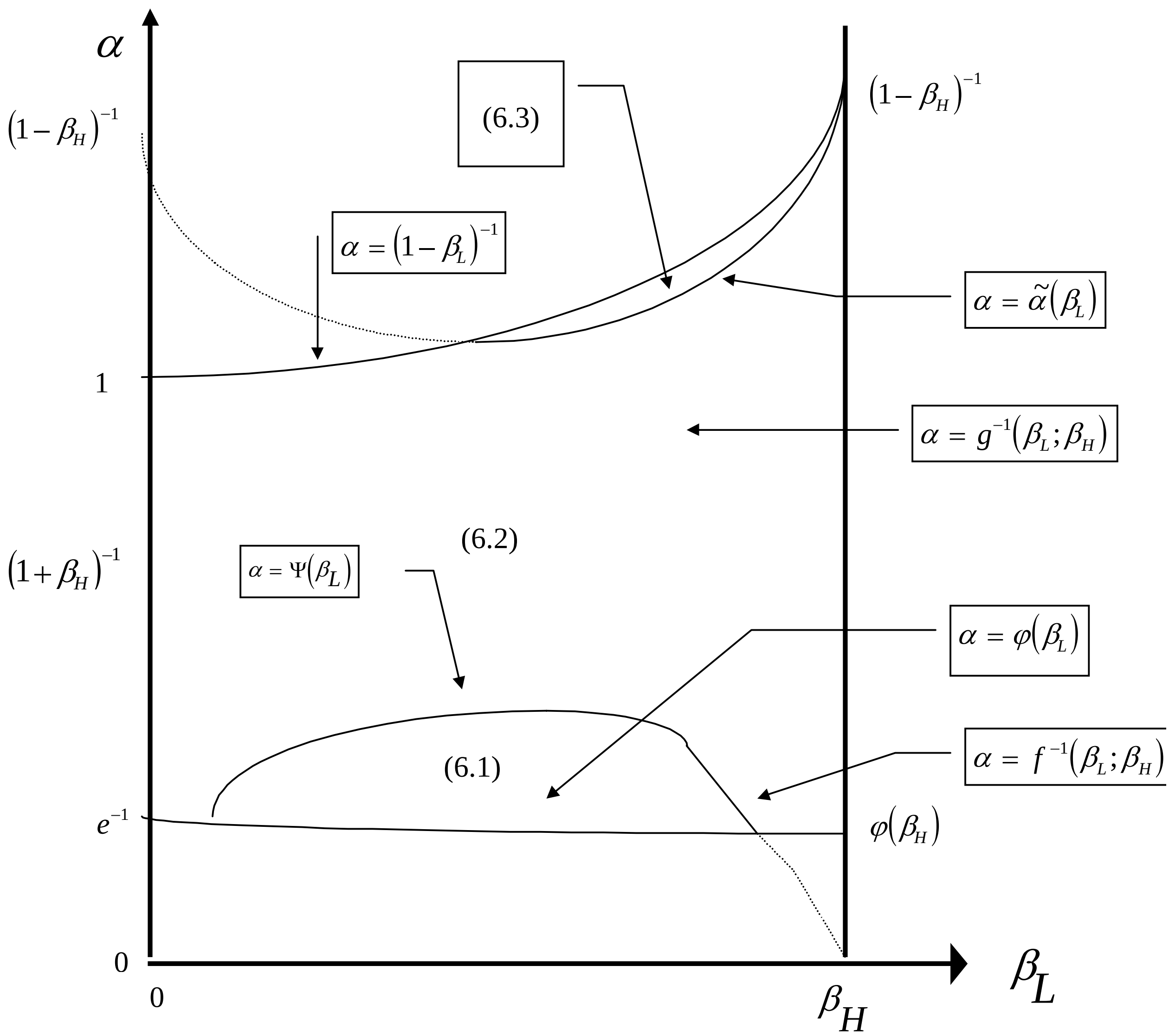

Figure 7. Efficiency with an equilibrium refinement.

Bücher des Forschungsschwerpunkts Marktprozeß und Unternehmensentwicklung

Books of the Research Area Market Processes and Corporate Development

(nur im Buchhandel erhältlich/available through bookstores)

Horst Albach, Ulrike Görtzen, Rita Zobel (Hg.)
Information Processing as a Competitive Advantage of Japanese Firms
1999, edition sigma

Dieter Köster
Wettbewerb in Netzproduktmärkten
1999, Deutscher Universitäts-Verlag/Gabler Verlag

Christian Wey
Marktorganisation durch Standardisierung: Ein Beitrag zur Neuen Institutionenökonomik des Marktes
1999, edition sigma

Horst Albach, Meinolf Dierkes, Ariane Berthoin Antal, Kristina Vaillant (Hg.)
Organisationslernen – institutionelle und kulturelle Dimensionen
1998, edition sigma

Lars Bergman, Chris Doyle, Jordi Gual, Lars Hultkrantz, Damien Neven, Lars-Hendrik Röller, Leonard Waverman
Europe's Network Industries: Conflicting Priorities - Telecommunications
Monitoring European Deregulation 1
1998, Centre for Economic Policy Research

Manfred Fleischer
The Inefficiency Trap
Strategy Failure in the
German Machine Tool Industry
1997, edition sigma

Christian Göseke
Information Gathering and Dissemination
The Contribution of JETRO to
Japanese Competitiveness
1997, Deutscher Universitäts-Verlag

Andreas Schmidt
Flugzeughersteller zwischen globalem Wettbewerb und internationaler Kooperation Der Einfluß von Organisationsstrukturen auf die Wettbewerbsfähigkeit von Hochtechnologie-Unternehmen
1997, edition sigma

Horst Albach, Jim Y. Jin, Christoph Schenk (eds.)
Collusion through Information Sharing? New Trends in Competition Policy
1996, edition sigma

Stefan O. Georg
Die Leistungsfähigkeit japanischer Banken Eine Strukturanalyse des Bankensystems in Japan
1996, edition sigma

Stephanie Rosenkranz
Cooperation for Product Innovation
1996, edition sigma

Horst Albach, Stephanie Rosenkranz (eds.)
Intellectual Property Rights and Global Competition - Towards a New Synthesis
1995, edition sigma.

David B. Audretsch
Innovation and Industry Evolution
1995, The MIT Press.

Julie Ann Elston
US Tax Reform and Investment: Reality and Rhetoric in the 1980s
1995, Avebury

Horst Albach
The Transformation of Firms and Markets: A Network Approach to Economic Transformation Processes in East Germany
Acta Universitatis Upsaliensis, Studia Oeconomiae Negotiorum, Vol. 34
1994, Almqvist & Wiksell International (Stockholm).

Horst Albach
"Culture and Technical Innovation: A Cross-Cultural Analysis and Policy Recommendations"
Akademie der Wissenschaften zu Berlin (Hg.)
Forschungsbericht 9, S. 1-597
1994, Walter de Gruyter.

Horst Albach
Zerissene Netze. Eine Netzwerkanalyse des ostdeutschen Transformationsprozesses
1993, edition sigma.

Zoltan J. Acs/David B. Audretsch (eds)
Small Firms and Entrepreneurship: An East-West Perspective
1993, Cambridge University Press.

Anette Boom
Nationale Regulierungen bei internationalen Pharma-Unternehmen: Eine theoretische Analyse der Marktwirkungen
1993, Nomos Verlagsgesellschaft.

DISCUSSION PAPERS 1998

Horst Albach	**Unternehmensgründungen in Deutschland Potentiale und Lücken**	FS IV 98 - 1
Dietmar Harhoff	**Vertical Organization, Technology Flows and R&D Incentives - An Exploratory Analysis**	FS IV 98 - 2
Karel Cool Lars-Hendrik Röller Benoit Leleux	**Der Einfluß des tatsächlichen und des potentiellen Wettbewerbs auf die Rentabilität von Unternehmen der pharmazeutischen Industrie**	FS IV 98 - 3
Horst Albach	**Blühende Landschaften? Ein Beitrag zur Transformationsforschung**	FS IV 98 - 4
Shiho Futagami Tomoki Waragai Thomas Westphal	***Shukko* in Japanese Companies and its Economic and Managerial Effects**	FS IV 98 - 5
Dietmar Harhoff Timm Körting	**Lending Relationships in Germany: Empricial Results from Survey Data**	FS IV 98 - 6
Johan Lagerlöf	**Are We Better Off if Our Politicians Know How the Economy Works?**	FS IV 98 - 7
Justus Haucap Christian Wey Jens Barmbold	**Location Costs, Product Quality, and Implicit Franchise Contracts**	FS IV 98 - 8
Manfred Fleischer	**Patenting and Industrial Performance: The Case of the Machine Tool Industry**	FS IV 98 - 9
Dieter Köster	**Was sind Netzprodukte? - Eigenschaften, Definition und Systematisierung von Netzprodukten**	FS IV 98 - 10
Andreas Blume	**Coordination and Learning with a Partial Language**	FS IV 98 - 11
Andreas Blume Uri Gneezy	**An Experimental Investigation of Optimal Learning in Coordination Games**	FS IV 98 - 12
Andreas Blume Douglas V. DeJong George R. Neumann Nathan E. Savin	**Learning in Sender-Receiver Games**	FS IV 98 - 13
Hans Mewis	**The Stability of Information Cascades: How Herd Behavior Breaks Down**	FS IV 98 - 14
Lars-Hendrik Röller Mihkel M. Tombak Ralph Siebert	**The Incentives to Form Research Joint Ventures: Theory and Evidence**	FS IV 98 - 15
Christine Zulehner	**Econometric Analysis of Cattle Auctions**	FS IV 98 - 16
Catherine Matraves	**Market Structure, R&D and Advertising in the Pharmaceutical Industry**	FS IV 98 - 17

DISCUSSION PAPERS 1999

Suchan Chae Paul Heidhues	**Bargaining Power of a Coalition in Parallel Bargaining: Advantage of Multiple Cable System Operators**	FS IV 99 - 1
Christian Wey	**Compatibility Investments in Duopoly with Demand Side Spillovers under Different Degrees of Cooperation**	FS IV 99 - 2
Horst Albach	**Des paysages florissants? Une contribution à la recherche sur la transformation**	FS IV 99 - 3
Jeremy Lever	**The Development of British Competition Law: A Complete Overhaul and Harmonization**	FS IV 99 - 4
Damien J. Neven Lars-Hendrik Röller Zhentang Zhang	**Union Power and Product Market Competition: Evidence from the Airline Industry**	FS IV 99 - 5
Justus Haucap Uwe Pauly Christian Wey	**The Incentives of Employers' Associations to Raise Rivals' Costs in the Presence of Collective Bargaining**	FS IV 99 - 6
Jianbo Zhang Zhentang Zhang	**Asymptotic Efficiency in Stackelberg Markets with Incomplete Information**	FS IV 99 - 7
Justus Haucap Christian Wey	**Standortwahl als Franchisingproblem**	FS IV 99 - 8
Yasar Barut Dan Kovenock Charles Noussair	**A Comparison of Multiple-Unit All-Pay and Winner-Pay Auctions Under Incomplete Information**	FS IV 99 - 9
Jim Y. Jin	**Collusion with Private and Aggregate Information**	FS IV 99 - 10

Absender/From:

Versandstelle - WZB
Reichpietschufer 50

D-10785 Berlin

BESTELLSCHEIN / ORDERFORM

Bitte schicken Sie bei Ihren Bestellungen von WZB-Papers ***unbedingt*** eine ***1-DM-Briefmarke pro paper*** und einen an Sie adressierten ***Aufkleber*** mit. Danke.

For each paper you order please send a ***"Coupon-Réponse International"*** (international money order) plus ***a self-addressed adhesive label***. Thank You.

Bitte schicken Sie mir aus der Liste der Institutsveröffentlichungen folgende Papiere zu:

Please send me the following papers from your Publication List:

Paper Nr./No.	Autor/Author + Kurztitel/Short Title