Essays on Political Economy, Information, and Welfare

by

Johan Lagerlöf

A Dissertation for the
Degree of Doctor of Philosophy, Ph.D.

Department of Economics
Stockholm School of Economics
1997

STOCKHOLM SCHOOL
OF ECONOMICS
THE ECONOMIC RESEARCH INSTITUTE

ISBN 91-7258-465-3

Stockholm 1997

For my parents

Acknowledgments

This thesis has developed while I have been a graduate student at the Departments of Economics of (first) Stockholm University and (later) the Stockholm School of Economics. My main supervisor at both places and throughout the thesis work has been Jörgen Weibull. I am in great intellectual debt to him. Jörgen has taught me so much about economic modeling and how to write a paper. He has also been an invaluable source of encouragement, constructive criticism, and useful suggestions.

My secondary supervisors have been Torsten Persson (Chapter 1) and Karl Wärneryd (Chapters 2 and 3). I am deeply indebted to both of them for introducing me to the field of political economy, for encouragement and advice, and for helpful comments on my papers.

I have also benefited from comments from and discussions with numerous other individuals in Stockholm and elsewhere. Many of these people are mentioned in the acknowledgments of each separate chapter. Administrative personnel at Stockholm University and the Stockholm School of Economics have made my life as a graduate student easier. In particular I would like to thank Bibi Dahlbeck, Ritva Kiviharju, and Kerstin Niklasson.

Financial support from Jan Wallander and Tom Hedelius' Research Foundation is gratefully acknowledged. During the academic year 1997/98 I am visiting the W. Allen Wallis Institute of Political Economy at the University of Rochester. I am very grateful to the Wallis Institute for hosting me, and to *Svenska Institutet* and *Finanspolitiska Forskningsinstitutet* for providing financial support.

Finally I would like to express my gratitude to some people outside the world of economics for being much greater friends than I deserve. Special thanks go to Bele, Cristina, and Lasse.

Johan Lagerlöf

Rochester, December 1997

Contents

Chapter 1

Lobbying, information, and private and social welfare*

* This paper previously appeared in the European Journal of Political Economy 13 (1997): 615-637. It is reprinted here by permission of Elsevier Science.

Lobbying, Information, and Private and Social Welfare

Johan Lagerlöf*

Received Sept 1996 Revised Dec 1996 Accepted Jan 1997

Abstract

Lobbying is potentially welfare improving if it is a form of information transmission. This paper studies when and for whom this is the case. In the model, a government (G) can redistribute income between two interest groups (A and B). Only one of them (A) has the opportunity to lobby. Lobbying is understood as A's acquiring and strategically revealing policy-relevant information to G. Among the results is that A may be worse off as a consequence of lobbying while B is better off. It can also be the case that lobbying makes both the active and passive interest groups and the government all worse off.

JEL classification: D72; D82

Keywords: Lobbying; Persuasion; Influence activities; Interest groups; Disclosure; Information acquisition

1. Introduction

The last few years have witnessed the emergence of a literature modelling lobbying as a form of strategic information transmission (Austen-Smith and Wright, 1992; Potters and van Winden, 1992; Ainsworth, 1993; Rasmusen, 1993).[1] The starting point for this literature is the very plausible idea that one reason why lobbyists are able to influence public policy is that they either have or can acquire

* Internet address: nejla@hhs.se; postal address: Department of Economics, Stockholm School of Economics, P.O. Box 6501, S-113 83 Stockholm, Sweden.

[1] See also Ball (1995), Lohmann (1995), Sloof (1995a, 1995b), and Sloof and van Winden (1996). For a survey of informational and other kinds of lobbying models, see Potters and van Winden (1993).

information that is relevant to the politician in his policy making. This information may for instance concern the electorate's preferences regarding different policies or the cost of public good production. Lobbying which is perceived as being a form of information transmission is potentially welfare improving. In this paper I investigate when and for whom this is the case. Answering this question is not only of normative interest, but may also explain positive issues such as why lobbying is often institutionalized rather than being prohibited.

I consider two interest groups, A and B, who each have exogenous income. B's income is certain. A, if confronted by a bad state of nature, however, suffers a loss. Initially no-one knows whether a good or a bad state will occur. A government, G, can redistribute income between the two groups. A and B have preferences over their own net incomes, and G cares about the well-being of both A and B.[2] G is willing to provide A with insurance against the realization of a bad state of nature. A may thus have an incentive to try to change G's beliefs about the state of nature that has been realized. A's opportunity to do this—to lobby—is modelled as follows (B is assumed politically passive, and not to have an opportunity to lobby).[3] A chooses whether or not to conduct an investigation, which involves a certain cost. The investigation may be successful or may not, the probability of success being exogenous and fixed. In case of success, A learns the true state of nature perfectly; in case of failure, A learns nothing. After a successful investigation A chooses whether to report the true state to G. In doing this, A cannot make false assertions. (I discuss this assumption in the concluding section.)

If the cost of information acquisition is not too great, there will exist an equilibrium where A investigates and then reports when having learned that the state is bad, but only then. Accordingly, when A lobbies in this sense, G will either decide on the size of the transfer knowing that the true state is bad or make this decision under uncertainty. In the latter case it is assumed that G uses Bayes' rule to update his prior beliefs about the true state. I compare the ex ante payoffs in such an equilibrium of the lobbying model with the equilibrium of a benchmark model where information acquisition (and hence lobbying) is not possible.

Even though I specifically refer to lobbying, the model is consistent with a wide range of interpretations other than interest groups and government. For instance, A may be the head of a division of an organization pushing G, the director of the organization, for a larger share of the organization's budget. If A is allocated a larger share, it is at the expense of B, the passive head of another division of the

[2] G's objective function may be thought of as a reduced form of some non-modelled electoral process; see Section 2.

[3] The reason why B cannot lobby may be that B is a large and diffused group, which is less able to overcome the free rider problem.

organization. Second, one may interpret the model as a sales encounter: A is a seller and G is a buyer who is to purchase some quantity of a commodity having an unknown quality. A third interpretation is that A is a child asking its parent, G, for more pocket money. In the last two examples, B may be thought of as someone socially related to A or G, such as the buyer's wife or the child's sister or brother.

The main message of the paper is that, even where lobbying plays an exclusively informational role, it is not necessarily welfare improving. It may namely be that A, the lobbyist himself, is *worse* off from his opportunity to lobby. Moreover, if A is worse off, it may be that also B and G are worse off, in which case we can conclude that lobbying is unambiguously bad, only using the Pareto criterion. The result that A may be worse off is due to an informational assumption in the model: G cannot observe whether A has chosen to investigate. Hence, A may be trapped into a situation where his best response is to acquire costly information, but where he would be better off if he were able to commit himself not to do this. In effect, the benchmark model allows for such a commitment. In order to obtain the result that also G and B are worse off, we need to assume that G cares about the well-being of A to the extent that G is willing to compensate A for (some of) the costs of information acquisition. This is the case if G simply maximizes a weighted sum of A's and B's payoffs. Then, if A is trapped into the situation where he investigates, G will respond by passing on some of the costs to B. This in turn makes also B worse off while it only mitigates A's loss; and since both A and B are worse off, so is G. It may also be the case that some people in society are better off from lobbying while others are worse off; but here the gainers are not members of the lobbying group (A), but the passive *non-lobbying* group (B).

The lobbying model underlying the welfare analysis in this paper is closely related to the persuasion game in Milgrom (1981). He models a sales encounter where the buyer is to purchase some quantity of a commodity having an unknown quality. For each level of quality, there is a quantity which maximizes the buyer's payoff; and the seller's payoff is increasing in the quantity purchased. It is assumed that the seller has some private information about the quality of the commodity, represented by n stochastic variables, each of which may take on any value on the real line. Prior to the buyer's purchase decision, the seller may report or conceal any of the variables, but he cannot misreport them.[4] In my model I simplify this setting by assuming that $n = 1$ and by saying that this single variable may take on one of only two values. I also extend the Milgrom model by assuming that at

[4]Grossman (1981) also analyses a similar model of disclosure. Milgrom and Roberts (1986) and Okuno-Fujiwara, Postlewaite, and Suzumura (1990) use the same basic framework and allow for more than one interested party. Lipman and Seppi (1995) examine the case where there are many interested parties whose ability to prove claims is limited.

the outset of the game neither the seller nor the buyer (A respectively G in my model) knows the realization of the stochastic variable, but the seller may make an attempt to acquire this information by privately incurring a designated cost.

The stylized way of modelling the information acquisition in my extension of Milgrom's (1981) persuasion game is borrowed from Laffont's and Tirole's (1991; 1993 ch. 11) model of regulatory capture: A is either perfectly informed about the realization of the stochastic variable or he is as ignorant as G; and each one of these possibilities occurs with an exogenous and fixed probability. In the Laffont and Tirole setting, however, whether or not to acquire information is not a choice variable. In providing a rationale for advocacy, Dewatripont and Tirole (1995) do endogenize the acquisition of information, and they also retain the above stylized way of modelling this acquisition. In contrast to my model, however, Laffont and Tirole and Dewatripont and Tirole adopt a principal-agent approach: they assume that the parties can design and commit to a long-term contract.[5]

My model in also related to Crawford and Sobel (1982) and other so-called cheap talk games, i.e., sender-receiver games where signalling has no cost to the sender other than that inherent in its effect on the receiver's choice of action. Crawford and Sobel assume that the sender has private information about a variable relevant for the receiver's decision, and prior to this decision the sender costlessly chooses an element belonging to some arbitrary set of messages. Austen-Smith (1994) extends the Crawford and Sobel model, which makes for a model somewhat more similar to mine. First, by endogenizing the sender's information acquisition, he assumes that the receiver does not know whether the sender is informed or not. Second, if the sender is informed he is able to prove this to the receiver at no cost, whereas it is impossible for the sender to prove that he is uninformed. In contrast to my model, however, the sender is still able to choose any message from an arbitrary set. Austen-Smith shows that in this setting there exist equilibria where information is transmitted under circumstances where such equilibria do not exist in the Crawford and Sobel model.

Ball (1995) and Lohmann (1993) develop welfare analyses of models where information transmission plays an important role. Ball sets up a model of lobbying where an interest group's contribution to a policy maker is a signal about the public's preferences. The benefit of lobbying lies in its information transmitting function, and the cost is that it may drive the chosen policy away from the social optimum. In Ball's model, lobbying may or may not be socially beneficial depending on parameter values. Lohmann considers a model where people may take costly political action (e.g., sign petitions, take part in demonstrations, or participate in violent riots) prior to a vote (or, in another version of the model, a political leader's decision). Each individual is assumed to have some private and

[5]Shavell (1994) also studies a model of costly acquisition of information; see footnote 13.

independent information relevant to the decision to be taken. By taking political action, the individual may be able to signal his information to the other voters (the political leader). Lohmann shows that in this model, political action may be both under- and oversupplied in equilibrium.

The remainder of the paper is organized as follows. Section 2 presents the lobbying model underlying the welfare analysis. In Section 3 a benchmark model is described and analysed, and the equilibrium behaviour of the lobbying model is characterized. Section 4 carries out the welfare analysis. In Section 5 an alternative specification of the agents' preferences is considered. Section 6 briefly summarizes and concludes.

2. The Model

Consider a society consisting of a government, G, and two interest groups, A and B. G cares about the well-being of A as well as of B in a way soon to be specified. G also has the power to redistribute money between the two groups. A transfer received by A and financed by a tax on B is denoted by t.[6] A has income $y_A > 0$. However, if A faces a bad state of nature he suffers a loss $L > 0$. Initially no-one can observe the true state of nature, but it is common knowledge that a bad (respectively good) state occurs with probability q (respectively $1-q$), where $q \in (0,1)$. B has income $y_B > 0$, and he faces no uncertainty about his income. Prior to G's deciding on the size of t, A may make an attempt to learn the true state of nature; and if succeeding in this, A chooses whether or not to lobby—that is, whether or not to share his knowledge with G. The formal game-theoretic model is described in greater detail below. First note, however, that since G and A are the only ones who make choices, we do not have to consider B as a player of the game. Still, at the end of the game, B does receive a payoff. This renders it possible to study the effects of A's opportunity to lobby on the welfare of B.

First A chooses whether to conduct an investigation. Formally, A selects an *investigation strategy* $\iota \in \{I, NI\}$, where $\iota = I$ means investigating and $\iota = NI$ means not investigating. By investigating, A privately incurs a certain monetary cost $c > 0$. With an exogenous and fixed probability $\tau \in (0,1)$, the investigation is a success and reveals the true state for A. With the complementary probability $(1-\tau)$ the investigation is a failure, in which case A learns nothing.[7] Two important informational assumptions are made: G cannot observe (i) A's choice

[6]The formal direction of this transfer—if it is from A to B or from B to A—is not important since t may take negative values.

[7]The assumption that one can never know for sure whether the investigation will bear fruit (i.e., the assumption that $\tau < 1$) seems realistic, and it guarantees that beliefs can always be updated with the help of Bayes' rule.

whether or not to investigate nor (ii) whether an investigation is a success or a failure.

In case $\iota = I$ and the investigation is a success, A chooses whether or not to report the result of the investigation to G. Formally, if having learned the true state of nature, A chooses a pair of reporting strategies $\rho_i \in \{R, NR\}, i = b, g$, where b is short for bad and g is short for good. $\rho_i = R$ means reporting when knowing that the state is i, and $\rho_i = NR$ means not reporting when knowing that the state is i. It is thus assumed that A is not allowed to leave an incorrect report; he may, however, withhold information that he has. As a consequence, when having received a report, G knows the true state with certainty. Otherwise G will only know that A did not submit a report. G will *not* know whether this is because A chose to withhold information that he actually has or because A is just as ignorant as G is. Moreover, in case A indeed is ignorant about the true state, G will not know whether this is due to a failed investigation or to the fact that A chose not to investigate. Let α denote the probability G assigns to the state of nature being bad conditional on not having received a report.

After A's having chosen his reporting strategies, G chooses three *decision strategies* $t_i \in \Re$, $i = b, g, N$, where t_i is the transfer G decides on after having learned that the state is bad (t_b), good (t_g), or after not having received a report (t_N). Finally, after G has chosen t_i, the agents' payoffs are realized. A and B have preferences over their net income, their payoff functions being $U^A(y_A - LD_1 - cD_2 + t)$ and $U^B(y_B - t)$; D_1 is an index variable which equals one if the state of nature is bad and zero otherwise, and D_2 is an index variable which equals one if A has investigated and zero otherwise.

G's payoff is the weighted sum of B's utility and A's utility *as it were had A incurred a fraction $\theta \in \{0,1\}$ of the cost c:*

$$U^G(t, D_1, \theta) \equiv U^A(y_A - LD_1 - \theta c D_2 + t) + \lambda U^B(y_B - t)\,. \tag{2.1}$$

By assuming that c may not enter G's payoff function, I allow for the possibility that G does care about the well-being of A, but only to a certain extent: G is willing to support A when A unluckily has suffered a loss in income, but G is not willing to compensate A for costs that he incurs in an attempt to convince G that a loss indeed has occurred. The parameter $\lambda > 0$ measures the relative weight of B's utility in G's objective function. G's payoff may be thought of as a reduced form of some non-modelled electoral process. One possibility is that the real objective of G is to be re-elected, and redistributing money so that (2.1) is maximized is a means to this end.[8] Alternatively G may be a citizen genuinely having the altruistic preferences specified in (2.1), and the political process is such

[8]For models of an electoral competition where, in equilibrium, some social welfare function is maximized, see Coughlin (1992) or Lindbeck and Weibull (1987).

that, in equilibrium, G is chosen policy maker.[9] In any of these cases, the question whether the cost c should enter G's (reduced form) payoff function or not should depend on details of the specific model one has in mind, and my personal view is that none of the alternatives seems on a priori grounds more reasonable than the other. Therefore, by letting θ equal either zero or one, both possibilities are incorporated into the model. We will see that varying the parameter θ in this manner alters the results of the welfare analysis qualitatively.

I assume that $U^A(\cdot)$ and $U^B(\cdot)$ are characterized by constant absolute risk aversion (CARA). That is, $U^i(C) = -\frac{1}{\sigma_i}e^{-\sigma_i C}$ $(i = A, B)$, where the parameter $\sigma_i > 0$ is the degree of absolute risk aversion (i.e., $\sigma_i = -U''/U'$) and C is the argument. Although convenient to work with, this functional form does unfortunately not satisfy the usual Inada condition $\lim_{C \to 0} U'(C) = +\infty$. Therefore, in order to avoid corner solutions, we will have to impose certain restrictions on the parameters. In Section 3 this will be done implicitly, and in Section 4 I will explicitly state the restrictions so that one may assure oneself that they are not inconsistent with any assumptions needed for the results of the welfare analysis.

Finally a remark. I have restricted attention to behaviour where the players do not randomize in their choices. Concerning A's reporting strategies and G's decision strategies, this is without loss of generality when the preferences are as specified.[10] However, even with these preferences, there may exist equilibria where A randomizes between investigating and not investigating.

3. Equilibrium Behaviour

3.1. A Benchmark Case: No Lobbying

Before analysing the model presented in the preceding section, I will look at a benchmark case where information acquisition is not allowed for and where there accordingly is no scope for lobbying. (In terms of the above model this may be thought of as setting $\tau = 0$.) In particular I am interested in A's, B's, and G's ex ante payoffs in the equilibrium of this benchmark model, since I wish to relate these to the ex ante payoffs in an equilibrium of the lobbying model.

When information acquisition is not allowed for (and this fact is common knowledge), A does not have an active role in the game since he cannot investigate

[9] Besley and Coate (1995) develop a model where participation in the policy making process is derived endogenously. First citizens decide whether or not to declare themselves as candidates for public office. Then citizens vote over the declared candidates. Finally the winning candidate selects a policy alternative. Besley and Coate show that in this setting, political competition will tend to sort in more altruistic citizens.

[10] In the alternative specification of the preferences considered in Section 5, this claim is true only for G's decision strategies.

and thus not report anything. Consequently A's strategies are irrelevant and, at the final stage of the game, G's problem is to maximize his expected payoff only knowing the prior distribution of the state of nature. Hence G will set the transfer equal to t_{BM} (BM stands for benchmark), where

$$t_{BM} \equiv \arg\max_{t \in \Re} (1-q) U^A (y_A + t) + qU^A (y_A - L + t) + \lambda U^B (y_B - t) \tag{3.1}$$

$$= \frac{\sigma_B y_B - \sigma_A y_A}{\sigma_A + \sigma_B} + \ln \left(\frac{1 - q + qe^{\sigma_A L}}{\lambda} \right)^{\frac{1}{\sigma_A + \sigma_B}}. \tag{3.2}$$

This means that the ex ante payoffs of A, B, and G—denoted by Π^A_{BM}, Π^B_{BM}, and Π^G_{BM}, respectively—equal

$$\Pi^A_{BM} \equiv (1-q) U^A (y_A + t_{BM}) + qU^A (y_A - L + t_{BM}), \tag{3.3}$$

$$\Pi^B_{BM} \equiv U^B (y_B - t_{BM}), \tag{3.4}$$

$$\Pi^G_{BM} \equiv \Pi^A_{BM} + \lambda \Pi^B_{BM}. \tag{3.5}$$

By substituting t_{BM} as given in (3.2) into these expressions and using the specified functional form, one obtains algebraic expressions for Π^A_{BM}, Π^B_{BM}, and Π^G_{BM}. Such expressions will be presented in Table 1 in Section 4.

3.2. The Lobbying Model

The solution concept that will be employed in the analysis of the lobbying model is that of perfect Bayesian equilibrium. A perfect Bayesian equilibrium requires that (i) the players make optimal choices at all information sets of the extensive form of the game, and (ii) when not having received a report, G updates his beliefs about the true state of nature using Bayes' rule. (Since $\tau < 1$, the expression for G's posterior beliefs is always defined.) A list of strategies and beliefs meeting those requirements are denoted by $(\iota^e, \rho^e_b, \rho^e_g, t^e_b, t^e_g, t^e_N)$ and α^e.

Since in Section 2 I restricted attention to equilibria where the players do not randomize in their choices, we have the following remaining possibilities. First, A may either investigate or not. Second, A may report (i) always, (ii) never, (iii) only when the state of nature is bad, or (iv) only when the state of nature is good. However, it should be intuitively clear that A never has an incentive to report about a good state of nature, and he always has an incentive to report about a bad state of nature. Accordingly an equilibrium must belong to one of the following two categories:

1. *Investigation and Partial Revelation*: Equilibria where A investigates and then reports only when knowing that the true state of nature is bad.

2. *No Investigation*: Equilibria where A does not investigate. Out of equilibrium he reports only when knowing that the true state of nature is bad.

Let us start with considering equilibria belonging to the first category, Investigation and Partial Revelation (or only Partial Revelation for short). What will the size of the transfer be in such an equilibrium? On the equilibrium path, G will in this case receive a report when the investigation is a success and when the state of nature is bad. Then G knows with certainty that A has incurred a loss (and that he has investigated) and thus sets the transfer equal to t_b^*, where

$$\begin{aligned} t_b^* &\equiv \arg\max_{t_b \in \Re} U^A\left(y_A - L - \theta c + t_b\right) + \lambda U^B\left(y_B - t_b\right) \\ &= \frac{\sigma_B y_B - \sigma_A\left(y_A - \theta c - L\right)}{\sigma_A + \sigma_B} - \ln \lambda^{\frac{1}{\sigma_A + \sigma_B}}. \end{aligned} \tag{3.6}$$

Note that since the investigation is a success with probability τ and a bad state of nature occurs with probability q, in an equilibrium of this category the transfer will (from an ex ante perspective) be set equal to t_b^* with probability τq.

In the out of equilibrium event that G were to receive a report saying that the state of nature is good, G sets the transfer equal to t_g^*, where

$$\begin{aligned} t_g^* &\equiv \arg\max_{t_g \in \Re} U^A\left(y_A - \theta c + t_g\right) + \lambda U^B\left(y_B - t_g\right) \\ &= \frac{\sigma_B y_B - \sigma_A\left(y_A - \theta c\right)}{\sigma_A + \sigma_B} - \ln \lambda^{\frac{1}{\sigma_A + \sigma_B}} \\ &= t_b^* - \frac{\sigma_A}{\sigma_A + \sigma_B} L. \end{aligned} \tag{3.7}$$

Of course we have $t_b^* > t_g^*$: A receives a greater transfer when G knows that A will incur a loss. Also note that the difference $t_b^* - t_g^*$ equals a certain fraction $\frac{\sigma_A}{\sigma_A + \sigma_B}$ of the loss L, the fraction being greater the more risk averse A is and the less risk averse B is.

When A's investigation is either a failure or has shown that the true state of nature is good, G does not receive a report (this will happen with probability $1 - \tau q$). Then, taking into account A's incentives to strategically withhold information, G updates his prior beliefs about the true state using Bayes' rule:

$$\alpha^e = \frac{\Pr\left(NR \mid bad\right) \Pr\left(bad\right)}{\Pr\left(NR\right)} = \frac{\left(1 - \tau\right) q}{1 - \tau q} \equiv \widetilde{\alpha}. \tag{3.8}$$

(The new notation $\widetilde{\alpha}$ is convenient and will be used now and then throughout the paper.) The posterior is smaller than the prior, the intuition being that if A has not reported it may be because he knows that the state of nature is good but withholds this information. It is easily verified that $\widetilde{\alpha}$ is strictly increasing in q and strictly decreasing in τ. In particular we have that in the limit, when $\tau \longrightarrow 1$ ($\tau \longrightarrow 0$), then $\widetilde{\alpha} \longrightarrow 0$ ($\widetilde{\alpha} \longrightarrow q$). That is to say, when the investigation is very likely to have been a success, A's keeping silence is very likely to be due to his knowing that the good state has occurred. Being aware of this, G assigns a low probability to the state of nature being bad. (And similarly for the case $\tau \longrightarrow 0$.)

Thus, when G has not received a report, the transfer t is set so as to maximize G's expected payoff when assigning the probability $\widetilde{\alpha}$ to the state of nature being bad, yielding the optimal transfer $\widetilde{t}$, where

$$\begin{aligned}
\widetilde{t} &\equiv \arg\max_{t_N \in \Re} (1-\widetilde{\alpha}) U^A (y_A - \theta c + t_N) + \widetilde{\alpha} U^A (y_A - L - \theta c + t_N) + \lambda U^B (y_B - t_N) \\
&= \frac{\sigma_B y_B - \sigma_A (y_A - \theta c)}{\sigma_A + \sigma_B} + \ln \left(\frac{1 - \widetilde{\alpha} + \widetilde{\alpha} e^{\sigma_A L}}{\lambda} \right)^{\frac{1}{\sigma_A + \sigma_B}} . \qquad (3.9)
\end{aligned}$$

Hence the transfer $\widetilde{t}$ is set as a compromise between the transfers chosen under full information: $t_b^* > \widetilde{t} > t_g^*$.

When does an equilibrium with Partial Revelation exist? A condition that is both necessary and sufficient is that A indeed has an incentive to investigate:

$$\begin{aligned}
\Pi_{PR}^A &\equiv \tau q U^A (y_A - L - c + t_b^*) + (1-\tau) q U^A \left(y_A - L - c + \widetilde{t}\right) + (1-q) U^A \left(y_A - c + \widetilde{t}\right) \\
&\geq q U^A \left(y_A - L + \widetilde{t}\right) + (1-q) U^A \left(y_A + \widetilde{t}\right) . \qquad (3.10)
\end{aligned}$$

The left-hand side of inequality (3.10), Π_{PR}^A, is A's ex ante payoff in an equilibrium with Partial Revelation. The right-hand side is his ex ante payoff if unilaterally choosing not to investigate. Then A will never be able to report, and G will hence always choose the transfer $\widetilde{t}$. By carrying out some algebra, one may show that inequality (3.10) may equivalently be written $c \leq c_{PR}$, where

$$c_{PR} \equiv \ln \left(\frac{\left(1 - \widetilde{\alpha} + \widetilde{\alpha} e^{\sigma_A L}\right)^{\frac{-\sigma_A}{\sigma_A + \sigma_B}} \left(1 - q + q e^{\sigma_A L}\right)}{\tau q e^{\frac{\sigma_A \sigma_B}{\sigma_A + \sigma_B} L} + (1 - \tau q) \left(1 - \widetilde{\alpha} + \widetilde{\alpha} e^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A + \sigma_B}}} \right)^{\frac{1}{\sigma_A}} . \qquad (3.11)$$

It always holds that $c_{PR} > 0$.[11] We have the following proposition.

Proposition 1. *If and only if $c \leq c_{PR}$ does there exist an equilibrium with Investigation and Partial Revelation (i.e., then there exists an equilibrium $\left(\iota^e, \rho_b^e, \rho_g^e, t_b^e, t_g^e, t_N^e\right) = \left(I, R, NR, t_b^*, t_g^*, \widetilde{t}\right)$ and $\alpha^e = \widetilde{\alpha}$).*

[11] Since we have $c_{PR} > \underline{c}$, this is shown indirectly when in the Appendix I show that $\underline{c}$ (which is defined in (4.6)) is greater than zero.

Now let us consider equilibria belonging to the second category, No Investigation. In this case A will not be able to submit a report. In the out of equilibrium event that G did receive a report, G would set t equal to t_b^* or t_g^*, depending on whether the state of nature is bad or good. Always in equilibrium, however, G will set t equal to t_{BM}, the transfer chosen in the benchmark model. This is because when A does not investigate, applying Bayes' rule yields a posterior which equals the prior, $\alpha^e = q$. This result makes perfect sense since if A is as ignorant as G about the true state of nature, then there is nothing G can learn from A's not reporting.

Analogously with the case with equilibria of the first category, a condition that is both necessary and sufficient for an equilibrium of the second category to exist is that A must not have an incentive to deviate unilaterally by investigating:

$$\begin{aligned}\Pi_{BM}^{A} \geq\ & \tau q U^{A}\left(y_{A}-L-c+t_{b}^{*}\right)+(1-\tau) q U^{A}\left(y_{A}-L-c+t_{BM}\right) \\ & +(1-q) U^{A}\left(y_{A}-c+t_{BM}\right) \end{aligned} \tag{3.12}$$

or, equivalently,

$$e^{\sigma_A c} \geq \frac{\left(1-q+qe^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}}}{\tau q e^{\frac{-\sigma_A\left(\sigma_A \theta c-\sigma_B L\right)}{\sigma_A+\sigma_B}}+(1-\tau q)\left(1-\widetilde{\alpha}+\widetilde{\alpha} e^{\sigma_A L}\right)\left(1-q+qe^{\sigma_A L}\right)^{\frac{-\sigma_A}{\sigma_A+\sigma_B}}}. \tag{3.13}$$

The left-hand side of inequality (3.12), Π_{BM}^{A}, is A's ex ante payoff in the benchmark model as defined in (3.3) or, equivalently, in an equilibrium with No Investigation. The right-hand side is his ex ante payoff if unilaterally choosing to investigate. Then he will submit a report only when the investigation is a success and the state of nature is bad. We have the following proposition.

Proposition 2. *If and only if inequality (3.13) holds does there exist an equilibrium with No Investigation (i.e., then there exists an equilibrium* $\left(\iota^e, \rho_b^e, \rho_g^e, t_b^e, t_g^e, t_N^e\right) = \left(NI, R, NR, t_b^*, t_g^*, t_{BM}\right)$ *and* $\alpha^e = q$*).*

One may easily show that, if $\theta = 0$, then there is an overlap between the region where an equilibrium with Partial Revelation exists and the corresponding region for an equilibrium with No Investigation. However, if $\theta = 1$, inequality (3.13) cannot be solved for c explicitly, and it is not straightforward to see whether equilibria of the two categories may co-exist. Still, the kind of equilibrium which is relevant for the welfare analysis is the one with Partial Revelation, i.e., that where A actually lobbies. (In an equilibrium with No Investigation, the ex ante payoffs will be identical to the ones in the benchmark model.) Thus, in the following analysis I shall only consider the subset of the parameter space where such an equilibrium exists, i.e., the subset satisfying $c \leq c_{PR}$.

4. Welfare Analysis

The welfare analysis entails studying the signs of the expressions $\Pi^i_{PR} - \Pi^i_{BM}$, $i = A, B, G$. The notation Π^i_{BM}, which expresses i's ex ante payoff in the benchmark model, was defined in Section 3. Π^i_{PR} is i's ex ante payoff in an equilibrium of the lobbying model where there is Partial Revelation; Π^A_{PR} was defined in the left-hand side of inequality (3.10), and Π^B_{PR} and Π^G_{PR} are defined analogously.

As pointed out in Section 2, the CARA utility functions do not guarantee interior solutions to G's maximization problems. Hence, in the preceding section, I implicitly imposed certain restrictions on the parameters. It is important to ensure that those restrictions are not inconsistent with any assumptions needed for the results of the welfare analysis. The following inequalities need to hold in order to avoid corner solutions:

$$0 < \frac{\sigma_B}{\sigma_A + \sigma_B} (y_A + y_B) - L + \ln \left(\frac{1 - q + q e^{\sigma_A L}}{\lambda} \right)^{\frac{1}{\sigma_A + \sigma_B}}, \tag{4.1}$$

$$0 < \frac{\sigma_A}{\sigma_A + \sigma_B} (y_A + y_B) - \ln \left(\frac{1 - q + q e^{\sigma_A L}}{\lambda} \right)^{\frac{1}{\sigma_A + \sigma_B}}, \tag{4.2}$$

$$c < \frac{\sigma_B (y_A + y_B) - (\sigma_A + \sigma_B) L}{\sigma_A (1 - \theta) + \sigma_B} + \ln \left(\frac{1 - \widetilde{\alpha} + \widetilde{\alpha} e^{\sigma_A L}}{\lambda} \right)^{\frac{1}{\sigma_A (1-\theta) + \sigma_B}}, \tag{4.3}$$

$$\theta c < y_A + y_B - L + \ln \lambda^{\frac{1}{\sigma_A}}. \tag{4.4}$$

The first two conditions guarantee that A's and B's incomes always exceed zero in an equilibrium of the benchmark model (the inequalities (4.1) and (4.2) are equivalent to $y_A - L + t_{BM} > 0$ and $y_B - t_{BM} > 0$, respectively); and the third and fourth conditions guarantee the same thing concerning equilibria with Partial Revelation in the lobbying model (the inequalities (4.3) and (4.4) are equivalent to $y_A - L - c + \widetilde{t} > 0$ and $y_B - t^*_b > 0$, respectively). All four conditions are satisfied if $y_A + y_B$ is sufficiently great. Assuming that this is the case is non-restrictive in the sense that—as we will soon see—the qualitative results of the welfare analysis do not depend on the magnitude of $y_A + y_B$.

Table 1 lists the agents' ex ante payoffs. Although the expressions for these payoffs are complex, the different agents' payoff levels relate to each other in a simple manner. For instance, one sees that in the benchmark model, A's ex ante payoff equals $\frac{\sigma_B}{\sigma_A}\lambda$ times the one of B, regardless of the values of the parameters q, y_A, y_B, and L. Thus, in the special case where $\lambda = 1$ and $\sigma_A = \sigma_B$—i.e., when A's and B's utilities have the same weight in G's objective function and the groups are equally risk averse—A's and B's ex ante payoffs have the same magnitude.

The same is true in equilibria with Partial Revelation when either the cost of information acquisition, c, is set equal to zero or when $\theta = 1$. G's ex ante payoff is the sum of A's and B's, as it were had A only incurred a fraction θ of the cost c.

[TABLE 1 ABOUT HERE]

From inspection of the table two important observations immediately follow. First, B is better off from lobbying if and only if G is: $\Pi^G_{PR} > \Pi^G_{BM} \Leftrightarrow \Pi^B_{PR} > \Pi^B_{BM}$. Second, if $\theta = 1$, then either all three agents are better off from lobbying or all three agents are worse off from lobbying:

$$\Pi^G_{PR} > \Pi^G_{BM} \Leftrightarrow \Pi^B_{PR} > \Pi^B_{BM} \Leftrightarrow \Pi^A_{PR}(\theta = 1) > \Pi^A_{BM}(\theta = 1)\,. \qquad (4.5)$$

To start with I shall look at the case $\theta = 0$. First consider the welfare of G and B (the proposition is proven in the Appendix):

Proposition 3. *If $\theta = 0$, then G and B are always strictly better off with lobbying than without (i.e., $\Pi^G_{PR} > \Pi^G_{BM}$ and $\Pi^B_{PR} > \Pi^B_{BM}$).*

The key to understanding the result for G is that, when $\theta = 0$, the cost of information acquisition does not enter G's payoff function. This fact and the fact that G acts last in effect make G's situation tantamount to a single agent decision problem. Ex post, G can guarantee himself the same payoff as in the benchmark model by simply setting $t = t_{BM}$. However, since G is a Bayesian, he always does better than that—whether he has received a report or not. Hence G is better off from lobbying also ex ante.

Let us now turn to B. Perhaps it seems somewhat counter-intuitive that although A and B have diametrically opposed interests concerning the transfer t, and A does what he can to manipulate the decision-making process in his favour, B is always (as long as $\theta = 0$) better off from his opponent's opportunity to carry on these manipulations. Moreover, this is true regardless which weight $\lambda > 0$ B's utility has in G's objective function. In order to understand this result, first note that when A lobbies, G is better informed than otherwise. This makes G better off from lobbying. The question, then, is: does G's being better informed induce G to make decisions which are (from an ex ante perspective) better for B?

The answer to this question will depend on the relation between G's and B's preferences; A's preferences only influence which different kinds of equilibria will exist. Hence, for B to be better off from A's opportunity to lobby, B's and G's preferences must be sufficiently aligned. At least with the functional form specified here (CARA), it turns out that any $\lambda > 0$ meets this requirement.[12] In summary, when A has the opportunity to lobby, then G always gains; and since a strictly positive weight is assigned to B's utility in G's objective function, a strictly positive fraction of these gains is allocated to B.

Still assuming that $\theta = 0$, we now turn to A. Define $\underline{c}$ as follows:

$$\underline{c} \equiv \ln \left(\frac{\left(1 - q + q e^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A + \sigma_B}}}{\tau q e^{\frac{\sigma_A \sigma_B}{\sigma_A + \sigma_B} L} + (1 - \tau q)\left(1 - \widetilde{\alpha} + \widetilde{\alpha} e^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A + \sigma_B}}} \right)^{\frac{1}{\sigma_A}}. \tag{4.6}$$

In the Appendix I show that $\underline{c} > 0$. The following proposition follows directly from Table 1.

Proposition 4. *Suppose that $\theta = 0$. Then A is strictly better (worse) off with lobbying than without if and only if $c < \underline{c}$ ($c > \underline{c}$) (i.e., $\Pi^A_{PR} \gtreqless \Pi^A_{BM}$ as $c \lesseqgtr \underline{c}$).*

Of course, the result that A is worse off from his opportunity to lobby when $c > \underline{c}$ would not be very interesting if an equilibrium with Partial Revelation did not exist for such values of c. However, straightforward algebra shows that the critical value for such an equilibrium to exist, c_{PR}, is strictly greater than $\underline{c}$. Hence there is a non-empty set $(\underline{c}, c_{PR}]$ such that if $c \in (\underline{c}, c_{PR}]$ (and if $\theta = 0$), then there exists an equilibrium with Partial Revelation (i.e., where A lobbies) but where A would be strictly better off under the benchmark model (i.e., if lobbying were not possible).[13]

Recall that when $\theta = 0$, an equilibrium with Partial Revelation and an equilibrium with No Investigation co-exist for some values of c. One may thus ask oneself if the phenomenon that A is worse off from his opportunity to lobby may

[12] However, it is not true in general that an interested party is ex ante better off from the decision maker's having access to more information. In Section 5 an example is provided where the decision maker's and the interested party's preferences are *not* sufficiently aligned in the above sense.

[13] This result is related to a result in Shavell (1994). Shavell examines a model with many sellers and many buyers. Before meeting a buyer, each seller may acquire costly information about the good. As in my model, the buyer cannot observe whether the seller has chosen to do this or not. The information is assumed to be socially valuable in the sense that it allows an action to be taken that increases the value of the good. Among Shavell's results is that when disclosure of information is voluntary, the amount spent by sellers on acquiring information is socially excessive. Shavell's measure of social welfare is the net value obtained from use of the good minus the information acquisition cost. Hence, he does not consider the question whether a seller (in my model, A) may be *individually* worse off when having access to the information technology. Another difference between my model and Shavell's is that he assumes risk neutral agents. Also, Shavell does not look at the welfare of a third party (in my model, B) whose payoff also is affected by the buyer's (G's) decision.

arise when an equilibrium with Partial Revelation is the *only* equilibrium of the lobbying model. The answer is yes. One may show that for any c belonging to a non-empty subset of $(\underline{c}, c_{PR}]$, an equilibrium with No Investigation does *not* exist.[14] I conjecture that for such a c, nor does there exist an equilibrium where A randomizes between investigating and not.[15] This remains to be shown, however.

The assumption driving the result that A may be worse off from lobbying is that G cannot observe A's choice whether to investigate. Suppose that G has not received a report but believes that A has investigated. Then G will take this into account when forming his posterior beliefs about the true state of nature. These beliefs affect the decision t_N made by G, in a direction which is to A's disadvantage (compared to the case when G does not think that A has investigated). At this point A may find it worthwhile to actually incur the cost c, i.e., to investigate. For if he does, and if the investigation is a success, A may—by reporting selectively to G about the result—be able to induce G to make a decision which on average is better for A. A's choice to incur c in turn confirms G's belief that A indeed investigated. In other words, because of G's inability to observe A's choice whether to investigate, A may be trapped into an equilibrium where he is worse off than he would be if he were able to commit himself not to investigate. In effect, the benchmark model allows for such a commitment.

Let us now turn to the case $\theta = 1$. Recall that under this assumption either all three agents are better off from lobbying or all three are worse off. It turns out that both these outcomes are possible to obtain. Let $\bar{c} \equiv \left(1 + \frac{\sigma_A}{\sigma_B}\right)\underline{c}$. (Thus $\bar{c} > \underline{c} > 0$.) The following proposition follows directly from Table 1.

Proposition 5. *Suppose that $\theta = 1$. Then A, B, and G are all strictly better (worse) off with lobbying than without if and only if $c < \bar{c}$ ($c > \bar{c}$) (i.e., $\Pi^i_{PR} \gtreqless \Pi^i_{BM}$ as $c \lesseqgtr \bar{c}$, $i = A, B, G$).*

Similarly to Proposition 4, the result that all three agents are worse off from A's opportunity to lobby when $c > \bar{c}$ would not be very interesting if an equilibrium with Partial Revelation did not exist for such values of c. Again, however, by carrying out some straightforward algebra one may show that the critical value for such an equilibrium to exist, c_{PR}, is strictly greater than $\bar{c}$. Hence there is a non-empty set $(\bar{c}, c_{PR}]$ such that if $c \in (\bar{c}, c_{PR}]$ (and if $\theta = 1$), then there exists an equilibrium with Partial Revelation (i.e., where A lobbies) but where

[14] By letting inequality (3.13) hold with equality and setting $\theta = 0$, one obtains a critical value of c such that for any c below this an equilibrium with No Investigation does not exist. By carrying out some algebra one may then check that this critical value is strictly greater than $\underline{c}$.

[15] More specifically I conjecture that this kind of equilibrium exists only when the other two co-exist.

everyone would be better off under the benchmark model (i.e., if lobbying were not possible).

May the phenomenon that everyone is worse off from A's opportunity to lobby arise when an equilibrium with Partial Revelation is the *only* equilibrium of the lobbying model? The answer is yes. In the Appendix it is shown that if c is close to $\bar{c}$ and if L is small, then an equilibrium with No Investigation does not exist. (The question is left open, however, whether for such a c there exists an equilibrium where A randomizes between investigating and not investigating.)

The intuition behind the result that everyone may be worse off from lobbying is straightforward. The reason why A may be worse off is the same as above (i.e., as in the case $\theta = 0$). Now, when $\theta = 1$, G takes into account the fact that A has incurred the cost c, and may thus be willing to compensate him for this when deciding on the transfer t. However, such compensation will be at the expense of B. The fact that G in this way passes on some of the costs to B also makes B worse off, while it only mitigates A's loss; and since both A and B are worse off, so is G.

As to the question concerning the desirability of lobbying, the welfare analysis in this section has revealed two possibilities: Either there is unanimity about A's opportunity to lobby; or B and G are better off while A is worse off. Thus, the popular view that those who gain from lobbying are members of the lobbying group (and those only) is not compatible with this model. However, the next section provides an example where B, independently of the welfare of A and G, may be either worse off or better off from A's opportunity to lobby; and whether B is better or worse off depends on the degree to which B's and G's preferences are aligned in a specific sense.

5. The Government Deciding on a Public Good

Consider the following circumstances, which are slightly different from those described in Section 2. We have again a government, G, one active (potentially lobbying) interest group, A, and one passive interest group, B. However, G's choice now concerns the quantity of a public good to produce. This quantity is denoted by x. G, A, and B have preferences over x respectively described by the payoff functions

$$U^G = -(x - x_G - D_1)^2, \tag{5.1}$$

$$U^A = -(x - x_A - D_1)^2 - cD_2, \tag{5.2}$$

$$U^B = -(x - x_B - aD_1)^2. \tag{5.3}$$

The notation c, D_1, and D_2 have the same meaning as before; we here make the interpretation $D_1 = 0$ ($D_1 = 1$) when the cost of public good production is high (low). $x_i + D_1$ is i's ($i = A, G$) ideal (or bliss) point: When the cost is low, i wants $x_i + 1$ units of the public good to be produced, and otherwise only x_i units. Without loss of generality, assume that $x_A \leq x_G$. B's ideal point is given by $x_B + aD_1$. The parameter a measures the responsiveness of B to changes in the cost of public good production. Hence, whereas A's and G's ideal points change in a parallel way as the cost of public good production changes, B's ideal point may change more (if $a > 1$) or less (if $a < 1$) than A's and G's ideal points.[16] It is natural to think of a as being strictly positive, but in principle it may be zero or even negative. The sequence of events is the same as in Section 2.

In an earlier working paper (Lagerlöf, 1996) I show the following. When A's and G's ideal points are close ($x_G - x_A$ is small) and when the cost of information acquisition (c) is not too great, then there exists an equilibrium where A investigates and then always reports when knowing the true state (Full Revelation); when $x_G - x_A$ is sufficiently great and c is sufficiently small, there exists an equilibrium where A investigates and then reports if and only if he knows that the cost is high (Partial Revelation); and if c is sufficiently great, A will not investigate in equilibrium. I also show that the welfare results for A and G are similar to the ones found in Section 4, in the special case $\theta = 0$: G is always better off from A's opportunity to lobby, while A himself may be worse off. For B we have the following result.

Proposition 6. *If and only if $a > \frac{1}{2}$, then B is strictly better off in an equilibrium with Full Revelation than in an equilibrium with Partial Revelation, and then B is strictly better off in an equilibrium with Partial Revelation than under the benchmark model.*

Thus, for B to be better off from A's opportunity to lobby (or from G's having access to more information), B's low cost ideal point must to a sufficient degree be greater than the high cost ideal point. In order to understand the mechanism behind this result, let us consider the simpler story where the comparison is between a situation where G only knows the prior distribution of the cost of public good production and a situation where G is perfectly informed about the true cost. One may verify that Proposition 6 extends to this alternative story. That is, B strictly prefers the situation where G is perfectly informed if and only if $a > \frac{1}{2}$.

The comparison is illustrated in Figure 1. When only knowing the prior, G will make the decision $x_G + q$. When being perfectly informed, G will either (when the cost is high) make the decision x_G or (when the cost is low) make the decision

[16] This kind of quadratic loss function is common in the literature. However, the assumption that people may differ as to their responsiveness to changes in the stochastic variable is, to the best of my knowledge, only found in Schultz (1996).

$x_G + 1$. B's ideal point equals x_B when the cost is high and $x_B + a$ when the cost is low. Thus, when G only knows the prior, B's ex post loss is the square of either (with probability $1 - q$) the distance denoted by V_h in the figure or (with probability q) the distance denoted by V_l. When G is perfectly informed, B's ex post loss is the square of either (with probability $1 - q$) the distance W_h or (with probability q) the distance W_l.

[FIGURE 1 ABOUT HERE]

Why does B ex ante want G to be informed if and only if $a > \frac{1}{2}$? To see this, let us decompose B's gain from G's being informed into two parts: (i) the gain B would obtain if he were risk neutral (i.e., if his loss were linear in the distance between his ideal point and G's decision) and (ii) the gain that is due to B's being risk averse. If B were risk neutral, then he would only care about the expected value of the distances marked in the figure. Let us assume that $a \geq 0$ and that $x_B \geq x_G + 1$. Then it turns out that these expected values are the same whether G is informed or not:[17]

$$(1 - q) V_h + qV_l = x_B - x_G + q(a - 1) = (1 - q) W_h + qW_l. \tag{5.4}$$

Hence, the gain B would obtain if he were risk neutral equals zero, and B's total gain from G's being informed must exclusively be attributed to B's being risk averse. B's risk aversion manifests itself in his wanting the distances marked in the figure to be as equal in size as possible. Taken together, this means that B's gain from G's being better informed is positive if and only if G's being better informed results in a smaller difference in the distances between B's ideal point and G's decision—that is, if and only if

$$|W_l - W_h| < |V_l - V_h|. \tag{5.5}$$

However, since $|W_l - W_h| = |1 - a|$ and $|V_l - V_h| = a$, condition (5.5) is equivalent to $a > \frac{1}{2}$.

6. Concluding Remarks

To sum up, the welfare analysis in Section 4 was carried out under two different assumptions concerning G's payoff function. Under the first assumption G maximized the weighted sum of B's utility and A's utility as it were had A not incurred the cost of information acquisition. Under the second assumption this

[17] This is due to the quadratic functional form.

cost did enter G's payoff function. When the cost does *not* enter, then (i) G and B always gain from A's opportunity to lobby; and (ii) A may be worse off from this opportunity. When the cost of information acquisition does enter G's payoff function either all three agents are better off from lobbying or all three agents are worse off from lobbying, and which of these outcomes will occur depends on the magnitude of the cost parameter.

We may thus conclude that even in a setting where lobbying plays an exclusively informational roll, it is not necessarily welfare improving. It may be that A, the lobbyist himself, is trapped into a situation where the best he can do is to acquire costly information, but where he would be better off if he could commit himself not to do this. Not only A may be worse off; if the cost of information acquisition enters G's payoff function, then G's willingness to compensate A for his costs may render also G and B worse off from A's opportunity to lobby. Hence, in this case we can conclude that lobbying is unambiguously bad, only using the Pareto criterion.

Moreover, when the cost of information acquisition does not enter G's payoff function, it may be that some people in society are better off from lobbying while others are worse off. However, in contrast to the popular view that those who gain are members of the lobbying interest group, here the *non-lobbying* group is the winner. This result is obtained in spite of the fact that A's and B's interests concerning the transfer are diametrically opposed. The reasons why B is better off are that G is better off and, since G is assigning a strictly positive weight to B's utility, some of these gains are allocated to B. This gives us to the perhaps trivial but still important insight that the question whether B is better or worse off from A's lobbying depends on the relation between B's and G's preferences; A's preferences only matter with regard to determining the different kinds of equilibria that will exist. This, in turn, suggests that it might be restrictive to take G's preferences as given. Hence, an interesting topic for future research should be to make the identity of G endogenous by integrating a model of electoral competition with an informational model of lobbying.

In Section 5 I considered an alternative specification of the agents' payoffs which gave the result that B may be worse off due to A's lobbying, even when G is better off. More specifically, whether B is better or worse off from lobbying (and from G's being better informed) depends on how much B's ideal point changes as the state of nature changes.

Throughout the paper the assumption that A cannot misrepresent was maintained. In some situations such an assumption may very well be reasonable. For instance, it may be commonly known that the information resulting from a successful investigation has the character of a logical argument (the investigation may show that a proposed bill is in conflict with the constitution, something that was not previously known but which is an obvious fact once discovered). Alternatively, it may be commonly known that the content of a successful investigation consists of documents which are not possible to forge.

In other situations, however, the assumption that the lobbyist cannot misrepresent may *not* be reasonable. One may then wonder to what extent the assumption drives the results of the paper. Some of the results, in particular those in Section 5, concern the welfare effects on B of A's having the opportunity to lobby. These results should continue to hold if one instead assumed that A may indeed misreport but that there is some other mechanism (e.g., an inspection game) guaranteeing that at least *some* information is transmitted in the lobbying process. If so, the welfare analysis amounts to asking whether G's having access to this information induces G to make decisions which are (from an ex ante perspective) better to B. This is precisely the question posed in this paper.

The result that A may be worse off from having the opportunity to lobby may be more sensitive to the assumption that A may not misreport. However, the result still holds in the following variation of the set-up in Section 2. Suppose that whether or not A knows the true state, he can report that the state is bad or good or choose not to leave a report: $\rho_i = \{b, g, NR\}$ $(i = b, g, N)$. However, if A reports "bad" when the true state is good, or vice versa, G will find out about this with some exogenous probability $\gamma \in (0, 1]$; if G "finds out", G learns the true state and A must pay a penalty $\delta > 0$. If A does not submit a report, he does not run the risk of having to pay the penalty. The rest of the game is as in Section 2. Clearly, if γ and/or δ are sufficiently great, A will report the truth if knowing that the state is bad and otherwise not report, and we thus reobtain the result that A may be worse off from his opportunity to lobby.

I have also assumed a great deal of rationality on the part of G. Whenever A does not submit a report, G uses Bayes' rule to update his prior beliefs about the true state of nature. One alternative to this specification is to assume that G is *naively credulous* in the sense of Milgrom and Roberts (1986, p. 23): If A does not submit a report, then G believes that A is as ignorant as G. Thus we always have $\alpha^e = q$. I conjecture that such an assumption would invalidate many of the results in the paper. For instance, at least when the cost of information acquisition does not enter G's payoff function, A will always be better off when having the opportunity to lobby. However, it can be shown that the result that G always is better off from lobbying whenever the cost does not enter his payoff function (Proposition 3) still holds.[18]

[18] If the cost does not enter G's payoff function, then $t_N^e = t_{BM}$. Substituting $\tilde{t}$ for t_{BM} in the proof of Proposition 3 does not invalidate any of the arguments showing that $\Pi^G_{PR} > \Pi^G_{BM}$.

Appendix

Proof of Proposition 3: Since we know that $\Pi_{PR}^{G} > \Pi_{BM}^{G} \Leftrightarrow \Pi_{PR}^{B} > \Pi_{BM}^{B}$, it suffices to show that under the assumption $\theta = 0$, $\Pi_{PR}^{G} > \Pi_{BM}^{G}$. However, when $\theta = 0$, $\Pi_{PR}^{G} > \Pi_{BM}^{G}$ may equivalently be written

$$\begin{aligned} 0 &< \tau q\left[U^{G}\left(t_{b}^{*},1,0\right)-U^{G}\left(t_{BM},1,0\right)\right] \\ &\quad +(1-\tau q)\left\{\left[\widetilde{\alpha}U^{G}\left(\widetilde{t},1,0\right)+(1-\widetilde{\alpha})\,U^{G}\left(\widetilde{t},0,0\right)\right]\right. \\ &\quad \left.-\left[\widetilde{\alpha}U^{G}\left(t_{BM},1,0\right)+(1-\widetilde{\alpha})\,U^{G}\left(t_{BM},0,0\right)\right]\right\}, \end{aligned} \tag{6.1}$$

where $U^{G}(t, D_1, \theta)$ is defined in equation (2.1). The right-hand side of inequality (6.1) is strictly positive since U^G is strictly concave and, per definition, t_b^* and $\widetilde{t}$ maximize $U^{G}(t_b, 1, 0)$ and $\widetilde{\alpha}U^{G}(t,1,0)+(1-\widetilde{\alpha})\,U^{G}(t,0,0)$, respectively.□

Proof of the claim that $\underline{c}$> 0: We know from Proposition 3 that when $\theta = 0$, $\Pi_{PR}^{B} > \Pi_{BM}^{B}$. Using the expressions for Π_{PR}^{B} and Π_{BM}^{B} in Table 1 and carrying out some algebra, we may write

$$\Pi_{BM}^{B} < \Pi_{PR}^{B} \Leftrightarrow \tag{6.2}$$

$$\theta c < \frac{\sigma_A+\sigma_B}{\sigma_A\sigma_B}\ln\left(\frac{\left(1-q+qe^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}}}{\tau q e^{\frac{\sigma_A\sigma_B}{\sigma_A+\sigma_B}L}+(1-\tau q)\left(1-\widetilde{\alpha}+\widetilde{\alpha}e^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}}}\right) \Leftrightarrow \tag{6.3}$$

$$\underline{c} > 0, \tag{6.4}$$

where the last equivalence was obtained by setting $\theta = 0$ and using the definition of $\underline{c}$ in (4.6).□

Proof of the uniqueness claim at the end of Section 4: The claim is that there is a $c \in (\overline{c}, c_{PR}]$ such that an equilibrium with No Investigation does not exist if $\theta = 1$. If $c = \overline{c}$ and an equilibrium with No Investigation does not exist, then the following inequality must hold (see condition (3.13)):

$$e^{\sigma_A\overline{c}} < \frac{\left(1-q+qe^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}}}{\tau q e^{\frac{-\sigma_A(\sigma_A\theta\overline{c}-\sigma_B L)}{\sigma_A+\sigma_B}}+(1-\tau q)\left(1-\widetilde{\alpha}+\widetilde{\alpha}e^{\sigma_A L}\right)\left(1-q+qe^{\sigma_A L}\right)^{\frac{-\sigma_A}{\sigma_A+\sigma_B}}}. \tag{6.5}$$

Substituting the definition of $\overline{c}$ for the $\overline{c}$ appearing in the left-hand side of inequality (6.5), setting $\theta = 1$, and then simplifying yield:

$$\begin{aligned} &\tau q e^{\frac{-\sigma_A(\sigma_A\overline{c}-\sigma_B L)}{\sigma_A+\sigma_B}}\left(1-q+qe^{\sigma_A L}\right)^{\frac{\sigma_A}{\sigma_A+\sigma_B}}+(1-\tau q)\left(1-\widetilde{\alpha}+\widetilde{\alpha}e^{\sigma_A L}\right) \\ &< \left[\tau q e^{\frac{\sigma_A\sigma_B}{\sigma_A+\sigma_B}L}+(1-\tau q)\left(1-\widetilde{\alpha}+\widetilde{\alpha}e^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}}\right]^{\frac{\sigma_A+\sigma_B}{\sigma_B}}. \end{aligned} \tag{6.6}$$

Setting $L = 0$ in inequality (6.6) yields:

$$\tau q e^{\frac{-\sigma_A^2 \bar{c}}{\sigma_A + \sigma_B}} + 1 - \tau q < 1, \tag{6.7}$$

which is always satisfied since $\bar{c} > 0$ and $\tau q > 0$. Thus, if $c = \bar{c}$ and $L = 0$, then an equilibrium with No Investigation does not exist. By continuity, the same must be true for some $c > \bar{c}$ and some $L > 0$.□

Acknowledgments

I am deeply indebted to my supervisor Jörgen Weibull for encouragement, constructive criticism, and many useful suggestions. I have also benefited from discussions with and comments from a large number of other people, including Abhijit Banerjee, Tim Besley, Bouwe Dijkstra, Tore Ellingsen, Jacob Gyntelberg, Arye Hillman, Bengt Holmström, Yeongjae Kang, Nippe Lagerlöf, Ariane Lambert-Mogiliansky, Michael Lundholm, Eric Maskin, Motty Perry, Torsten Persson, Thomas Piketty, Christian Schultz, Björn Segendorff, Johan Stennek, Patrik Säfvenblad, Peter Sørensen, Jean Tirole, Karl Wärneryd, and two anonymous referees. Financial support from Jan Wallander's and Tom Hedelius' Research Foundation is gratefully acknowledged.

References

Ainsworth, S., 1993, Regulating lobbyist and interest group influence, Journal of Politics 55, 41-56.

Austen-Smith, D., 1994, Strategic transmission of costly information, Econometrica 62, 955-963.

Austen-Smith, D. and J.R. Wright, 1992, Competitive lobbying for a legislator's vote, Social Choice and Welfare 9, 229-257.

Ball, R., 1995, Interest groups, influence and welfare, Economics and Politics 7, 119-146.

Besley, T. and S. Coate, 1995, An economic model of representative democracy, Working paper 95-02 (Center for Analytical Research in Economics and the Social Sciences, University of Pennsylvania).

Coughlin, P.J., 1992, Probabilistic voting theory (Cambridge University Press, Cambridge).

Crawford, V.P. and J. Sobel, 1982, Strategic information transmission, Econometrica 50, 1431-1451.

Dewatripont, M. and J. Tirole, 1995, Advocates, Mimeo (University of Toulouse).

Grossman, S.J., 1981, The informational role of warranties and private disclosure about product quality, Journal of Law and Economics 24, 462-483.

Laffont, J.-J. and J. Tirole, 1991, The politics of government decision-making: A theory of regulatory capture, Quarterly Journal of Economics 106, 1089-1127.

Laffont, J.-J. and J. Tirole, 1993, A theory of incentives in procurement and regulation (MIT Press, Cambridge).

Lagerlöf, J., 1996, A welfare analysis of strategic information revelation, Working paper no. 109 (Department of Economics, Stockholm School of Economics), presented at 1996 European Public Choice Society Meetings, Tiberias and Bar-Ilan University, Israel.

Lindbeck, A. and J.W. Weibull, 1987, Balanced-budget redistribution as the outcome of political competition, Public Choice 52, 273-297.

Lipman, B.L. and D.J. Seppi, 1995, Robust inference in communication games with partial provability, Journal of Economic Theory 66, 370-405.

Lohmann, S., 1993, A welfare analysis of political action, in: W. Barnett, M. Hinich, and N. Schofield, eds., Political economy: Institutions, competition and representation (Cambridge University Press, Cambridge) 437-461.

Lohmann, S., 1995, Information, access, and contributions: A signaling model of lobbying, Public Choice 85, 267-284.

Milgrom, P.R., 1981, Good news and bad news: Representation theorems and applications, Bell Journal of Economics 12, 380-391.

Milgrom, P. and J. Roberts, 1986, Relying on the information of interested parties, Rand Journal of Economics 17, 18-32.

Okuno-Fujiwara, M., A. Postlewaite, and K. Suzumura, 1990, Strategic information revelation, Review of Economic Studies 57, 25-47.

Potters, J. and F. van Winden, 1992, Lobbying and asymmetric information, Public Choice 74, 269-292.

Potters, J. and F. van Winden, 1993, Models of interest groups: Four different approaches, TRACE discussion paper TI 93-204 (University of Amsterdam), forthcoming in: N. Schofield, ed., Social Choice and Political Economy (Kluwer, Boston).

Rasmusen, E., 1993, Lobbying when the decision maker can acquire independent information, Public Choice 77, 899-913.

Schultz, C., 1996, Political competition and polarization, Mimeo (University of Copenhagen).

Shavell, S., 1994, Acquisition and disclosure of information prior to sale, RAND Journal of Economics 25, 20-36.

Sloof, R., 1995a, Lobbying when the decision maker can acquire independent information: A comment, Discussion paper TI 1-95-233 (Tinbergen Institute), forthcoming in Public Choice.

Sloof, R., 1995b, Competitive lobbying for a legislator's vote: A comment, Mimeo (University of Amsterdam).

Sloof, R. and F. van Winden, 1996, Show them your teeth first: A game-theoretic analysis of lobbying and pressure, Mimeo (University of Amsterdam), presented at 1996 European Public Choice Society Meetings, Tiberias and Bar-Ilan University, Israel.

$\Pi^A_{BM} = \frac{\sigma_B}{\sigma_A}\lambda\Pi^B_{BM}$
$\Pi^B_{BM} = -\frac{1}{\sigma_B}\lambda^{\frac{-\sigma_B}{\sigma_A+\sigma_B}}\left(1-q+qe^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}} e^{\frac{-\sigma_A\sigma_B}{\sigma_A+\sigma_B}(y_A+y_B)}$
$\Pi^G_{BM} = \frac{\sigma_A+\sigma_B}{\sigma_A}\lambda\Pi^B_{BM}$
$\Pi^A_{PR} = \frac{\sigma_B}{\sigma_A}\lambda e^{\sigma_A(1-\theta)c}\Pi^B_{PR}$
$\Pi^B_{PR} = \frac{-\lambda^{\frac{-\sigma_B}{\sigma_A+\sigma_B}}}{\sigma_B} e^{\frac{\sigma_A\sigma_B}{\sigma_A+\sigma_B}(\theta c-y_A-y_B)}\left[\tau q e^{\frac{\sigma_A\sigma_B}{\sigma_A+\sigma_B}L} + (1-\tau q)\left(1-\tilde{\alpha}+\tilde{\alpha}e^{\sigma_A L}\right)^{\frac{\sigma_B}{\sigma_A+\sigma_B}}\right]$
$\Pi^G_{PR} = \frac{\sigma_A+\sigma_B}{\sigma_A}\lambda\Pi^B_{PR}$

Table 1. A's, B's, and G's ex ante payoffs.

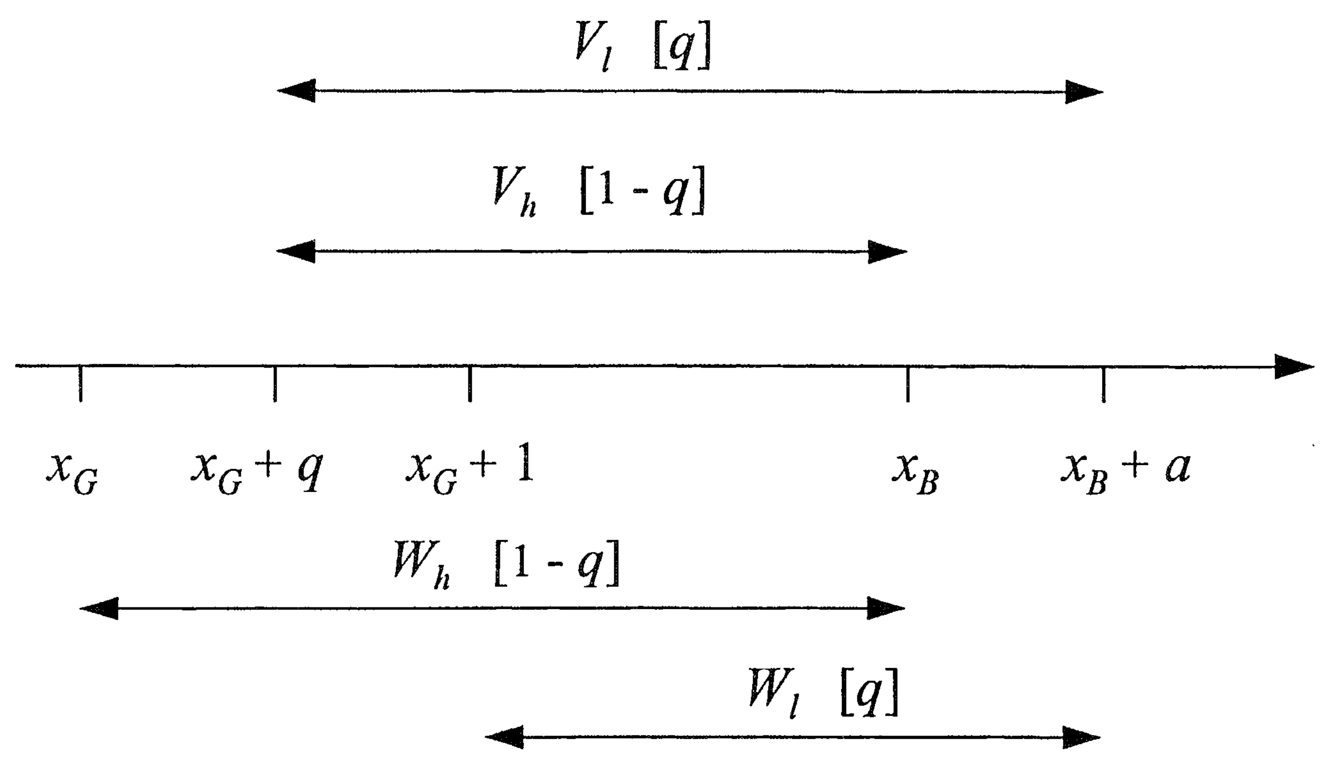


Fig. 1. When does *B* want *G* to be informed?

Chapter 2

Are we better off if our politicians know how the economy works?

Are we better off if our politicians know how the economy works?

Johan Lagerlöf*

University of Rochester
Stockholm School of Economics

First version: August 23, 1997
This version: December 4, 1997

Abstract

This paper concerns public policy and welfare in a society where citizens' preferences over public policy depend, in varying degrees, on some unknown state of the world. That is, people are heterogenous with respect to their responsiveness to the unknown state. Public policy is decided on by a policymaker who is elected among the citizens by majority vote. Given this framework it is asked whether the citizens would be better off if the amount of uncertainty that the policymaker is facing were smaller. Among the results is that those who are sufficiently responsive to the unknown state may be worse off if the variance of the stochastic variable decreases. *//[Doc: av-dog.tex]//*

JEL classification: D69, D78, D89

Keywords: Information acquisition; Welfare; Delegation

* Correspondence to: Johan Lagerlöf, W. Allen Wallis Institute of Political Economy, University of Rochester, 110 Harkness Hall, Rochester, NY 14627, USA. E-mail: jlgf@troi.cc.rochester.edu. I have benefited from helpful discussions with and comments from Jonas Björnerstedt, Yeongjae Kang, Nippe Lagerlöf, Susanne Lohmann, César Martinelli, Björn Persson, David Sundén, Karl Wärneryd, and Jörgen Weibull. Financial support from Jan Wallander and Tom Hedelius' Foundation is gratefully acknowledged.

1. Introduction

This paper is concerned with public policy and welfare in a society where citizens' preferences over public policy depend, in varying degrees, on some unknown state of the world. That is, people are heterogenous with respect to their *responsiveness* to the unknown state: whereas some citizens' ideal public policy depends strongly on the true state, others' ideal policy is much less state dependent. Public policy is decided on by a policymaker who is elected among the citizens by majority vote. Given this setup the questions are posed whether the citizens would be better or worse off ex ante (i) if the policymaker had access to more information about the state of the world and (ii) if the randomness in the state of the world were lower.

One example of a story which fits into the above setup is the following. People have preferences over a public good (roads, say) and a public bad (pollution), and these have different weights in different citizens' utility functions. An elected policymaker decides directly only on the amount of roads. Indirectly, however, this decision affects also the amount of pollution: More roads give rise to more pollution. The exact relationship between the amount of roads and pollution is unknown though. Together, each citizen's utility function and the stochastic relationship between roads and pollution give rise to induced preferences over roads only. Those citizens who care relatively more about the amount of pollution (the "environmentalists") will be more responsive to changes in the unknown stochastic variable.

Another example concerns monetary policy.[1] Citizens have preferences over inflation and employment, although the relative weight on these two issues differ among them. An elected policymaker sets the rate of inflation directly; the level of employment is affected indirectly if the actual inflation differs from the expected inflation. The level of employment is also affected by some external shock, the exact magnitude of which is unknown by the policymaker when setting the inflation. The stochastic relationship between inflation and employment (the expectations-augmented Phillips curve) together with a citizen's preferences over inflation and employment give rise to induced preferences over inflation only; and for those citizens for which employment is relatively more important (for those less "conservative" in Rogoff's (1985) terminology), the ideal inflation level will

[1] There is an extensive literature on credibility in monetary policy which uses a modeling framework that is compatible with this example and the analysis in Section 5 of the paper. This literature was initiated by Kydland and Prescott (1977) and Barro and Gordon (1983), and it is surveyed by e.g. Persson and Tabellini (1990).

depend to a greater extent on the stochastic variable.

Concerning the two examples above one may ask oneself two questions. First, would all people in the economy be better off (ex ante) if the policymaker had more information about the relationship between roads and pollution respectively the expectations-augmented Phillips curve? The analysis of this paper shows that this is not the case. Given that "more information" is understood as a more informative signal about the realization of the stochastic variable,[2] only those members of the society which are sufficiently responsive to the stochastic variable gain from the policymaker's having more information. Those who are not sufficiently responsive — in the sense that they only to a small extent care about pollution respectively inflation — would be worse off if the policymaker had access to a more informative signal.

Second, would all people in the economy be better off (ex ante) if the randomness in the relationship between roads and pollution respectively inflation and employment decreased? The analysis of the paper shows that if a decrease in randomness is understood as a smaller variance of the stochastic variable and if the policymaker can improve upon the informativeness of the signal that he observes by making a greater effort, then those who are sufficiently responsive (i.e., the environmentalists respectively the "liberals") may be worse off from a lower variance. The reason for this is that the environmentalists and the liberals want the policymaker to make a great effort, thereby getting access to a more informative signal. However, a lower variance of the stochastic variable induces the policymaker to make a smaller effort.

It is important to note that the question whether people in the economy would be better off if the policymaker had access to more information is quite different from the question whether additional information would be good for the policymaker himself. It is well known that a player in a non-zero sum game can be worse off from having more information; for examples of this phenomenon in a political framework, see Reed (1989) or, in a Cournot duopoly setting, Sakai (1985). This finding is perhaps more surprising than the result that others than the policymaker may be worse off. Nevertheless, the latter result may be at least as important. One example of a field where this kind of result is of significance is the literature on informational lobbying (Austen-Smith and Wright, 1992; Pot-

[2] More precisely, in the analysis that follows, "more information" will be understood as an increase in the square of the correlation coefficient between the signal and the unknown stochastic variable.

ters and van Winden, 1992).[3] This literature takes as its point of departure that lobbyists have access to information that is relevant to the politician in his policy making. Hence, by strategically transmitting this information to the policymaker, the lobbyists may be able to influence public policy. Typically, in the equilibria of the models in this literature, at least some information is transmitted to the policymaker. A welfare analysis of the lobbyist's opportunity to lobby then amounts to asking whether the policymaker's having access to this information induces him to make decisions which are better to people in the society.[4]

Besides studying the welfare effects of a policymaker's being better informed and of a decrease in the randomness of the state of the world, this paper also considers another question, namely the relation between the policymaker's and the electorate's degree of responsiveness. One of the few papers in the literature which explicitly models this kind of heterogeneity is Schultz (1996). In his model the electorate is homogenous with regard to the degree of responsiveness, but the two political parties differ from the electorate in that they are *less* responsive.[5] This difference between the electorate and the parties is exogenous to Schultz's analysis. However, Schultz shows that if such a difference exists and if one considers a dynamic environment with incomplete information about the incumbent party's preferences, then more polarization (i.e., a greater difference between the parties' preferences) makes the electorate worse off. Since this result is driven by the assumption that candidates are less responsive than the electorate, one may wonder what could give rise to such a difference.

If public policy concerns the rate of inflation, then one reason for the electorate to delegate the formulation of public policy to a policymaker that is less responsive than the median voter can be found in the existing literature on credibility in monetary policy (see the references in footnote 1 above). Rogoff (1985) shows that the inflationary bias that arises in the models in this literature may be mitigated if the task of conducting monetary policy is delegated to someone more conservative. However, a policymaker that is more conservative in the sense of Rogoff is also less responsive in the sense of Schultz. Hence, if monetary policy is conducted by a politician who is elected by the electorate, then the difference between the politician and the electorate that is postulated by Schultz should arise endogenously.[6]

[3] Two recent surveys can be found in Austen-Smith (1997) and Sloof (1997).

[4] This question is also studied in Lagerlöf (1997). The present paper extends that analysis.

[5] Similar assumptions are made in Martinelli (1997) and Martinelli and Matsui (1997).

[6] In Rogoff's model it is a benevolent government — not an electorate — that appoints the

The present paper models this mechanism together with another mechanism which has not, to my knowledge, been studied in the previous literature and which may also lead to a difference in the degree of responsiveness between the policymaker and the electorate. However, the other mechanism leads the policymaker to be *more* responsive than the median voter. The mechanism works like this. Suppose that the policymaker, after having taken office but before having decided on public policy, observes a noisy signal about the realization of the stochastic variable. Moreover, by making a costly effort, the policymaker can improve upon the informativeness of the signal. If so, voters may have an incentive to delegate the task of deciding on public policy to a policymaker who is more responsive since such a person has a greater incentive to make an effort.

The remainder of the paper is organized as follows. In Sections 2, 3, and 4, I consider a model compatible with the environmental story told above. In Section 2, the basic model, where the signal's informativeness is exogenous, is presented. In Section 3, this model is analyzed and the first results are stated. Section 4 studies an extension of the model where the signal's informativeness is endogenous. In Section 5 I consider a slightly different model which is compatible with the monetary policy story told above. Section 6 briefly summarizes and concludes. Proofs are found in an appendix.

2. The basic model

Consider a society with a continuum of citizens each having preferences over two public goods, provided in quantities π and x. Citizen i's preferences are described by the von Neumann–Morgenstern utility function

$$U_i(\pi, x) = -(\pi - \overline{\pi})^2 - \lambda_i (x - \overline{x})^2, \tag{0.1}$$

where $\overline{\pi}$, $\overline{x}$, and λ_i are fixed parameters. The citizens differ from each other only with regard to the parameter λ_i. The distribution of λ_i among the citizens is described by a cumulative probability distribution function G whith support $\Re_+$. The (finite) mean and the median of λ_i are denoted by $\overline{\lambda}$ and λ_m, respectively.

Public policy is decided on by a policymaker. The policymaker can control only π. However, there is a stochastic relationship between π and x, given by

$$x = \beta\pi - \varepsilon. \tag{0.2}$$

policymaker/central banker. For models where the policymaker deciding on monetary policy is elected, see Alesina and Grilli (1992), Björnerstedt (1995), and Waller and Walsh (1996).

Here $\beta > 0$ is a fixed parameter and ε is a stochastic variable with zero mean. The model is compatible with the first example given in the introduction.[7] That is, we may think of π as the amount of roads in a country, and x as the amount of pollution caused by the traffic on these roads (or perhaps rather the adverse environmental effects of the pollution). Everybody has some ideal amount of roads, $\overline{\pi}$, and some ideal amount of pollution, $\overline{x}$. The uncertainty as to the exact relationship between the amount of roads and pollution may be due to the fact that the technology giving rise to the relationship is not perfectly known, or to the fact that the amount of pollution also depends on weather conditions which vary in an unpredictable manner.

Substituting (0.2) into (0.1) yields citizen i's induced preferences over π only:

$$\begin{aligned} u_i(\pi,\varepsilon) &= -(\pi-\overline{\pi})^2 - \lambda_i(\beta\pi-\varepsilon-\overline{x})^2 \\ &= -\left(1+\lambda_i\beta^2\right)\left[\pi-\psi(\lambda_i)-\varphi(\lambda_i)\varepsilon\right]^2 - \frac{\lambda_i(\beta\overline{\pi}+\varepsilon+\overline{x})^2}{1+\lambda_i\beta^2}, \end{aligned} \tag{0.3}$$

where

$$\psi(\lambda_i) = \frac{\overline{\pi}+\lambda_i\beta\overline{x}}{1+\lambda_i\beta^2} \tag{0.4}$$

and

$$\varphi(\lambda_i) = \frac{\lambda_i\beta}{1+\lambda_i\beta^2}. \tag{0.5}$$

This means that if ε were known, citizen i would like the policymaker to set π equal to

$$\widehat{\pi} = \psi(\lambda_i) + \varphi(\lambda_i)\varepsilon. \tag{0.6}$$

Hence, since $\varphi(0) = 0$ and $\varphi' > 0$, the parameter λ_i measures how responsive a citizen is to changes in ε. Someone who has a low λ_i (i.e., someone who does not care much about pollution) would like the policymaker to make $\widehat{\pi}$ contingent on ε to a lesser degree than someone for whom λ_i is large. In the following, the parameter λ_i will often be called citizen i's responsiveness parameter.

The policymaker is elected among the citizens by majority vote. More specifically, in a *political equilibrium*, the policymaker is a citizen having a responsiveness parameter λ_i such that he cannot be beaten in a pair-wise comparison when each citizen votes for the one of the two candidates that gives him the highest expected

[7] The preferences in equation (0.1) are also compatible with the second example (the one on monetary policy). However, the stochastic relationship between π and x, given by equation (0.2), does not fit that story.

utility. Hence, like all other citizens, the policymaker has preferences according to equation (0.1), and these will govern his choice of π; he cannot commit himself to any electoral platform other than his ideal policy.

Concerning the informational structure and the timing of events, the following is assumed. First the policymaker is elected. The stochastic variable ε cannot be observed by anyone, neither before nor after the elections. However, after having taken office, the policymaker observes a signal s, which is correlated with ε. Then the policymaker decides on π. All citizens' preferences are known by all citizens.

Let F be the joint cumulative distribution function of ε and s, with density f. The following notation is used:

$$\mu_s = E(s), \tag{0.7}$$

$$\sigma^2 = Var(\varepsilon), \tag{0.8}$$

$$\sigma_s^2 = Var(s), \tag{0.9}$$

and

$$\rho = \frac{Cov(\varepsilon, s)}{\sigma\sigma_s}. \tag{0.10}$$

(Recall that the expected value of ε equals zero, $E(\varepsilon) = 0$.) $\rho \in [-1, 1]$ is thus the correlation coefficient between s and ε.

The policymaker is assumed to be a Bayesian updater. Thus, after having observed the signal s, the policymaker's beliefs about ε are described by the conditional density function $f(\varepsilon \mid s)$, defined by

$$f(\varepsilon \mid s) = \frac{f(\varepsilon, s)}{f(s)}, \tag{0.11}$$

where $f(s) = \int_{-\infty}^{\infty} f(\varepsilon, s)\, d\varepsilon$ is the marginal density of s. The conditional expectation function is defined by $E(\varepsilon \mid s) = \int \varepsilon f(\varepsilon \mid s)\, d\varepsilon$. Assume that F is such that ε has *linear regression* with regard to s, i.e., that $E(\varepsilon \mid s)$ is a linear (affine) function of s.[8] It is well known that if ε has linear regression with regard to s (and if $E(\varepsilon) = 0$), then

$$E(\varepsilon \mid s) = \rho\frac{\sigma}{\sigma_s}(s - \mu_s). \tag{0.12}$$

[8]For instance, a bivariate normal distribution has this property.

3. Beginning the analysis

Let us denote the policymaker's responsiveness parameter by λ_g (where g stands for government). At the last stage, conditional on his having observed the signal s, the policymaker will implement the policy π which maximizes his expected utility:

$$\max_{\pi} \int u_g(\pi, \varepsilon) f(\varepsilon \mid s)\, d\varepsilon. \tag{0.13}$$

The unique solution to this problem is given by

$$\pi_g^* = \psi(\lambda_g) + \varphi(\lambda_g) E(\varepsilon \mid s). \tag{0.14}$$

Now consider a citizen/voter. At the time of the elections, this person only knows the prior distribution of s and ε. However, he anticipates that a policymaker with responsiveness parameter λ_g will set π equal to π_g^*. Hence, citizen i's expected utility at the time of the elections, denoted by Eu_i, may be written as

$$\begin{aligned} Eu_i &= \int\int u_i\left(\pi_g^*, \varepsilon\right) f(\varepsilon, s)\, d\varepsilon ds \\ &= -\left(1+\lambda_i\beta^2\right)\left\{\psi(\lambda_g)\left[\psi(\lambda_g) - 2\psi(\lambda_i)\right]\right. \\ &\quad \left. +\rho^2\sigma^2\varphi(\lambda_g)\left[\varphi(\lambda_g) - 2\varphi(\lambda_i)\right]\right\} - \overline{\pi}^2 - \lambda_i\left(\overline{x}^2+\sigma^2\right). \end{aligned} \tag{0.15}$$

The expression after the second equality sign in equation (0.15) was obtained by using equations (0.3), (0.12), (0.14), and by carrying out some algebra.

Eu_i represents citizen i's preferences over a potential policymaker. These potential policymakers differ from each other along only one dimension, $\lambda_g \in [0, \infty)$. Moreover, in the proof of Lemma 1 below it is shown that Eu_i is single peaked in λ_g. Hence, we can invoke the median voter theorem (see e.g. Mueller, 1989), which states that if those two conditions (i.e., one dimension and single-peakedness) are met then the median voter's favorite policymaker cannot lose under majority rule. This means that, in a political equilibrium, the policymaker will be a citizen preferred by the median citizen/voter. Not surprisingly, the responsiveness parameter of this preferred citizen equals the median voter's, $\lambda_g = \lambda_m$; there is no reason for any member of the electorate to delegate the task of deciding on public policy to someone with other preferences than the member himself.

Lemma 1. *The policymaker's responsiveness parameter is the same as the median citizen's, $\lambda_g = \lambda_m$.*

Let us now investigate whether members of the society would be better off if the policymaker had access to a more informative signal about the stochastic variable ε. The welfare evaluation will be made ex ante; that is, I will consider citizen i's expected utility, as measured by Eu_i in equation (0.15) (with $\lambda_g = \lambda_m$). The expression "more informative signal" will be understood as an increase in ρ^2.

Let $\widetilde{\lambda}$ be defined by

$$\widetilde{\lambda} = \frac{\lambda_m}{2 + \lambda_m \beta^2}. \tag{0.16}$$

Proposition 1. *An increase in ρ^2 benefits those with $\lambda_i > \widetilde{\lambda}$ and makes those with $\lambda_i < \widetilde{\lambda}$ worse off (i.e., $\frac{\partial Eu_i}{\partial \rho^2} |_{\lambda_g = \lambda_m} \gtreqless 0$ as $\lambda_i \gtreqless \widetilde{\lambda}$).*

Accordingly, those members of the electorate who have a sufficiently low responsiveness parameter λ_i are *worse* off if the policymaker has access to better information about the relationship between π and x, in the sense that ρ^2 is larger.[9] Before looking at the intuition for this result, let us consider the question whether a *majority* of the citizens may be worse off from an increase in ρ^2. Since $\widetilde{\lambda} < \lambda_m/2$ (see equation 0.16), it follows immediately that the answer to this question is no: Everyone with a responsiveness parameter $\lambda_i \in [\lambda_m/2, \infty)$ is strictly better off from a larger ρ^2, and this group of citizens form a majority.

However, it may be that a social welfare function that assigns an equal weight to the expected utility of all citizens is decreasing in ρ^2. Let EW be defined by

$$EW = \int_0^\infty Eu_i dG(\lambda_i). \tag{0.17}$$

Since the expression for Eu_i in equation (0.15) is a linear (affine) function of λ_i (cf. the first line of equation (0.3)), EW is obtained by simply substituting $\overline{\lambda}$ (i.e., the responsiveness parameter of the average citizen) for λ_i in equation (0.15):

$$\begin{aligned} EW \;=\; & -\left(1 + \overline{\lambda}\beta^2\right) \left\{ \psi(\lambda_g) \left[\psi(\lambda_g) - 2\psi\left(\overline{\lambda}\right) \right] \right. \\ & \left. + \rho^2 \sigma^2 \varphi(\lambda_g) \left[\varphi(\lambda_g) - 2\varphi\left(\overline{\lambda}\right) \right] \right\} - \overline{\pi}^2 - \overline{\lambda}\left(\overline{x}^2 + \sigma^2\right). \end{aligned} \tag{0.18}$$

[9] The result in Proposition 1 is related to a result in Lagerlöf (1997). In that paper, however, the identity of the policymaker is exogenous, and the differences in responsiveness between citizens is not — as in this paper — derived from differences in the relative weights on two policy issues and the stochastic relationship between them. Also, in Lagerlöf (1997) the stochastic variable has a Bernoulli distribution.

For EW to be decreasing in ρ^2, the distribution G must be sufficiently skewed to the right, so that λ_m is to a sufficient extent greater than $\overline{\lambda}$. If so, it might be that $\overline{\lambda} < \widetilde{\lambda}$.

In order to understand the intuition behind the result that those citizens having a low responsiveness parameter are worse off if ρ^2 is larger, let us consider the extreme case where $\lambda_i = 0$. Such a citizen only cares about π, and he does not want the policy to be conditioned on ε at all. Instead, his ideal policy always equals $\overline{\pi}$ (cf. equation (0.6)). Now consider a policymaker having a responsiveness parameter $\lambda_g > 0$. If this policymaker can observe a signal about the realization of ε, then he will condition his decision on the signal, and thus make the decision

$$\pi_g^* = \psi(\lambda_g) + \varphi(\lambda_g) E(\varepsilon \mid s). \tag{0.19}$$

From an ex ante perspective this means that the decision will vary, since the citizen observes the signal only ex post. If the policymaker could not observe the signal, then he would make the decision

$$\widetilde{\pi}_g^* = \psi(\lambda_g). \tag{0.20}$$

Clearly this decision will not vary.

Let us decompose the citizen's gain from the policymaker's not having access to the signal into two parts: (i) the gain the citizen would obtain if he were risk neutral and (ii) the gain that is due to the citizen's being risk averse. If the citizen were risk neutral, he would only care about the expected policy. However, it is easy to see that the expected policy is the same regardless of the policymaker's having access to the signal or not:[10]

$$E_s(\pi_g^*) = E_s[\psi(\lambda_g) + \varphi(\lambda_g) E(\varepsilon \mid s)] = \psi(\lambda_g). \tag{0.21}$$

Hence, the gain the citizen would obtain if he were risk neutral equals zero, and his total gain from the policymaker's not having access to the signal must exclusively be attributed to the citizen's being risk averse. But the citizen's being risk averse manifests itself in his not wanting any variation in the policymaker's decision. Thus the citizen's gain from the policymaker's not having access to the signal is always positive. The same is true for citizens having a responsiveness parameter λ_i that is strictly positive but still relatively small (smaller than $\widetilde{\lambda}$).[11]

[10] This is due to the quadratic functional form.

[11] The intuition for the result in Proposition 1 is related to the intuition for a result in Freixas

Before finishing this section, let us consider the question whether all citizens would be better off ex ante if σ^2, the variance of ε, were lower. Not surprisingly, it turns out that this is indeed the case. However, in the next section the model will be expanded by making the informativeness of the signal that the policymaker observes endogenous, and in that model a smaller variance may be harmful. This finding will be easier to understand in light of the result stated in the following observation, which assumes that the signal's informativeness is exogenous.

Observation 1. *All citizens are always better off from a lower variance of* ε *(i.e.,* $\frac{\partial E u_i}{\partial \sigma^2} |_{\lambda_g=\lambda_m} < 0$ *for all* λ_i*).*

4. The signal's informativeness being endogenous

In this section it is assumed that the policymaker can make a costly effort and thereby improve upon the informativeness of the signal that he observes. The informational structure and the timing of events in this extension of the model is as follows. First the policymaker is elected. After having taken office, the policymaker first decides on an effort level e. Then he observes the signal s, which is correlated with ε. Finally the policymaker decides on π. The stochastic variable ε cannot be directly observed by anyone, neither before nor after the elections.

It is assumed that $e \equiv \rho^2$, where as before ρ is the correlation coefficient between s and ε; hence $e \in [0,1]$. Thus, by making a greater effort, the policymaker can improve upon the informativeness of the signal. However, making an effort is costly for the policymaker; the disutility that he incurs from exerting effort level e equals $C(e)$, where $C' > 0$ and $C'' > 0$, with $C'(0) = 0$.

---

and Kihlstrom (1984). They consider a situation in which a patient must choose a doctor in the face of imperfect information about the distribution of service quality across doctors. In particular they study the effect of risk aversion on demand for information about this distribution. They write (p. 93): "On this issue, intuition is inconclusive since it suggests that the final effect is a combination of two conflicting effects. On the one hand, more risk-averse decision-makers should have a stronger preference for the *ex post* reduction in uncertainty accomplished by acquiring information. But uncertainty is reduced only *ex post*, i.e. only after the informative message has been received. When the decision to buy information is made, the buyer does not yet know whether he will receive good news or bad when the information arrives. Thus, *ex ante*, the returns to information are uncertain, and more risk averse buyers should be less willing to accept the risks associated with its acquisition." Freixas and Kihlstrom find that, in their model, an increase in the degree of risk aversion unambiguously reduces information demand.

Let $E\widetilde{u}_g$ denote the policymaker's expected utility at the stage where he is to decide on the effort level e. It follows from the expression for Eu_i in equation (0.15) that $E\widetilde{u}_g$ may be written as

$$E\widetilde{u}_g = \left(1+\lambda_g\beta^2\right)\left[\psi^2\left(\lambda_g\right)+e\sigma^2\varphi^2\left(\lambda_g\right)\right]-\overline{\pi}^2-\lambda_g\left(\overline{x}^2+\sigma^2\right)-C\left(e\right), \quad (0.22)$$

where the last term is the postulated cost of information acquisition.[12] The policymaker solves the problem of maximizing $E\widetilde{u}_g$ in equation (0.22) with respect to e, subject to the constraint $e\in[0,1]$. Throughout I shall assume that this problem has an interior solution.[13] This interior solution, e^*, is implicitly defined by

$$C'\left(e^*\right) \equiv \frac{\beta^2\sigma^2\lambda_g^2}{1+\lambda_g\beta^2}. \quad (0.23)$$

Note for future use that

$$\frac{\partial e^*}{\partial\lambda_g} = \frac{\beta^2\sigma^2\lambda_g\left(2+\lambda_g\beta^2\right)}{C''\left(e^*\right)\left(1+\lambda_g\beta^2\right)^2} > 0 \quad (0.24)$$

and

$$\frac{\partial e^*}{\partial\sigma^2} = \frac{\beta^2\lambda_g^2}{C''\left(e^*\right)\left(1+\lambda_g\beta^2\right)} > 0. \quad (0.25)$$

That is, as expected, a policymaker who cares more about the environment (has a larger λ_g) makes a greater effort to learn about how much the environment is adversely affected by building roads. Similarly, a larger variance of the stochastic variable also induces the policymaker to make a greater effort.

Let η be defined by

$$\eta = \frac{C''\left(e^*\right)e^*}{C'\left(e^*\right)}, \quad (0.26)$$

[12] Note that the gross value of information (i.e., $E\widetilde{u}_g$ if not counting the cost $C(e)$) is linear in $e=\rho^2$. However, had we assumed that $e=\rho$, then the gross value of information would have been a convex function of e. This phenomenon is closely related to a result in Radner and Stiglitz (1984). They show that for an important class of decision problems, the value of information is nonconcave. In particular, see their first example where they consider a linear prediction problem.

[13] The problem has an interior solution if the Inada condition $\lim_{e\to 1} C'(e)=\infty$ holds or if this limit is finite and $\lambda_g<\lambda_c$, where λ_c is defined by $\lambda_c^2\beta^2\sigma^2 \equiv \left(1+\lambda_c\beta^2\right)C'(1)$. An example of such a cost function is $C(e)=\left(1-\sqrt{1-e}\right)^2$. The reason why I do not simply assume that this Inada condition holds is that, when studying some examples later in this section, it will be convenient to let $C(e)=e^a$ for $a>1$.

and let $Z = e^* \left(1 + \frac{1}{\eta}\right)$. By using equations (0.23) and (0.25), one may show that

$$\frac{1}{\eta} = \frac{\partial e^*}{\partial \sigma^2} \frac{\sigma^2}{e^*}. \tag{0.27}$$

Hence, η not only measures the curvature of the cost function C, but is also equal to the inverse of the elasticity of information demand with respect to σ^2.

Now consider again a citizen/voter with responsiveness parameter λ_i. His expected utility if the policymaker has a responsiveness parameter λ_g and accordingly exerts effort $e^*(\lambda_g)$ is denoted by $E\widetilde{u}_i$, and it is obtained by simply substituting e^* for ρ^2 in equation (0.15):

$$\begin{aligned} E\widetilde{u}_i &= -\left(1 + \lambda_i \beta^2\right) \left[\psi\left(\lambda_g\right) \left[\psi\left(\lambda_g\right) - 2\psi\left(\lambda_i\right)\right]\right. \\ &\left. + e^* \sigma^2 \varphi\left(\lambda_g\right) \left[\varphi\left(\lambda_g\right) - 2\varphi\left(\lambda_i\right)\right]\right] - \overline{\pi}^2 - \lambda_i \left(\overline{x}^2 + \sigma^2\right). \end{aligned} \tag{0.28}$$

$E\widetilde{u}_i$ thus represents i's induced preferences over policymakers. Again, for the median voter theorem to hold, these preferences must be single peaked in λ_g. In the Appendix I show that sufficient conditions for this are that $C(e) = e^a$, $a \in \left(1, \frac{3}{2}\right)$, $\overline{x} \neq \beta\overline{\pi}$, and that σ^2 is sufficiently close to zero. Here, I will confine myself with showing that when $E\widetilde{u}_i$ is single peaked in λ_g, then the policymaker's responsiveness parameter does not necessarily equal the median voter's.

To see this, let us differentiate $E\widetilde{u}_i$ with respect to λ_g and evaluate at $\lambda_i = \lambda_m$:

$$\begin{aligned} \frac{\partial E\widetilde{u}_i}{\partial \lambda_g} \mid_{\lambda_i = \lambda_m} &= -2\left(1 + \lambda_m \beta^2\right) \left[\psi'\left(\lambda_g\right) \left[\psi\left(\lambda_g\right) - \psi\left(\lambda_m\right)\right]\right. \\ &\left. + e^* \sigma^2 \varphi'\left(\lambda_g\right) \left[\varphi\left(\lambda_g\right) - \varphi\left(\lambda_m\right)\right]\right] - \left(1 + \lambda_m \beta^2\right) \sigma^2 \varphi\left(\lambda_g\right) \left[\varphi\left(\lambda_g\right) - 2\varphi\left(\lambda_m\right)\right] \frac{\partial e^*}{\partial \lambda_g}. \end{aligned} \tag{0.29}$$

When $E\widetilde{u}_i$ is single peaked, then the median voter theorem applies, and in a political equilibrium the policymaker will be the favorite of the median voter. That is, λ_g will be such that the right-hand side of equation (0.29) equals zero. Now suppose that $\frac{\partial e^*}{\partial \lambda_g} = 0$. It follows immediately from equation (0.29) that then the result from Lemma 1 is reobtained, $\lambda_g = \lambda_m$. However, if $\frac{\partial e^*}{\partial \lambda_g} > 0$, then we must have $\lambda_g > \lambda_m$. This can seen by evaluating (0.29) at $\lambda_g = \lambda_m$:

$$\frac{\partial E\widetilde{u}_i}{\partial \lambda_g} \mid_{\lambda_i = \lambda_g = \lambda_m} = \left(1 + \lambda_m \beta^2\right) \sigma^2 \left[\varphi\left(\lambda_m\right)\right]^2 \frac{\partial e^*}{\partial \lambda_g} \mid_{\lambda_g = \lambda_m}. \tag{0.30}$$

Since this expression is strictly positive, it must be that $\lambda_g > \lambda_m$. The intuition for this result is clear. A policymaker who cares more about the environment will make a greater effort finding information about the environmental effects of building roads, and it will be in the median voter's interest that the policymaker has access to such information. Thus, the median voter can gain by delegating the task of deciding on public policy to somebody that cares more about the environment than himself.[14]

Let us now turn to the question whether a citizen would be better or worse off if σ^2, the variance of ε, were smaller.

Proposition 2. *Suppose that $2\beta Z\varphi(\lambda_g) \leq 1$. Then a decrease in σ^2 (strictly) benefits all citizens. Suppose that $2\beta Z\varphi(\lambda_g) > 1$. Then a decrease in σ^2 (strictly) benefits citizen i if and only if*

$$\varphi(\lambda_i) < \frac{\beta Z[\varphi(\lambda_g)]^2}{2\beta Z\varphi(\lambda_g) - 1}. \tag{0.31}$$

Inequality (0.31) does not need to hold when the condition $2\beta Z\varphi(\lambda_g) > 1$ is met. That is, it may be that a citizen is *worse* off if the variance of ε is smaller. To illustrate this I will consider two numerical examples. In both of them it is assumed that $C(e) = e^a$, for $a > 1$. This implies that

$$e^* = \left[\frac{\beta^2\sigma^2\lambda_g^2}{a(1+\lambda_g\beta^2)}\right]^{\frac{1}{a-1}} \tag{0.32}$$

and that

$$Z = \frac{ae^*}{a-1}. \tag{0.33}$$

Now consider the first example.

Example 1. $\beta = \lambda_g = 1$, $a = \frac{3}{2}$, *and* $\sigma^2 = \frac{5}{2}$.

Given the parameter values specified in Example 1, we get $e^* = \frac{25}{36}$, $Z = \frac{25}{12}$, $\varphi(\lambda_g) = \frac{1}{2}$, and $\varphi(\lambda_i) = \frac{\lambda_i}{1+\lambda_i}$. Hence, $2\beta Z\varphi(\lambda_g) = \frac{25}{12} > 1$; and condition (0.31) now becomes

$$\frac{\lambda_i}{1+\lambda_i} < \frac{25}{52} \Leftrightarrow \lambda_i < \frac{25}{27}. \tag{0.34}$$

[14]For another example of strategic delegation in a political context, see Persson and Tabellini (1994).

Thus, all citizens with a responsiveness parameter larger than $\lambda_i = \frac{25}{27}$ are strictly worse off if the variance decreases. In order to understand the intuition for this result, let us decompose the total welfare effect of a decrease in the variance into two parts: the effect on welfare which arises in the hypothetical case that the signal's informativeness is given; and the effect that is due to the informativeness actually being chosen by the policymaker. In Section 3 (Observation 1) we saw that the first part always is positive. Concerning the second effect, note that if the variance decreases, this will induce the policymaker to make a smaller effort and thus get access to a less informative signal. This comparative statics result follows from equation (0.25). From Section 3 (Proposition 1), however, we know that, everything else equal, those citizens having a sufficiently large responsiveness parameter are worse off from a less informative signal. Hence, for those citizens, the second is effect is negative. The algebra shows that the second effect may in fact be stronger than the first effect, making the most responsive citizens worse off from a lower variance of the stochastic variable.

Example 1 shows that also a citizen with the same responsiveness parameter as the policymaker, $\lambda_i = \lambda_g = 1$, may be worse off from a lower variance. My second example helps us understand what is required for this particular result to obtain.

Example 2. $\lambda_i = \lambda_g$.

Condition (0.31) now becomes

$$Z < \frac{1}{\lambda_i \beta^2} + 1. \tag{0.35}$$

Thus, a necessary condition for a voter with the same responsiveness parameter as the policymaker to be worse off from a lower variance is that $Z > 1$.[15] This highlights the point that essential for our result that some citizens may be worse off from a lower variance is that the magnitude of $1/\eta$, the policymaker's elasticity of information demand with respect to σ^2, is sufficiently great. This is in line with our intuition: The reason why a larger σ^2 may be good is that it induces the policymaker to make a greater effort.

[15] The condition $2\beta Z\varphi(\lambda_g) > 1$ is automatically satisfied if $Z > \frac{1}{\lambda_i \beta^2} + 1$, since, if $\lambda_g = \lambda_i$, the former condition can be rewritten as $2Z > \frac{1}{\lambda_i \beta^2} + 1$.

5. Monetary policy with rational expectations

In this section I consider a model of credibility in monetary policy. The model is only slightly different from the one described and analyzed in the preceding sections. It is also very similar to many models in the existing literature on monetary policy; see the references in footnote 1, in particular Chapter 2 of Persson and Tabellini (1990).

Again there is a continuum of citizens each having preferences over a public good and a public bad, provided in quantities x respectively π. π is now interpreted as the rate of inflation and x as the level of employment. These preferences are described by the von Neumann-Morgenstern utility function

$$V_i(\pi, x) = -\pi^2 - \lambda_i (x - \overline{x})^2. \tag{0.36}$$

That is, citizen i's utility is decreasing in the rate of inflation and decreasing in the deviation from an employment goal $\overline{x} > 0$. The parameter λ_i is as before a weight. The policymaker can control only π. However, by setting the inflation rate different from the expected rate of inflation, π^e, the policymaker can indirectly affect the level of employment. That is, the following expectations-augmented Phillips curve is assumed:

$$x = \beta(\pi - \pi^e) - \varepsilon, \tag{0.37}$$

where as before $\beta > 0$ is a fixed parameter and ε is a stochastic variable with zero mean. The policymaker observes only a signal s, correlated with ε. The expected rate of inflation, π^e, is given by

$$\pi^e = E_{\varepsilon,s}(\pi); \tag{0.38}$$

that is, π^e equals the expected value of the actual rate of inflation at the stage where only the prior distribution of ε and s is known. Equations (0.37) and (0.38) and the fact that the policymaker but not the wage setters can observe the signal s imply that the policymaker can stabilize employment by choosing to "surprise inflate."

All other notation and model features are identical to the model in Section 2. In particular, the citizens have different weights λ_i. The timing is also the same. That is, first the policymaker is elected, then he observes s, and finally he decides on π. Hence, the main difference between the model described in this section and the one in Section 2 is the form of the stochastic relationship between x and π, given by equations (0.37) and (0.2) respectively.[16]

[16] The only other differences are that in this section I have set $\overline{\pi} = 0$ and I require that $\overline{x} > 0$.

Substituting (0.37) into (0.36) yields

$$\begin{aligned} v_i(\pi,\varepsilon) &= -\pi^2 - \lambda_i\left(\beta\left(\pi-\pi^e\right)-\varepsilon-\overline{x}\right)^2 \\ &= -\left(1+\lambda_i\beta^2\right)\left[\pi-\left(\beta\pi^e+\varepsilon+\overline{x}\right)\varphi\left(\lambda_i\right)\right]^2 - \frac{\lambda_i\left(\beta\pi^e+\varepsilon+\overline{x}\right)^2}{1+\lambda_i\beta^2} \end{aligned} \tag{0.39}$$

where $\varphi(\lambda_i)$ is as defined in equation (0.5). Hence, exactly as in the model in Section 2, the parameter λ_i measures i's responsiveness to changes in the stochastic variable ε.

5.1. Analysis

Let us again denote the policymaker's responsiveness parameter by λ_g. At the last stage, the policymaker will implement the policy π which maximizes his expected utility conditional on his having observed the signal s:

$$\max_{\pi} \int v_g(\pi,\varepsilon) f(\varepsilon \mid s)\, d\varepsilon. \tag{0.40}$$

Taking the first-order condition of this problem and then solving for π yields

$$\pi = \frac{\lambda_g\beta\left[\beta\pi^e + E(\varepsilon \mid s) + \overline{x}\right]}{1+\lambda_g\beta^2}. \tag{0.41}$$

The expected rate of inflation is obtained by taking expectations with respect to s of both sides of equation (0.41), using the fact that $E_s(E(\varepsilon \mid s)) = E(\varepsilon) = 0$, and then solving for π^e. Doing this yields

$$\pi^e = \lambda_g\beta\overline{x}. \tag{0.42}$$

Substituting this expression for π^e into equation (0.41) in turn yields

$$\pi_g^* = \lambda_g\beta\overline{x} + \varphi(\lambda_g) E(\varepsilon \mid s). \tag{0.43}$$

That is, on average, the equilibrium rate of inflation equals $\lambda_g\beta\overline{x}$, which typically is greater than zero — the ideal level according to equation (0.36). This "inflationary bias" arises because an average inflation rate of zero is not credible (or time consistent). For at $\pi = 0$, the marginal benefit of surprise inflation exceeds the marginal cost of inflation. For the marginal cost of inflation just to balance the marginal gain from an increase in employment, it must be that the average inflation equals $\lambda_g \beta \overline{x}$. Thus, the zero rate of inflation would not be time inconsistent if the employment goal were equal to the "natural" rate of employment, normalized to zero in equation (0.37).

The inflationary bias is also greater the greater is the policymaker's responsiveness parameter λ_g. Thus, by electing a policymaker having a zero responsiveness parameter, the electorate may eliminate the inflationary bias. However, such a policymaker would not stabilize employment. Hence, the optimal trade-off is to delegate the task of conducting monetary policy to a policymaker having a responsiveness parameter that is positive but smaller than one's own. This was first shown by Rogoff (1985).[17] The result is formally demonstrated in Lemma 2 below. Let us first, however, derive a citizen's expected utility as to the stage where neither ε nor s can be observed.

Consider a citizen/voter with responsiveness parameter λ_i. His expected utility from a policy maker with responsiveness parameter λ_g, denoted by Ev_i, may be written as

$$\begin{aligned} Ev_i &= \int\int v_i\left(\pi^*, \varepsilon\right) f\left(\varepsilon, s\right) d\varepsilon ds \\ &= -\left(\lambda_g \beta \overline{x}\right)^2 + \frac{\lambda_g \beta^2 \rho^2 \sigma^2}{\left(1+\lambda_g \beta^2\right)^2}\left[\lambda_i\left(2+\lambda_g \beta^2\right) - \lambda_g\right] - \lambda_i\left(\overline{x}^2+\sigma^2\right) \\ &= -\left(\lambda_g \beta \overline{x}\right)^2 + \rho^2\sigma^2\left(1+\lambda_i\beta^2\right)\varphi\left(\lambda_g\right)\left[2\varphi\left(\lambda_i\right) - \varphi\left(\lambda_g\right)\right] - \lambda_i\left(\overline{x}^2+\sigma^2\right) \end{aligned} \tag{0.44}$$

That is, the expression for Ev_i is very similar to the expression for Eu_i in equation (0.15). In fact, we have

$$\frac{\partial Eu_i}{\partial \rho^2} = \frac{\partial Ev_i}{\partial \rho^2} \tag{0.45}$$

and

$$\frac{\partial Eu_i}{\partial \sigma^2} = \frac{\partial Ev_i}{\partial \sigma^2}. \tag{0.46}$$

This means that the results stated in Propositions 1 and Observation 1 hold also in this alternative model, where the relationship between x and π is given by equation (0.37) instead of equation (0.2). However, the derivatives of Eu_i and Ev_i with respect to λ_g are not the same, which means that the incentives when electing a policymaker differ between the models. This is illustrated by the following lemma.

[17] See footnote 6 though.

Lemma 2. *The policymaker's responsiveness parameter* λ_g *is implicitly defined by*

$$\overline{x}^2\lambda_g\left(1+\lambda_g\beta^2\right)^3 = \rho^2\sigma^2\left(\lambda_m-\lambda_g\right). \qquad (0.47)$$

The result stated in Lemma 2 is a generalization of a result in Alesina and Grilli (1992).[18] Note that if we let $\overline{x}$ approach zero, then the median's favorite is $\lambda_g=\lambda_m$. In general, as long as $\rho^2>0$, $\lambda_g\in(0,\lambda_m)$. Thus, the policymaker is always less responsive than the median voter. However, when I now extend the model, allowing for an endogenous degree of informativeness of the policymaker's signal, this result will be altered.

5.2. The signal's informativeness being endogenous

Analogously to the analysis in Section 4, I now assume that the policymaker can make a costly effort and thereby improve upon the informativeness of the signal that he observes. The acquisition of information is modeled exactly as in Section 4.

We can immediately conclude that Proposition 2 stated in Section 4 holds also in this alternative model. This is because when equation (0.45) holds, the marginal benefit of making an effort is the same in either model; hence, since also the cost function is the same, a policymaker will make the same effort in both models. However, similarly to above, the incentives when electing a policymaker differ between the models. This is shown in the remaining part of this subsection.

The policymaker's expected utility at the stage where he is to choose the effort level e is given by

$$\begin{aligned} E\widetilde{v}_g &= \int\int v_g\left(\pi^*,\varepsilon\right)f\left(\varepsilon,s\right)d\varepsilon ds - C\left(e\right) \\ &= -\left(\lambda_g\beta\overline{x}\right)^2+\frac{\lambda_g^2\beta^2\sigma^2 e}{1+\lambda_g\beta^2}-\lambda_g\left(\overline{x}^2+\sigma^2\right)-C\left(e\right) \qquad (0.48)\end{aligned}$$

The problem of maximizing $E\widetilde{v}_g$ with respect to e subject to the constraint $e\in[0,1]$ has the same solution e^* as the corresponding problem in Section 4. That is, e^* is implicitly defined by equation (0.23).

Now consider again a voter with responsiveness parameter λ_i. His expected loss if the policymaker has responsiveness parameter λ_g and accordingly exerts

[18] Their result is obtained if one sets $\beta=\rho^2=1$.

effort $e^*(\lambda_g)$ may be written as

$$E\widetilde{v}_i = -(\lambda_g \beta \overline{x})^2 + \frac{\lambda_g \beta^2 e^* \sigma^2}{(1+\lambda_g \beta^2)^2} \left[\lambda_i \left(2+\lambda_g \beta^2\right) - \lambda_g\right] - \lambda_i \left(\sigma^2 + \overline{x}^2\right). \tag{0.49}$$

As with the model considered in Section 4, one may show that sufficient conditions for $E\widetilde{v}_i$ to be single peaked in λ_g are that $C(e) = e^a$ for $a \in \left(1, \frac{3}{2}\right)$ and that σ^2 is sufficiently close to zero. Here I confine myself with showing that when single-peakedness holds then, depending on parameter values, we may have any relation between λ_g and λ_m.

When the median voter theorem applies, then the policymaker's responsiveness parameter λ_g is implicitly defined by the following identity:

$$-2\beta^2 \overline{x}^2 \lambda_g + \frac{2\beta e^* \sigma^2 (\lambda_m - \lambda_g)}{(1+\lambda_g \beta^2)^3} + \frac{\lambda_g \beta^2 \sigma^2}{(1+\lambda_g \beta^2)^2} \left[\lambda_m \left(2+\lambda_g \beta^2\right) - \lambda_g\right] \frac{\partial e^*}{\partial \lambda_g} = 0. \tag{0.50}$$

The left-hand side of equation (0.50) was obtained by differentiating $E\widetilde{v}_i$ with respect to λ_g and evaluating at $\lambda_i = \lambda_m$. Inspecting equation (0.50), we can identify three different effects regarding the median voter's incentives to appoint a policymaker with certain preferences, each effect corresponding to one of the three terms of the left-hand side of equation (0.50). The first term is the only one containing $\overline{x}$, and it vanishes if $\overline{x} = 0$. This term captures the "Rogoff effect," i.e., the median voter's incentives to appoint a policymaker that is less flexible than himself, in order to mitigate the inflationary bias. The third term captures the "information acquisition effect." If the condition $\lambda_m (2 + \lambda_g \beta^2) > \lambda_g$ holds, then this effect counteracts the Rogoff effect; this condition, which is identical to the condition in Proposition 1, guarantees that it is in the median voter's interest that the policymaker acquires more information. If we had $\overline{x} = \frac{\partial e^*}{\partial \lambda_g} = 0$, then both the first and the third term would vanish, and we would reobtain the result of Section 3, $\lambda_g = \lambda_m$.

To see when the information acquisition effect is stronger than the Rogoff effect, so that $\lambda_m < \lambda_g$, let us evaluate the left-hand side of the identity (0.50) at $\lambda_m = \lambda_g$. Doing this yields

$$-2\lambda_m \beta^2 \overline{x}^2 + \frac{\beta^2 \lambda_m^2 \sigma^2}{1+\lambda_m \beta^2} \frac{\partial e^*}{\partial \lambda_g}. \tag{0.51}$$

In order to have $\lambda_m < \lambda_g$, this expression must be greater than zero, or

$$\frac{\partial e^*}{\partial \lambda_g} > \frac{2\overline{x}^2 (1+\lambda_m \beta^2)}{\lambda_m \sigma^2}. \tag{0.52}$$

Hence, if $\overline{x}$ is small relative to $\frac{\partial e^*}{\partial \lambda_g}$, it may be that the median voter delegates the task of conducting monetary policy to someone more responsive (or, in Rogoff's (1985) terminology, less conservative) than himself.

6. Concluding remarks

To summarize, this paper has considered two models. In the first one, the policymaker decides on the amount of a public good. This public good has an adverse effect on the environment, but the exact relationship between the public good and the environment is unknown. The second model concerns monetary policy with rational expectations. Here there is some uncertainty about an additive term in the expectations-augmented Phillips curve. In both models there is heterogeneity among the citizens with respect to how serious an issue one thinks the environment respectively full employment is — or, equivalently, concerning one's responsiveness to changes in the stochastic variable. Moreover, in both models, the policymaker is elected among the citizens by a majority vote.

Concerning both models two questions were posed. First, would all citizens be better off ex ante if the policymaker, when making the decision, were having more information about the realization of the stochastic variable? Second, would all citizens be better off ex ante if the variance of the stochastic variable were smaller? The second question was studied in two different environments. In the first one the policymaker can, prior to making the decision, observe a noisy signal about the stochastic variable, and the informativeness of this signal is given exogenously. In the second environment the policymaker can improve upon the informativeness of the signal by making a greater effort.

It was found that the answers to the questions are the same regardless which one of the two models is considered. Concerning the first question it was shown that only those citizens which are sufficiently responsive to the stochastic variable gain from a more informative signal. That is, the "non-environmentalists" or the conservatives are worse off. However, it turns out that a majority of the citizens are always better off from the policymaker's having access to a more informative signal. Concerning the second question it was found that, in the environment where the signal's informativeness is exogenous, everyone gains from the variance of the stochastic variable being smaller. However, when the signal's informativeness is endogenous, those people who are sufficiently responsive may be worse off from a smaller variance of the stochastic variable.

This paper has also considered another question, namely the relation between the policymaker's and the electorate's degree of responsiveness. Specifically, it was demonstrated that, in the environment where the policymaker can improve upon the informativeness of the signal by making a greater effort, a voter may have an incentive to delegate the task of deciding on public policy to a policymaker that is *more* responsive than himself. In particular, in the model on monetary policy, this means that the rate of inflation will be set by someone *less* conservative than the median voter.

Appendix

A. Proof of Lemma 1

To be able to invoke the median voter theorem one must show that Eu_i is single peaked in λ_g. Differentiate Eu_i in (0.15) with respect to λ_g:

$$\frac{\partial Eu_i}{\partial \lambda_g} = -2\left(1+\lambda_i\beta^2\right)\left[\psi'\left(\lambda_g\right)\left[\psi\left(\lambda_g\right)-\psi\left(\lambda_i\right)\right]+\sigma^2\rho^2\varphi'\left(\lambda_i\right)\left[\varphi\left(\lambda_g\right)-\varphi\left(\lambda_i\right)\right]\right]. \tag{0.53}$$

It is easy to check that $\varphi' > 0$ and that ψ' has the same sign as $(\overline{x}-\beta\overline{\pi})$. By inspecting equation (0.53) one sees that regardless of the sign of ψ' we have: $\frac{\partial Eu_i}{\partial \lambda_g} > 0$ for any $\lambda_g < \lambda_i$, $\frac{\partial Eu_i}{\partial \lambda_g} < 0$ for any $\lambda_g > \lambda_i$, and $\frac{\partial Eu_i}{\partial \lambda_g} = 0$ for $\lambda_g = \lambda_i$. Hence, Eu_i is single peaked in λ_g, and the peak is at $\lambda_g = \lambda_i$. □

B. Proof of Proposition 1

Differentiating Eu_i in (0.15) with respect to ρ^2 and evaluating at $\lambda_g = \lambda_m$ yield

$$\frac{\partial Eu_i}{\partial \rho^2}\mid_{\lambda_g=\lambda_m} = -\left(1+\lambda_i\beta^2\right)\sigma^2\varphi\left(\lambda_m\right)\left[\varphi\left(\lambda_m\right)-2\varphi\left(\lambda_i\right)\right], \tag{0.54}$$

which has the same sign as $\left(2\varphi\left(\lambda_i\right)-\varphi\left(\lambda_m\right)\right)$. By using the definition of φ and by carrying out some algebra, one may show that $\left(2\varphi\left(\lambda_i\right)-\varphi\left(\lambda_m\right)\right)$ in turn has the same sign as $\left(\lambda_i-\widetilde{\lambda}\right)$. □

C. Proof of Observation 1

Differentiating Eu_i in equation (0.15) with respect to σ^2 and evaluating at $\lambda_g = \lambda_m$ yield

$$\frac{\partial Eu_i}{\partial \sigma^2} \mid_{\lambda_g=\lambda_m} = -\rho^2 \varphi(\lambda_m)\left(1+\lambda_i\beta^2\right)\left[\varphi(\lambda_m) - 2\varphi(\lambda_i)\right] - \lambda_i < 0 \Leftrightarrow$$

$$\varphi(\lambda_i)\left[2\beta\rho^2\varphi(\lambda_m) - 1\right] < \beta\rho^2\varphi^2(\lambda_m). \tag{0.55}$$

If $2\beta\rho^2\varphi(\lambda_m) \leq 1$, then clearly $\frac{\partial Eu_i}{\partial \sigma^2} \mid_{\lambda_g=\lambda_m} < 0$. Suppose that $2\beta\rho^2\varphi(\lambda_m) > 1$. Then inequality (0.55) may be rewritten as

$$\varphi(\lambda_i) < \frac{\beta\rho^2\varphi^2(\lambda_m)}{2\beta\rho^2\varphi(\lambda_m) - 1}. \tag{0.56}$$

We must show that, when $2\beta\rho^2\varphi(\lambda_m) > 1$, inequality (0.56) always holds. To see that it does, note that the right-hand side of inequality (0.56) can be rewritten as follows:

$$\frac{\beta\rho^2\varphi^2(\lambda_m)}{2\beta\rho^2\varphi(\lambda_m) - 1} = \frac{1}{\beta\rho^2}\left[\frac{\left[\beta\rho^2\varphi(\lambda_m)\right]^2}{2\beta\rho^2\varphi(\lambda_m) - 1}\right] = \frac{1}{\beta\rho^2}\left[\frac{\left[\beta\rho^2\varphi(\lambda_m) - 1\right]^2}{2\beta\rho^2\varphi(\lambda_m) - 1} + 1\right], \tag{0.57}$$

which is greater than or equal to $\frac{1}{\beta}$. On the other hand, the left-hand side of inequality (0.56), $\varphi(\lambda_i)$, is strictly smaller than $\frac{1}{\beta}$. To see this, note that $\varphi' > 0$ and $\lim_{\lambda_i \to \infty} \varphi(\lambda_i) = \frac{1}{\beta}$. Hence inequality (0.56) must hold. □

D. Proof of the claim about single-peakedness in Section 4

Here I prove the claim made in Section 4 that $E\widetilde{u}_i$ is single peaked in λ_g if $C(e) = e^a$, $a \in \left(1, \frac{3}{2}\right)$, $\overline{x} \neq \beta\overline{\pi}$, and σ^2 is sufficiently close to zero.

If $C(e) = e^a$ then e^* is given by equation (0.32). It is a straightforward exercise to show that, under the assumption $a \in \left(1, \frac{3}{2}\right)$,

$$\lim_{\sigma^2 \to 0} e^* = \lim_{\sigma^2 \to 0} \frac{\partial e^*}{\partial \lambda_g} = \lim_{\sigma^2 \to 0} \frac{\partial^2 e^*}{\partial (\lambda_g)^2} = 0. \tag{0.58}$$

Now differentiate $E\widetilde{u}_i$ in equation (0.28) once with respect to λ_g:

$$\frac{\partial E\widetilde{u}_i}{\partial \lambda_g} = -2\left(1+\lambda_i\beta^2\right)\left[\psi'(\lambda_g)\left[\psi(\lambda_g) - \psi(\lambda_i)\right]\right.$$

$$+e^*\sigma^2\varphi'(\lambda_g)\left[\varphi(\lambda_g)-\varphi(\lambda_i)\right]\Big]-\left(1+\lambda_i\beta^2\right)\sigma^2\varphi(\lambda_g)\left[\varphi(\lambda_g)-2\varphi(\lambda_i)\right]\frac{\partial e^*}{\partial\lambda_g}. \tag{0.59}$$

Differentiating $E\widetilde{u}_i$ once more yields:

$$\begin{aligned}\frac{\partial^2 E\widetilde{u}_i}{\partial\lambda_g^2} &= -2\left(1+\lambda_i\beta^2\right)\left[\psi''(\lambda_g)\left[\psi(\lambda_g)-\psi(\lambda_i)\right]+\left[\psi'(\lambda_g)\right]^2\right.\\ &\quad e^*\sigma^2\left[\varphi''(\lambda_g)\left[\varphi(\lambda_g)-\varphi(\lambda_i)\right]+\left[\varphi'(\lambda_g)\right]^2\right]+2\sigma^2\varphi'(\lambda_g)\left[\varphi(\lambda_g)-\varphi(\lambda_i)\right]\left.\frac{\partial e^*}{\partial\lambda_g}\right]\\ &\quad-\left(1+\lambda_i\beta^2\right)\sigma^2\varphi(\lambda_g)\left[\varphi(\lambda_g)-2\varphi(\lambda_i)\right]\frac{\partial^2 e^*}{\partial\lambda_g^2}.\end{aligned} \tag{0.60}$$

The assumption $\overline{x}\neq\beta\overline{\pi}$ implies that $\psi'(\lambda_g)\neq 0$. This means that $\frac{\partial E\widetilde{u}_i}{\partial\lambda_g}=0$ if and only if

$$\begin{aligned}\left[\psi(\lambda_g)-\psi(\lambda_i)\right] &= \frac{-\sigma^2}{2\psi'(\lambda_g)}\left[2e^*\varphi'(\lambda_g)\left[\varphi(\lambda_g)-\varphi(\lambda_i)\right]\right.\\ &\quad\left.+\varphi(\lambda_g)\left[\varphi(\lambda_g)-2\varphi(\lambda_i)\right]\frac{\partial e^*}{\partial\lambda_g}\right]\\ &\equiv \Theta\left(\sigma^2\right)\end{aligned} \tag{0.61}$$

Substituting $\left[\psi(\lambda_g)-\psi(\lambda_i)\right]$ for $\Theta(\sigma^2)$ in equation (0.60) and then taking the limit $\sigma^2\to 0$ yield

$$\lim_{\sigma^2\to 0}\frac{\partial^2 E\widetilde{u}_i}{\partial\lambda_g^2}\,|_{[\psi(\lambda_g)-\psi(\lambda_i)]=\Theta(\sigma^2)}=-2\left(1+\lambda_i\beta^2\right)\left[\psi'(\lambda_g)\right]^2<0. \tag{0.62}$$

By continuity, $\frac{\partial^2 E\widetilde{u}_i}{\partial\lambda_g^2}$ evaluated at $\left[\psi(\lambda_g)-\psi(\lambda_i)\right]=\Theta(\sigma^2)$ is strictly negative also for some strictly positive σ^2, which proves the claim. □

E. Proof of Proposition 2

Differentiating $E\widetilde{u}_i(\lambda_g)$ in equation (0.28) with respect to σ^2 (and making use of equation (0.27)) yield

$$\frac{\partial E\widetilde{u}_i}{\partial\sigma^2}=-\varphi(\lambda_g)\left(1+\lambda_i\beta^2\right)\left[\varphi(\lambda_g)-2\varphi(\lambda_i)\right]Z-\lambda_i. \tag{0.63}$$

Thus, $\frac{\partial E\tilde{u}_i}{\partial \sigma^2} < 0$ is equivalent to

$$\varphi(\lambda_i)\left[2\beta Z\varphi(\lambda_g) - 1\right] < \beta Z\left[\varphi(\lambda_g)\right]^2. \tag{0.64}$$

If $2\beta Z\varphi(\lambda_g) \leq 1$, then clearly $\frac{\partial E\tilde{u}_i}{\partial \sigma^2} < 0$. Suppose that $2\beta Z\varphi(\lambda_g) > 1$. Then we may rewrite inequality (0.64) as (0.31). □

F. Proof of Lemma 2

I first show that Ev_i is single peaked in λ_g. Differentiate Ev_i once with respect to λ_g:

$$\frac{\partial Ev_i}{\partial \lambda_g} = -2\lambda_g\beta^2\overline{x}^2 + \frac{2\beta^2\rho^2\sigma^2(\lambda_i - \lambda_g)}{(1+\lambda_g\beta^2)^3}. \tag{0.65}$$

And once more:

$$\frac{\partial^2 Ev_i}{\partial \lambda_g^2} = -2\beta^2\overline{x}^2 - 2\beta^2\rho^2\sigma^2\frac{1+\lambda_g\beta^2+3\beta^2(\lambda_i - \lambda_g)}{(1+\lambda_g\beta^2)^4}. \tag{0.66}$$

Note that when $\frac{\partial Ev_i}{\partial \lambda_g} = 0$, we must have $\lambda_i > \lambda_g$. Thus, when evaluated at values of λ_g satisfying $\frac{\partial Ev_i}{\partial \lambda_g} = 0$, the second derivative $\frac{\partial^2 Ev_i}{\partial \lambda_g^2}$ is strictly negative. Hence Ev_i is single peaked in λ_g, and we can apply the median voter theorem. λ_g is accordingly given by the median citizen's favorite. That is, λ_g is defined by (0.47). □

References

Alesina, A., and V. Grilli, 1992, The European Central Bank: Reshaping monetary politics in Europe, in "Establishing a central bank. Issues in Europe and lessons from the US," Canzoneri, M., V., Grilli, and P. Masson (eds.), Cambridge University Press.

Austen-Smith, D., 1997, Interest groups: Money, information, and influence, in: "Perspectives on public choice: A handbook" D.C. Mueller (ed.), Cambridge University Press.

Austen-Smith, D., and J.R. Wright, 1992, Competitive lobbying for a legislator's vote, Social Choice and Welfare 9, 229-257.

Barro, R.J., and D.B. Gordon, 1983, A positive theory of monetary policy in a natural rate model, Journal of Political Economy 91: 589-610.

Björnerstedt, Jonas, 1995, Elections and the credibility of monetary policy, mimeo, Stockholm School of Economics.

Freixas, X., and R.E. Kihlstrom, 1984, Risk aversion and information demand, in: "Bayesian models in economic theory," M. Boyer and R.E. Kihlstrom (eds.), North-Holland.

Kydland, F.E., and E.C. Prescott, 1977, Rules rather than discretion: The inconsistency of optimal plans, Journal of Political Economy 85: 473-491.

Lagerlöf, J., 1997, Lobbying, information, and private and social welfare, European Journal of Political Economy 13: 615-637.

Martinelli, C., 1997, Information revelation in electoral competition, mimeo, Universidad Carlos III de Madrid.

Martinelli, C., and A. Matsui, 1997, Policy reversals: A Democratic Nixon and a Republican Clinton, mimeo, Universidad Carlos III de Madrid.

Mueller, D.C., 1989, Public Choice II, Cambridge: Cambridge University Press.

Persson, T., and G. Tabellini, 1990, Macroeconomic policy, credibility and politics, London: Harwood Publishers.

Persson, T., and G. Tabellini, 1994, Representative democracy and capital taxation, Journal of Public Economics 55, 53-70.

Potters, J., and F. van Winden, 1992, Lobbying and asymmetric information, Public Choice 74, 269-292.

Radner, R., and J.E. Stiglitz, 1984, A nonconcavity in the value of information, in: "Bayesian models in economic theory," M. Boyer and R.E. Kihlstrom (eds.), North-Holland.

Reed, W.R., (1989), Information in political markets: A little knowledge can be a dangerous thing, Journal of Law, Economics, and Organization 5, 355-374.

Rogoff, K., (1985), The optimal degree of commitment to an intermediate monetary target, Quarterly Journal of Economics 100, 1169-1190.

Sakai, Y., 1985, The value of information in a simple duopoly model, Journal of Economic Theory 36, 36-54.

Schultz, C., 1996, Political competition and polarization, mimeo, University of Copenhagen.

Sloof, R., 1997, Game-theoretic models of the political influence of interest groups, Thesis Publishers, Amsterdam.

Waller, C.J., and C.E. Walsh, 1996, Central-bank independence, economic behavior, and optimal term lengths, American Economic Review 86: 1139-1153.

Chapter 3

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

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

Johan Lagerlöf*

December 11, 1997

Abstract

Suppose that 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 an arbitrarily small 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.

//[Doc: SD-5.tex]//

JEL classification: D64; D82

Keywords: Altruism; Signaling; Samaritan's dilemma; Efficiency; Saving

---

* University of Rochester and Stockholm School of Economics. Correspondense to: W. Allen Wallis Institute of Political Economy, University of Rochester, 110 Harkness Hall, Rochester, NY 14627, USA; e-mail: jlgf@troi.cc.rochester.edu or nejla@hhs.se. I have benefited from helpful discussions with and/or comments from Jonas Björnerstedt, Yeongjae Kang, Nippe Lagerlöf, Björn Persson, Johan Stennek, Karl Wärneryd, and Jörgen Weibull. Financial support from Jan Wallander and Tom Hedelius' Research Foundation, *Svenska Institutet*, and *Finanspolitiska Forskningsinstitutet* is gratefully acknowledged.

1. Introduction

Suppose that an altruistic person, 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 more today and less tomorrow. And, since A has altruistic concerns for the welfare of B, this would make also A better off.[1,2]

The mechanism in the Samaritan's dilemma has been employed in an extensive literature, 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; 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.[3] In those models the government provides a transfer on two occasions over time. Since a cash transfer would be used in an inefficient manner, the government may give 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 toward their parents. When, as in the standard literature, parental saving is non-strategic, this kind of model is characterized by dynamic inefficiency, i.e.,

[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] The situation described may also be thought of as a situation where two agents sequentially make private contributions to a public good; see Varian (1994).

[3] 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.

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 mechanism in 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 external effect, 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.[4] 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 savings 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. Consequently, "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).

In spite of the extensive use in the literature of the mechanism in the Samaritan's dilemma, little attention has been paid to the question whether it is robust to, for example, changes in the assumption that information is complete. This paper is an attempt in that direction. In the version of the Samaritan's dilemma that I consider there are two agents, A and B, and two periods, 1 and 2. A, who has altruistic feelings for B, lives only in period 2. In period 1, B makes a saving decision. In period 2, A observes the level of B's savings and thereupon decides on a transfer to B. The important extension of the existing literature on the Samaritan's dilemma is that B has private information about a certain parameter in his utility function. In particular I consider the case where the uncertainty is arbitrarily small; hence the extension serves as a robustness test of the standard model.

[4]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.

The parameter in question, β, is such that the larger its magnitude, the greater is B's marginal utility of consumption in period 2. This means that A is willing to make a larger transfer to B if A believes that β is large. It also means that the larger is β, the more B wants to save (everything else being equal). B thus has an incentive to make A believe that β is large, 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. This mechanism counteracts the incentives to undersave in the standard Samaritan's dilemma model. Indeed, it turns out that in the presence of an arbitrarily small amount of uncertainty, a large number of different outcomes can be sustained as equilibria, and these outcomes may involve both undersaving and oversaving. Although full efficiency cannot be obtained, one may get arbitrarily close, in a sense made precise, to ex post incentive efficiency.

The remainder of the paper is organized as follows. In the next section a benchmark model with complete information is presented and analyzed. In Section 3 I consider the model with incomplete information. Section 4 briefly summarizes and concludes. Most of the proofs are found in an appendix.

2. A Benchmark: Complete Information

In this section I will consider a very simple model giving rise to the Samaritan's dilemma. The model 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 3, the model with incomplete information is considered.

2.1. 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, $t \in [0, \omega]$.[5] 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 von Neumann-Morgenstern utility function:

$$\begin{aligned} U_B(s,t) &= \log(c_{1B}) + \beta \log(c_{2B}) \\ &= \log(\omega - s) + \beta \log(s + t), \end{aligned} \tag{2.1}$$

where $\beta \in (0,1)$ is a fixed parameter.[6] A is altruistic in the sense that she has preferences over both her own consumption and B's utility level U_B. These preferences are described by the following von Neumann-Morgenstern utility function:

$$\begin{aligned} U_A(s,t) &= \gamma \log(c_A) + \alpha U_B(s,t) \\ &= \gamma \log(\omega - t) + \alpha \log(\omega - s) + \alpha\beta \log(s + t), \end{aligned} \tag{2.2}$$

where $\gamma > 0$ and $\alpha > 0$ are fixed parameters; the parameter α represents the altruistic concern of A for the welfare of B. The structure of the model and in particular the individuals' preferences are assumed to be common knowledge.

2.2. 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 Nash 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 (2.2) 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

$$\widehat{t} = \begin{cases} \frac{\alpha\beta\omega - \gamma s}{\gamma + \alpha\beta} & \text{for } s \leq \frac{\alpha\beta}{\gamma}\omega \\ 0 & \text{otherwise.} \end{cases} \tag{2.3}$$

Note that for any $s < \alpha\beta\omega/\gamma$, $\widehat{t}$ 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

[5]Hence a non-negativity constraint is imposed on the transfer, $t \geq 0$: A cannot *take* income from B.

[6]The assumption that β is smaller than unity will be made also in the model with incomplete information considered in Section 3. The assumption is not essential for my argument but it simplifies the analysis.

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 function is given by

$$U_B\left(s,\widehat{t}\right) = \log(\omega - s) + \begin{cases} \beta\log(\omega+s) + \beta\log\left(\frac{\alpha\beta}{\gamma+\alpha\beta}\right) & \text{for } s \leq \frac{\alpha\beta}{\gamma}\omega \\ \beta\log(s) & \text{otherwise.} \end{cases} \tag{2.4}$$

There are two cases to investigate: (i) $\alpha\beta \geq \gamma$ and (ii) $\alpha\beta < \gamma$. 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

$$\max_{s\in[0,\omega]} \log(\omega - s) + \beta\log(\omega+s) + \beta\log\left(\frac{\alpha\beta}{\gamma+\alpha\beta}\right). \tag{2.5}$$

Since by assumption $\beta < 1$, 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 (2.3) yields the equilibrium outcome of t, $t^* = \alpha\beta\omega/(\gamma+\alpha\beta)$.

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

$$\varphi(\beta,\gamma) = \frac{\gamma}{(1+\beta)^{\frac{1+\beta}{\beta}} - \beta}. \tag{2.6}$$

The following lemma, which is proven in the Appendix, states that the partial derivative of φ with respect to β is negative.

Lemma 1. *$\varphi_1'(\cdot,\gamma) < 0$ for all $\gamma \in \Re_{++}$.*

One may also show that $\lim_{\beta\to 0}\varphi(\beta,\gamma) = \gamma/e$ (where $e^{-1} \approx 0.37$)[7] and it is easy to see that $\lim_{\beta\to 1}\varphi(\beta,\gamma) = \gamma/3$. Thus, for any $\beta \in (0,1)$, $\varphi(\beta,\gamma)$ approximately equals one third of γ.

[7]This exercise is left to the reader.

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

$$(s^*,t^*) = \begin{cases} \left(\frac{\beta}{1+\beta}\omega, 0\right) & \text{for } \alpha < \varphi(\beta,\gamma) \\ \left(0, \frac{\alpha\beta}{\gamma+\alpha\beta}\omega\right) & \text{for } \alpha > \varphi(\beta,\gamma). \end{cases} \tag{2.7}$$

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

Figure 1 illustrates the results stated in the proposition. φ, the critical value of α defined in equation (2.6), 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 $\alpha\beta/(\gamma+\alpha\beta)$ of her income to B. For values of α exactly at the critical value, $\alpha = \varphi(\beta,\gamma)$, 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 (2.3), i.e., either not make any transfer or transfer the fraction $\alpha\beta/(\gamma+\alpha\beta)$ of her income.

2.3. 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 Nash 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 savings-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

$$U_i\left(s',t'\right) \geq U_i(s,t), \qquad \forall i \in \{A,B\} \tag{2.8}$$

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

That is, if A's degree of altruism takes on any value $\alpha \in \left(\varphi(\beta,\gamma), \frac{\gamma}{1-\beta}\right)$, then the equilibrium outcome is not Pareto efficient (cf. Figure 1). Recall that for these values of α, B saves nothing but receives a transfer $t^* = \frac{\alpha\beta}{\gamma+\alpha\beta}\omega$ 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 B's saving level and that she cannot precommit to any transfer level.

If the degree of altruism is either sufficiently low ($\alpha < \varphi(\beta,\gamma)$) or sufficiently high ($\alpha \geq \frac{\gamma}{1-\beta}$), then the equilibrium allocation is Pareto efficient. The intuition for this is straightforward. For $\alpha < \varphi(\beta,\gamma)$, 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 savings choice is not distorted. Hence, any other choice of (s,t) would make A strictly worse off. Similarly for the case $\alpha > \frac{\gamma}{1-\beta}$. For these parameter values the equilibrium outcome is again A's first-best alternative: Since she cares so much for the welfare of B, it is in her interest that B does not save anything himself.

3. Incomplete Information

3.1. The Model

In this section I consider an adaptation of the model described and analyzed in the previous section. Relative to that model 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. The 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 puts 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 2.1. In particular there are two periods. B lives in both of them while A lives only in period 2. The players 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 level of savings chosen by the low type and s_H is the level of savings chosen by the high type. A does not know B's type but observes his actual saving level, 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 von Neumann-Morgenstern utility functions, given B's type, by $U_A(s,t \mid \beta)$ and $U_B(s,t \mid \beta)$:

$$U_A(s,t \mid \beta) = \gamma \log(\omega - t) + \alpha \log(\omega - s) + \alpha\beta \log(s+t), \qquad \beta = \beta_L, \beta_H, \quad (3.1)$$

$$U_B(s,t \mid \beta) = \log(\omega - s) + \beta \log(s+t), \qquad \beta = \beta_L, \beta_H. \quad (3.2)$$

Again, $\alpha > 0$ and $\gamma > 0$ are fixed parameters.

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 sequential equilibrium. A sequential 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 $t^*(s)$:

$$s_i^* \in \arg\max_{s \in [0,\omega]} U_B(s, t^*(s) \mid \beta_i), \qquad \forall i \in \{L, H\}; \quad (3.3)$$

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

$$t^*(s) \in \arg\max_{t(s) \in [0,\omega]} (1 - \widetilde{\mu}(s)) U_A(s, t(s) \mid \beta_L) + \widetilde{\mu}(s) U_A(s, t(s) \mid \beta_H), \qquad \forall s \in [0,\omega]; \quad (3.4)$$

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

$$\widetilde{\mu}(s_H^*) = 1 \text{ and } \widetilde{\mu}(s_L^*) = 0 \text{ if } s_H^* \neq s_L^*, \quad (3.5)$$

$$\widetilde{\mu}(s_H^*) = \widetilde{\mu}(s_L^*) = \mu \text{ if } s_H^* = s_L^*. \quad (3.6)$$

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

3.2. Signaling

Γ_I is an example of a *signaling game*: One party (B) possesses private information. On the basis of this information he sends a signal to a second party (A), who thereupon takes an action. This kind of model was first analyzed by Spence (1973, 1974), and today there exist numerous applications of signaling games.[8] A pure-strategy equilibrium of a signaling game can be either *separating* or *pooling.* A separating equilibrium is an equilibrium where $s_L^* \neq s_H^*$, i.e., the types behave differently. A pooling equilibrium is an equilibrium where $s_L^* = s_H^*$, i.e., the types behave identically.

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

Assumption 1.

$$\alpha \in \left(\varphi(\beta_L, \gamma), \frac{\gamma}{1-\beta_L}\right). \quad (3.7)$$

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).

3.2.1. 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 $\alpha > \varphi(\beta_L, \gamma)$ the low type chooses not to save.

Lemma 2. *Suppose that $\alpha > \varphi(\beta_L, \gamma)$. Then, in any separating equilibrium,* $s_L^* = 0$.

[8] See, e.g., van Damme (1991) for references.

We are thus looking for an equilibrium where $s_L^* = 0$ and where s_H^* is positive. There are two possibilities to investigate: either s_H^* belongs to the set $\left(0, \frac{\alpha\beta_H\omega}{\gamma}\right]$, in which case the non-negativity constraint on A's transfer is not binding; or s_H^* belongs to the set $\left(\frac{\alpha\beta_H\omega}{\gamma}, \omega\right]$, in which case the constraint is binding. Let us call a separating equilibrium belonging to the first category a "type I equilibrium" and an equilibrium belonging to the second category a "type II equilibrium." Similarly to the case with the analysis of the benchmark model in Section 2, the analysis of the type II equilibria is fairly cumbersome and is therefore deferred to the Appendix; the results of the analysis carried out there will be reported later in this subsection though. Below follows an analysis of the type I equilibria.

In order to have $s_L^* = 0$ and $s_H^* \in \left(0, \frac{\alpha\beta_H\omega}{\gamma}\right]$ in a separating equilibrium there are two necessary conditions. First, the low type must not have an incentive to choose the high type's saving level:

$$U_B\left(0, \frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L} \mid \beta_L\right) \geq U_B\left(s_H^*, \frac{\alpha\beta_H\omega - \gamma s_H^*}{\gamma + \alpha\beta_H} \mid \beta_L\right). \tag{3.8}$$

Second, the high type must not have an incentive to choose the low type's saving level:

$$U_B\left(s_H^*, \frac{\alpha\beta\omega - \gamma s_H^*}{\gamma + \alpha\beta} \mid \beta_H\right) \geq U_B\left(0, \frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L} \mid \beta_H\right). \tag{3.9}$$

Figure 2 illustrates these two conditions in the savings-transfer space. The two straight lines in the figure are A's optimal transfers, as given by equation (2.3); the lower straight line corresponds to A's believing that $\beta = \beta_L$, and the upper one corresponds to A's believing that $\beta = \beta_H$. The figure also depicts two indifference curves through the point $(s,t) = \left(0, \frac{\alpha\beta_L\omega}{\gamma+\alpha\beta_L}\right)$, one for the low type and one for the high type. These levels of utility thus correspond to the types' choosing the low type's equilibrium saving level and receiving the low type's equilibrium transfer. Note that, for each type, an indifference curve located above and to the left of another one corresponds to a higher level of utility. Since A will learn B's type perfectly in any separating equilibrium, A will make a transfer according to either one of the two straight lines. 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 (i.e., for condition (3.8) to hold) we must have $s_H^* \geq s_I^{'}$; otherwise the low type could, by saving s_H^*, obtain a savings-transfer pair (s,t) which corresponds to some (not drawn) indifference curve above and to the left of the one drawn in the figure.

By using a similar argument, one may verify that for condition (3.9) to hold we must have $s^* \leq s_I''$. The critical values s_I' and s_I'' are implicitly defined by

$$\log\left(\omega - s_I'\right) + \beta_L \log\left(\omega + s_I'\right) \equiv \log(\omega) + \beta_L \log\left(\frac{\omega\beta_L(\gamma + \alpha\beta_H)}{\beta_H(\gamma + \alpha\beta_L)}\right) \qquad (3.10)$$

respectively

$$\log(\omega - s_I'') + \beta_H \log(\omega + s_I'') \equiv \log(\omega) + \beta_H \log\left(\frac{\omega\beta_L(\gamma + \alpha\beta_H)}{\beta_H(\gamma + \alpha\beta_L)}\right). \qquad (3.11)$$

These two identities were obtained by letting conditions (3.8) respectively (3.9) hold with equality, setting $s_H^* = s_I'$ respectively $s_H^* = s_I''$, and using the explicit functional form given in equation (3.2). Note that the left-hand side of (3.10) is strictly decreasing in s_I'. Also, the left-hand side of (3.10) tends to $-\infty$ as s_I' tends to ω; and as s_I' tends to 0 the left-hand side tends to $(1 + \beta_L)\log(\omega)$, which is strictly greater than the right-hand side of (3.10). Hence the identity in (3.10) uniquely defines s_I'. Similarly with s_I''. One may also verify that $0 < s_I' < s_I'' < \omega$.

If A's out-of-equilibrium beliefs are chosen appropriately, these two necessary conditions are also sufficient for having $s_L^* = 0$ and $s_H^* \in \left(0, \frac{\alpha\beta_H\omega}{\gamma}\right]$ in a separating equilibrium. For instance, one may let $\widetilde{\mu}(s) = 1$ for all $s \geq s_H^*$ and $\widetilde{\mu}(s) = 0$ for all $s < s_H^*$. These posterior beliefs are consistent with the equilibrium requirements. Hence, letting the set Ω_I^{sep} be defined by

$$\Omega_I^{sep} = \left(0, \frac{\alpha\beta_H\omega}{\gamma}\right] \cap \left[s_I', s_I''\right], \qquad (3.12)$$

we have shown that any $(s_L^*, s_H^*) \in \{0\} \times \Omega_I^{sep}$ is part of an outcome of a separating type I equilibrium. If $\alpha\beta_H \geq \gamma$, then the set Ω_I^{sep} is always non-empty (in fact, then $\Omega_I^{sep} = \left[s_I', s_I''\right]$). If $\alpha\beta_H < \gamma$, then Ω_I^{sep} is non-empty if and only if $s_I' \leq \frac{\alpha\beta_H\omega}{\gamma}$ or, equivalently,

$$\beta_L \log\left(\frac{\beta_H(\gamma + \alpha\beta_L)}{\gamma\beta_L}\right) \leq \log\left(\frac{\gamma}{\gamma - \alpha\beta_H}\right). \qquad (3.13)$$

Condition (3.13) was obtained by substituting $s_H^* = \frac{\alpha\beta_H\omega}{\gamma}$ into condition (3.8) and rewriting. If the stronger condition $s_I'' \leq \frac{\alpha\beta_H\omega}{\gamma}$ holds, then $\Omega_I^{sep} = \left[s_I', s_I''\right]$. By using condition (3.9) one may equivalently (and similarly to above) write the condition $s_I'' \leq \frac{\alpha\beta_H\omega}{\gamma}$ as

$$\beta_H \log\left(\frac{\beta_H\left(\gamma+\alpha\beta_L\right)}{\gamma\beta_L}\right) \leq \log\left(\frac{\gamma}{\gamma-\alpha\beta_H}\right). \tag{3.14}$$

If condition (3.14) does not hold but condition (3.13) does, then $\Omega_I^{sep} = \left[s_I', \frac{\alpha\beta_H\omega}{\gamma}\right]$.

In summary we have the following proposition.

Proposition 3a. *(Separating equilibria, type I) There exists a separating equilibrium of type I if and only if any one of the following three conditions is met:*

1. $\alpha\beta_H \geq \gamma$;
2. $\alpha\beta_H < \gamma$ *and (3.14) holds;*
3. $\alpha\beta_H < \gamma$, *(3.13) holds, but (3.14) does not hold.*

The outcome of a separating equilibrium of type I can be any

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

where

$\Omega_I^{sep} = \left[s_I', s_I''\right]$ *if either one of condition 1 or 2 above is met; and*

$\Omega_I^{sep} = \left[s_I', \frac{\alpha\beta_H\omega}{\gamma}\right]$ *if condition 3 is met.*

Hence, in any separating type I equilibrium, the low type saves nothing and the high type saves a positive amount $s_H^* \in \Omega_I^{sep}$. A then infers B's true type and makes a transfer t_L^* or t_H^*, as described by equation (2.3) with the appropriate value of β.

Let us now turn to the type II equilibria. Consider these two conditions:

$$\beta_H \log\left(\frac{\beta_H\left(\gamma+\alpha\beta_L\right)}{\alpha\beta_L}\right) \geq \left(1+\beta_H\right)\log\left(1+\beta_H\right), \tag{3.15}$$

$$\beta_H \log\left(\frac{\beta_H\left(\gamma+\alpha\beta_L\right)}{\gamma\beta_L}\right) < \log\left(\frac{\gamma}{\gamma-\alpha\beta_H}\right). \tag{3.16}$$

Let $s_{II}^{'}$ and $s_{II}^{''}$ be defined by

$$s_{II}^{'} = \min\left\{ s \mid \log(\omega - s) + \beta_H \log(s) = \log(\omega) + \beta_H \log\left(\frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L}\right)\right\} \quad (3.17)$$

and

$$s_{II}^{''} = \max\left\{ s \mid \log(\omega - s) + \beta_H \log(s) = \log(\omega) + \beta_H \log\left(\frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L}\right)\right\}. \quad (3.18)$$

Condition (3.16) is well-defined when $\gamma > \alpha\beta_H$; and if condition (3.15) is satisfied with strict inequality, then $s_{II}^{'}$ and $s_{II}^{''}$ are also well-defined.

The following proposition is proven in the Appendix.

Proposition 3b. *(Separating equilibria, type II) There exists a separating equilibrium of type II if and only if any one of the following three conditions is met:*

1. *(3.15) is satisfied with equality and* $\alpha \leq \frac{\gamma}{1+\beta_H}$;
2. *(3.15) is satisfied with strict inequality, (3.16) holds, and* $\alpha < \frac{\gamma}{1+\beta_H}$;
3. *(3.15) is satisfied with strict inequality, (3.16) does not hold, and* $\alpha\beta_H < \gamma$.

The outcome of a separating equilibrium of type II can be any

$$(s_L^*, t_L^*, s_H^*, t_H^*) \in \{0\} \times \left\{\frac{\alpha\beta_L}{\gamma + \alpha\beta_L}\right\} \times \Omega_{II}^{sep} \times \{0\},$$

where

$\Omega_{II}^{sep} = \left\{\frac{\beta_H\omega}{1+\beta_H}\right\}$ *if the condition 1 above is met;*

$\Omega_{II}^{sep} = \left[s_{II}^{'}, s_{II}^{''}\right]$ *if condition 2 is met; and*

$\Omega_{II}^{sep} = \left[\frac{\alpha\beta_H\omega}{\gamma}, s_{II}^{''}\right]$ *if condition 3 is met.*

3.2.2. Pooling Equilibria

Now consider pooling equilibria. We are looking for an equilibrium where both types make the same savings decision, say $s_L^* = s_H^* = s^*$. After having observed this savings level, A will make a transfer to B according to equation (2.3) where the parameter β has been substituted for the expected value of β, $E(\beta) \equiv \mu\beta_H + (1-\mu)\beta_L$. Again there are two possibilities to investigate: either s^* belongs to the set $\left[0, \frac{\alpha E(\beta)\omega}{\gamma}\right)$, in which case A's transfer will be positive; or s^* belongs to the set $\left[\frac{\alpha E(\beta)\omega}{\gamma}, \omega\right]$, in which case A's transfer is zero. It turns out that pooling equilibria where $s^* \in \left[\frac{\alpha E(\beta)\omega}{\gamma}, \omega\right]$ do not exist.[9]

Let us thus investigate whether and under what conditions we can have $s^* \in \left[0, \frac{\alpha E(\beta)\omega}{\gamma}\right)$ in a pooling equilibrium. It is fairly easy to see that one can always sustain $s^* = 0$ as part of a pooling equilibrium. If $s^* = 0$, then A's transfer is given by $t^* = \frac{\alpha E(\beta)\omega}{\gamma+\alpha E(\beta)}$. If A's beliefs are given by $\widetilde{\mu}(0) = \mu$ and $\widetilde{\mu}(s) = 0$ for $s \in (0, \omega]$, then neither the high nor the low type has an incentive to deviate from the savings level $s^* = 0$.

For $s^* \in \left(0, \frac{\alpha E(\beta)\omega}{\gamma}\right)$ in a pooling equilibrium, the following condition is necessary: the low type must not have an incentive to choose the saving level $s_L = 0$, thereby guaranteeing himself the transfer $\alpha\beta_L\omega/(\gamma+\alpha\beta_L)$:

$$U_B\left(0, \frac{\alpha\beta_L\omega}{\gamma+\alpha\beta_L} \mid \beta_L\right) \le U_B\left(s^*, \frac{\alpha E(\beta)\omega - \gamma s^*}{\gamma+\alpha E(\beta)} \mid \beta_L\right). \tag{3.19}$$

This necessary condition holds for all $s^* \in [0, \widehat{s}]$, where $\widehat{s}$ is obtained by letting inequality (3.19) hold with equality, setting $s^* = \widehat{s}$, and using the explicit functional form given in equation (3.2):

$$\log(\omega - \widehat{s}) + \beta_L \log(\omega + \widehat{s}) \equiv \log(\omega) + \beta_L \log\left[\frac{\beta_L\omega(\gamma+\alpha E(\beta))}{E(\beta)(\gamma+\alpha\beta_L)}\right]. \tag{3.20}$$

This identity uniquely defines $\widehat{s}$ (cf. the remarks after the definitions of s_I' and s_{II}'). Again, by appropriately choosing player A's out-of-equilibrium beliefs, this necessary condition is also sufficient. We can thus sustain any $s^* \in \Omega^{pool} \equiv \left[0, \frac{\alpha E(\beta)\omega}{\gamma}\right) \cap [0, \widehat{s}]$ as part of a pooling equilibrium. Note that we have $\widehat{s} \in (0, s_I')$. This means that the set Ω^{pool} is non-empty.

[9] This follows from Claim 1, stated in the proof of Proposition 3b in the Appendix.

Clearly, if $\alpha E(\beta) \geq \gamma$ or if $\alpha E(\beta) < \gamma$ and $\frac{\alpha E(\beta)\omega}{\gamma} \geq \widehat{s}$, then $\Omega^{pool} = [0, \widehat{s}]$; otherwise $\Omega^{pool} = \left[0, \frac{\alpha E(\beta)\omega}{\gamma}\right)$. One may verify that $\frac{\alpha E(\beta)\omega}{\gamma} \geq \widehat{s}$ is equivalent to

$$\log\left[\frac{\gamma}{\gamma - \alpha E(\beta)}\right] \geq \beta_L \log\left[\frac{E(\beta)(\gamma + \alpha\beta_L)}{\beta_L \gamma}\right]. \tag{3.21}$$

In summary we have the following proposition.

Proposition 4. *(Pooling equilibria) Pooling equilibria exist. The outcome of a pooling equilibrium can be any*

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

such that $s_L^* = s_H^* = s^*$, *where*

1. $\Omega^{pool} = [0, \widehat{s}]$ *if and only if either* $\alpha E(\beta) \geq \gamma$ *or* $\alpha E(\beta) < \gamma$ *and inequality (3.21) holds; and*
2. $\Omega^{pool} = \left[0, \frac{\alpha E(\beta)\omega}{\gamma}\right)$ *if and only if* $\alpha E(\beta) < \gamma$ *and inequality (3.21) does not hold.*

3.3. Efficiency

The aim of this subsection is to investigate whether allocations induced by the equilibria of the game Γ_I are efficient and, when not, whether B saves "too little" or "too much" in these equilibria. Similarly to the usage in Section 2, I will refer to a savings-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.[10] Following Holmström and Myerson (1983) I say that an allocation (s_L, t_L, s_H, t_H) is *incentive feasible* if

$$U_B(s_i, t_i \mid \beta_i) \geq U_B(s_j, t_j \mid \beta_i), \qquad \forall i, j \in \{L, H\} \text{ with } i \neq j \tag{3.22}$$

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

$$U_i\left(s_j', t_j' \mid \beta_j\right) \geq U_i(s_j, t_j \mid \beta_j), \qquad \forall j \in \{L, H\} \text{ and } \forall i \in \{A, B\}, \tag{3.23}$$

[10] See Holmström and Myerson (1983) for an elaborate discussion of the concept of efficiency in an incomplete information environment.

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.

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 (3.23).[11] And in a separating equilibrium the signaling mechanism does not affect the low type's savings choice (cf. Lemma 2). However, in a separating equilibrium the high type's choice is distorted upwards: $s_H^* > 0$; thus, it is conceivable that s_H^* 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. 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," given that μ tends to 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

$$U_i\left(s_H^{'}, t_H^{'} \mid \beta_H\right) \geq U_i\left(s_H, t_H \mid \beta_H\right), \qquad \forall i \in \{A, B\}, \tag{3.24}$$

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:

$$s_H^* \in \arg\max_{s\in[0,\omega]} U_B\left(s, t_H^* \mid \beta_H\right) \tag{3.25}$$

[11] Condition (3.23) 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 (3.23) 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 $\beta_L \neq \beta_H$.

and

$$t_H^* \in \arg\max_{t\in[0,\omega]} U_A\left(s_H^*, t \mid \beta_H\right). \tag{3.26}$$

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. It turns out that when

$$\alpha > \frac{\gamma}{1+\beta_H}, \tag{3.27}$$

conditions (3.25) and (3.26) are met only for

$$s_H^* = \frac{\gamma - \alpha\left(1-\beta_H\right)}{\gamma + \alpha\left(1+\beta_H\right)}\omega \equiv \widetilde{s}_I \tag{3.28}$$

and

$$t_H^* = \frac{\alpha\left(1+\beta_H\right) - \gamma}{\gamma + \alpha\left(1+\beta_H\right)}\omega \equiv \widetilde{t}_I. \tag{3.29}$$

(The assumption $\alpha < \frac{\gamma}{1-\beta_L}$ guarantees that $\widetilde{s}_I > 0$, and condition (3.27) guarantees that $\widetilde{t}_I > 0$.) And when

$$\alpha \leq \frac{\gamma}{1+\beta_H}, \tag{3.30}$$

conditions (3.25) and (3.26) are met only for

$$s_H^* = \frac{\beta_H}{1+\beta_H}\omega \equiv \widetilde{s}_{II} \tag{3.31}$$

and $t_H^* = 0 \equiv \widetilde{t}_{II}$. One may verify that when (3.27) holds, $\left(s_L^*, t_L^*, \widetilde{s}_I, \widetilde{t}_I\right)$ is indeed ex post incentive efficient for the high type; and when (3.30) holds, $\left(s_L^*, t_L^*, \widetilde{s}_{II}, \widetilde{t}_{II}\right)$ is indeed ex post incentive efficient for the high type.

To start with, let us assume that (3.27) holds, so that $\left(s_L^*, t_L^*, \widetilde{s}_I, \widetilde{t}_I\right)$ is ex post incentive efficient for the high type. Given this assumption, let us investigate whether we can sustain $(s_H^*, t_H^*) = \left(\widetilde{s}_I, \widetilde{t}_I\right)$ as part of an outcome of a separating equilibrium of type I.[12] For this to be possible we must have $\widetilde{s}_I \in \Omega_I^{sep}$. It turns out that a sufficient condition for this is that β_L is very close to zero. To see this, first note that if β_L tends to zero, Assumption 1 implies that $\alpha < \gamma$. This

[12] Clearly, since $\widetilde{t}_I > 0$, we cannot sustain $(s_H^*, t_H^*) = \left(\widetilde{s}_I, \widetilde{t}_I\right)$ as part of a separating type II equilibrium.

in turn implies that $\alpha\beta_H < \gamma$, which means that condition 1 in Proposition 3a is not met. Moreover, if β_L tends to zero, inequality (3.13) holds[13] while inequality (3.14) does not hold. This means that nor is condition 2 in Proposition 3a met, but condition 3 is. Hence $\Omega_I^{sep} = \left[s_I', \frac{\alpha\beta_H\omega}{\gamma}\right]$. By using the definition of s_I', one may show that $\widetilde{s}_I \geq s_I'$ is equivalent to

$$\log\left[\frac{\gamma+\alpha\left(1+\beta_H\right)}{2\alpha}\right] + \beta_L \log\left[\frac{\beta_L\left[\gamma+\alpha\left(1+\beta_H\right)\right]}{2\beta_H\left(\gamma+\alpha\beta_L\right)}\right] \geq 0. \tag{3.32}$$

And $\widetilde{s}_I \leq \frac{\alpha\beta_H\omega}{\gamma}$ is equivalent to

$$\gamma\left(\gamma-\alpha\right) \leq \alpha^2\beta_H\left(1+\beta_H\right). \tag{3.33}$$

Condition (3.32) is met for β_L sufficiently close to zero. To see this, note that as β_L tends to 0, the left-hand side of condition (3.32) tends to $\log\left[\frac{\gamma}{2\alpha} + \frac{1}{2}\left(1+\beta_H\right)\right]$, which is strictly positive due to Assumption 1. Hence, by continuity, the condition is met for some strictly positive β_L. Under the assumption that (3.27) holds, condition (3.33) is also met. To see this, first note that (3.27) is equivalent to $\beta_H > \frac{\gamma}{\alpha} - 1$ and that the right-hand side of (3.33) is strictly increasing in β_H. Also, evaluated at $\beta_H = \frac{\gamma}{\alpha} - 1$, condition (3.33) holds with equality; hence it holds for all β_H satisfying (3.27).

Let us now assume that (3.30) holds, so that $\left(s_L^*, t_L^*, \widetilde{s}_{II}, \widetilde{t}_{II}\right)$ is ex post incentive efficient for the high type, and let us investigate whether we can sustain $\left(s_H^*, t_L^*\right) = \left(\widetilde{s}_{II}, \widetilde{t}_{II}\right)$ as part of an outcome of a separating equilibrium of type II. Consider again the case where β_L is very small. It follows from conditions (3.15) and (3.16) that when β_L tends to 0, the conditions under 3 in Proposition 3b are met (the condition $\alpha\beta_H < \gamma$ is implied by (3.30)). Hence $\Omega_{II}^{sep} = \left[\frac{\alpha\beta_H\omega}{\gamma}, s_{II}''\right]$. The assumption that (3.30) holds guarantees that $\widetilde{s}_{II} \geq \frac{\alpha\beta_H\omega}{\gamma}$. It thus suffices to show that if β_L tends to 0, $\widetilde{s}_{II} \leq s_{II}''$. However, it follows from the definition of s_{II}'' in equation (3.18) that as β_L tends to 0, s_{II}'' tends to ω. This means that when (3.30) holds, we can indeed sustain $s_H^* = \widetilde{s}_{II}$ as part of an outcome of a separating equilibrium of type II.

In summary we have the following proposition.

Proposition 5. *There is an $\varepsilon > 0$ such that if $\beta_L \in (0, \varepsilon]$, then an allocation that is ex post incentive efficient for the hight type can be sustained as the outcome of a separating equilibrium.*

[13] To verify this one has to make use of the fact that $\lim_{\beta_L \to 0} \beta_L \log\left(\beta_L\right) = 0$.

One example of the case where $\widetilde{s}_I \in \Omega_I^{sep} = \left[s_I^{'}, s_I^{''}\right]$ is illustrated in Figure 2. There, the parameter values are $\alpha = \gamma = \omega = 1$, $\beta_L = \frac{1}{2}$, and $\beta_H = \frac{3}{4}$. As one can see from the figure, this gives a value of $\widetilde{s}_I$ which is roughly in the middle of $s_I^{'}$ and $s_I^{''}$, $\widetilde{s}_I = \frac{3}{11}$. It is also easy to see that inequality (3.27) is satisfied for these parameter values. Hence there is a separating equilibrium where $s_L^* = 0$ and $s_H^* = \frac{3}{11}$, which implies that (cf. equation (2.3)) $t_L^* = \frac{1}{3}$ and $t_H^* = \frac{3}{11}$. This outcome is ex post incentive efficient for the high type. Moreover, the equilibrium remains an equilibrium when we let the prior probability that B is the high type approach unity, and we may thus get arbitrarily close to efficiency from an ex ante perspective.

3.4. Refinements

In the previous subsection 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, for this to be possible, we must choose beliefs off the equilibrium path in a particular manner. For instance, suppose that $\widetilde{s}_I$ belongs to the interior of $\Omega_I^{sep} = \left[s_I^{'}, s_I^{''}\right]$ as in Figure 2. Then, in order to sustain $\widetilde{s}_I$ as part of a separating equilibrium, A must assign positive probablity to B's being the low type when observing any saving level $s \in \left[s_I^{'}, \widetilde{s}_I\right)$; otherwise the high type would have an incentive to deviate to any such savings level s. However, it is clear from Figure 2 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 savings level $s \in \left(s_I^{'}, s_I^{''}\right]$. Hence one may argue that any beliefs assigning positive probability to the low type saving $s \in \left(s_I^{'}, s_I^{''}\right]$ are unreasonable. But if we do not allow such beliefs the high type will deviate, and we cannot sustain any $s \in \left(s_I^{'}, s_I^{''}\right]$ as the high type's savings level in an equilibrium. In particular we cannot sustain the efficient saving level $\widetilde{s}_I$ as part of a separating equilibrium. The only saving level of the high type that survives this refinement is $s_H^* = s_I^{'}$.

The kind of reasoning described in the previous pharagraph would not have any bite on the set of separating equilibria if we assumed more than two types. However, there are other, stronger refinements that rule out all pooling equilibria as well as all separating equilibria except for the one where $s_H^* = s_I^{'}$ (see, e.g., Cho and Sobel (1990)). This means that if we employ such a refinement and if the high type's saving level $s_H^* = s_I^{'}$ does not happen to equal $\widetilde{s}_I$, we can no longer obtain the "almost efficiency" result. Still, since $s_I^{'} > 0$, the undersaving effect in the standard Samaritan's dilemma model is mitigated. Also, there is reason to be agnostic about this kind of stronger refinements. One such reason is that they are sensitive to the so-called Stiglitz critique:[14] "The argument is that players can 'have' their equilibrium values, and defections are to be thought of as a reasoned attempt to do better. (...) If a particular equilibrium is indeed suspect, then its values cannot be taken for granted."

4. Concluding Remarks

The main finding of this paper is that by adding an arbitrarily small amount of incomplete information to a standard Samaritan's dilemma model, a large number of different outcomes can be sustained as equilibria, and these outcomes typically involve both undersaving and oversaving. Although full efficiency cannot be obtained, one may get arbitrarily close to ex post incentive efficiency. The "almost efficiency" result is sensitive to refinements, but the result that the undersaving is mitigated continues to hold.

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. 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 are sensitive to the complete information assumption.

5. Appendix

5.1. Proof of Lemma 1

In order to prove Lemma 1 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 $\kappa : (0,1) \to \Re$ be defined by

$$\kappa(\beta) = \log(1+\beta) - \frac{\beta}{\sqrt{1+\beta}}. \tag{5.1}$$

[14] This formulation of the critique is borrowed from Kreps and Sobel (1994, p. 857).

I claim that $\kappa(\beta) < 0$. To prove this, first note that $\lim_{\beta \to 0} \kappa(\beta) = 0$. It thus suffices to show that $\kappa'(\beta) < 0$. Differentiating $\kappa(\beta)$ yields

$$\kappa'(\beta) = (1+\beta)^{-1} - (1+\beta)^{-\frac{1}{2}} + \frac{\beta}{2}(1+\beta)^{-\frac{3}{2}}. \tag{5.2}$$

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

$$\begin{aligned} \kappa'(\beta) &= \left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]^{-1}\left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right] \times \\ &\quad \times \left[(1+\beta)^{-1} - (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right] \\ &= \left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]^{-1}\left[(1+\beta)^{-2} - (1+\beta)^{-1}\left[\frac{2+\beta}{2(1+\beta)}\right]^{2}\right] \\ &= -\left[(1+\beta)^{-1} + (1+\beta)^{-\frac{1}{2}}\left[\frac{2+\beta}{2(1+\beta)}\right]\right]^{-1}\frac{\beta^2}{4(1+\beta)^3}. \end{aligned} \tag{5.3}$$

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

Now let the function $\xi : (0,1) \to \Re$ be defined by

$$\xi(\beta) = (1+\beta)^{\frac{1+\beta}{\beta}}. \tag{5.4}$$

I claim that $\xi'(\beta) > 1$. In order to prove this, first differentiate ξ:

$$\xi'(\beta) = \frac{\xi}{\beta^2}\left[\beta - \log(1+\beta)\right]. \tag{5.5}$$

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

$$\xi''(\beta) = \frac{\beta^2 \xi' - 2\beta\xi}{\beta^4}\left[\beta - \log(1+\beta)\right] + \frac{\xi}{\beta(1+\beta)}. \tag{5.6}$$

Substituting the expression for ξ' in (5.5) into (5.6) and then rewriting yields

$$\xi''(\beta) = \frac{\xi}{\beta^4}\left[\log(1+\beta) + \frac{\beta}{\sqrt{1+\beta}}\right]\left[\log(1+\beta) - \frac{\beta}{\sqrt{1+\beta}}\right]. \tag{5.7}$$

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 1. Differentiating the expression in (2.6) yields

$$\varphi_1'(\beta,\gamma) = -\frac{\gamma\left[\xi'(\beta)-1\right]}{\left[\xi(\beta)-\beta\right]^2}. \tag{5.8}$$

Thus, by the claim that $\xi'(\beta) > 1$, $\varphi_1'(\beta,\gamma) < 0$. □

5.2. Proof of Proposition 1

In order to prove Proposition 1 we must investigate the case where $\alpha\beta < \gamma$, since this was not done in subsection 2.2. When $\alpha\beta < \gamma$, B's indirect utility is given by either one of the two lines in equation (2.4), and which one it is depends on the choice of s itself. Let Θ^1 be defined by

$$\Theta^1 = \max_{s\in\left[0,\frac{\alpha\beta}{\gamma}\omega\right)} \log(\omega - s) + \beta\log(\omega+s) + \beta\log\left(\frac{\alpha\beta}{\gamma+\alpha\beta}\right). \tag{5.9}$$

That is, Θ^1 is the highest utility B can obtain if he is constrained to choose some $s \in \left[0, \frac{\alpha\beta}{\gamma}\omega\right)$. Denote the solution to the maximization problem in (5.9) (it does exist) by s^1. Since $\beta < 1$, we have $s^1 = 0$. Hence

$$\Theta^1 = (1+\beta)\log\omega + \beta\log\left(\frac{\alpha\beta}{\gamma+\alpha\beta}\right) \tag{5.10}$$

Similarly define Θ^2:

$$\Theta^2 = \max_{s\in\left[\frac{\alpha\beta}{\gamma}\omega,\omega\right]} \log(\omega-s) + \beta\log(s). \tag{5.11}$$

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

$$s^2 = \begin{cases} \frac{\beta}{1+\beta}\omega & \text{for } \alpha \le \frac{\gamma}{1+\beta} \\ \frac{\alpha\beta}{\gamma}\omega & \text{otherwise.} \end{cases} \tag{5.12}$$

Thus

$$\Theta^2 = \begin{cases} (1+\beta)\log\left(\frac{\omega}{1+\beta}\right) + \beta\log(\beta) & \text{for } \alpha \le \frac{\gamma}{1+\beta} \\ (1+\beta)\log(\omega) + \log\left(1-\frac{\alpha\beta}{\gamma}\right) + \beta\log\left(\frac{\alpha\beta}{\gamma}\right) & \text{otherwise.} \end{cases} \tag{5.13}$$

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

$$\Theta^1 \gtreqless \Theta^2 \Leftrightarrow \alpha \gtreqless \varphi(\beta, \gamma). \tag{5.14}$$

Thus, if $\alpha > \varphi(\beta,\gamma)$, then $s^* = 0$; and if $\alpha < \varphi(\beta,\gamma)$, then $s^* = \frac{\beta}{1+\beta}\omega$. If $\alpha = \varphi(\beta,\gamma)$, then there are two solutions two B's optimization problem, $s^* = 0$ and $s^* = \frac{\beta}{1+\beta}\omega$. From equation (2.3) one obtains the result that when $s^* = 0$, $t^* = \frac{\alpha\beta\omega}{\gamma+\alpha\beta}$; and when $s^* = \frac{\beta}{1+\beta}\omega$, then $t^* = 0$.

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

$$\begin{aligned} \Theta^1 &> \Theta^2 \Leftrightarrow \log\left(1 - \frac{\alpha\beta}{\gamma}\right) + \beta\log\left(1 + \frac{\alpha\beta}{\gamma}\right) < 0 \\ &\Leftrightarrow (1-\beta)\log\left(1 - \frac{\alpha\beta}{\gamma}\right) + \beta\log\left[1 - \left(\frac{\alpha\beta}{\gamma}\right)^2\right] < 0 \end{aligned} \tag{5.15}$$

Inequality (5.15) always holds since $\beta < 1$ and $\alpha\beta < \gamma$. Hence, we have $s^* = 0$ and $t^* = \frac{\alpha\beta\omega}{\gamma+\alpha\beta}$. Together with the analysis in subsection 2.2, this completes the proof of Proposition 1. □

5.3. Proof of Proposition 2

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

$$\max_{(t,s)\in[0,\omega]^2} U_A(s,t). \tag{5.16}$$

(This is the solution as long as $\alpha \leq \frac{\gamma}{1+\beta}$.) That is, this allocation is A's first-best choice. But for $\alpha < \varphi(\beta,\gamma)$, $(s,t) = \left(\frac{\beta}{1+\beta}\omega, 0\right)$ is also the outcome of the unique subgame-perfect Nash equilibrium; hence, this outcome is Pareto efficient. Similarly, for $\alpha = \varphi(\beta,\gamma)$, $(s,t) = \left(\frac{\beta}{1+\beta}\omega, 0\right)$ is one of two possible outcomes and it is Pareto efficient. Now assume that $\alpha \geq \frac{\gamma}{1-\beta}$. 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 (5.16). Again, this is also the outcome of the unique subgame-perfect Nash equilibrium, so the outcome is Pareto efficient. It remains to show that the allocation $(s,t) = \left(0, \frac{\alpha\beta}{\gamma+\alpha\beta}\omega\right)$ is not Pareto efficient for any $\alpha \in \left[\varphi(\beta,\gamma), \frac{\gamma}{1-\beta}\right)$. Suppose that $\alpha < \frac{\gamma}{1-\beta}$. Differentiate U_B with respect to s and evaluate at $(s,t) = \left(0, \frac{\alpha\beta}{\gamma+\alpha\beta}\omega\right)$:

$$\frac{\partial U_B(s,t)}{\partial s} \Big|_{(s,t)=\left(0,\frac{\alpha\beta}{\gamma+\alpha\beta}\omega\right)} = \frac{1}{\alpha\omega}\left[\gamma - \alpha(1-\beta)\right] > 0. \tag{5.17}$$

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 $(s,t) = \left(0, \frac{\alpha\beta}{\gamma+\alpha\beta}\omega\right)$ is not Pareto efficient for any $\alpha \in \left[\varphi(\beta,\gamma), \frac{\gamma}{1-\beta}\right)$. □

5.4. Proof of Lemma 2

Suppose, per contra, that $\alpha > \varphi(\beta_L,\gamma)$ and that $s_L^* > 0$ in a separating equilibrium. To start with, consider the case where $s_L^* \in \left(0, \frac{\alpha\beta_L}{\gamma}\omega\right]$. Then the low type receives a transfer from A according to the first line in equation (2.3) but with β_L substituted for β. The low type's utility is accordingly given by (cf. the first line of equation (2.4)):

$$V' = \log(\omega - s_L^*) + \beta_L \log(\omega + s_L^*) + \beta_L \log\left(\frac{\alpha\beta_L}{\gamma+\alpha\beta_L}\right). \tag{5.18}$$

However, if the low type instead chose $s = 0$, he would receive a transfer of at least $\frac{\alpha\beta_L\omega}{\gamma+\alpha\beta_L}$, which would give him the following utility (cf. again the first line of equation (2.4)):

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

V is strictly greater than V' for all $s_L^* \in \left(0, \frac{\alpha\beta_L}{\gamma}\omega\right]$, since V' is strictly decreasing in s_L^*. Now consider the case where $s_L^* \in \left(\frac{\alpha\beta_L}{\gamma}\omega, \omega\right]$. Then the low type receives a transfer from A according to the second line in equation (2.3) but with β_L substituted for β. However, as long as $\alpha > \varphi(\beta_L,\gamma)$, the low type is strictly better off from choosing $s = 0$ than from choosing any $s_L^* \in \left(\frac{\alpha\beta_L}{\gamma}\omega, \omega\right]$. This follows from the proof of Proposition 1. We thus have a contradiction, which proves the lemma. □

5.5. Proof of Proposition 3b

In order to have a separating equilibrium where $s_H^* \in \left[\frac{\alpha\beta_H\omega}{\gamma}, \omega\right]$, there are some necessary conditions. An obvious necessary condition is that $\alpha\beta_H < \gamma$, since otherwise either the set $\left[\frac{\alpha\beta_H\omega}{\gamma}, \omega\right]$ is empty or it equals the singleton $\{\omega\}$ (and $s_H^* = \omega$ can clearly not be an equilibrium). It must also be that the low type does not have an incentive to mimic the high type:

$$U_B\left(0, \frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L} \mid \beta_L\right) \geq U_B\left(s_H^*, 0 \mid \beta_L\right) \Leftrightarrow$$

$$\log\left(\omega - s_H^*\right) + \beta_L \log\left(s_H^*\right) \leq \log\left(\omega\right) + \beta_L \log\left(\frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L}\right). \tag{5.19}$$

Moreover, the high type must not have an incentive to mimic the low type:

$$U_B\left(s_H^*, 0 \mid \beta_H\right) \geq U_B\left(0, \frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L} \mid \beta_H\right) \Leftrightarrow$$

$$\log\left(\omega - s_H^*\right) + \beta_H \log\left(s_H^*\right) \geq \log\left(\omega\right) + \beta_H \log\left(\frac{\alpha\beta_L\omega}{\gamma + \alpha\beta_L}\right). \tag{5.20}$$

Together, these three conditions are also sufficient for a separating equilibrium of type II to exist. To see this, let A's out-of-equilibrium beliefs be given by $\widetilde{\mu}(s) = 1$ for all $s \geq s_H^*$ and $\widetilde{\mu}(s) = 0$ for all $s < s_H^*$. These beliefs are consistent with the equilibrium requirements.

Let Ω_{II}^{sep} be the set of s_H^* that satisfy $s_H^* \in \left[\frac{\alpha\beta_H\omega}{\gamma}, \omega\right]$ and conditions (5.19) and (5.20). It remains to investigate when Ω_{II}^{sep} is non-empty and, in those cases, what values of s_H^* that belong to this set. To start with I will show that condition (5.19) is not "binding" in the following sense:

Claim 1. Suppose that $\alpha > \varphi(\beta_L, \gamma)$. Then

$$\{s_H^* \in [0, \omega] \mid \text{inequality (5.19) holds}\} = [0, \omega] \tag{5.21}$$

for all ω, β_L, α, and γ.

Proof of Claim 1: The left-hand side of inequality (5.19) obtains its maximum at $s_H^* = \frac{\beta_L\omega}{1+\beta_L}$. Thus, it suffices to show that if $\alpha > \varphi(\beta_L, \gamma)$, inequality (5.19) holds for $s_H^* = \frac{\beta_L \omega}{1+\beta_L}$. Substituting $s_H^* = \frac{\beta_L \omega}{1+\beta_L}$ into inequality (5.19) and rewriting yields

$$(1+\beta_L)\log(1+\beta_L) - \beta_L \log\left(\frac{\gamma}{\alpha} + \beta_L\right) \geq 0. \tag{5.22}$$

Since $\alpha > \varphi(\beta_L, \gamma)$, inequality (5.22) holds if $\chi \geq 0$ holds, where

$$\chi = (1+\beta_L)\log(1+\beta_L) - \beta_L \log\left(\frac{\gamma}{\varphi(\beta_L,\gamma)} + \beta_L\right). \tag{5.23}$$

However, by using the definition of φ, one may easily verify that $\chi = 0$. This proves Claim 1.

Now turn to condition (5.20). This condition is satisfied for *some* $s_H^* \in [0, \omega]$ if and only if it is satisfied for $s_H^* = \frac{\beta_H \omega}{1+\beta_H}$, since this value of s_H^* maximizes the left-hand side of inequality (5.20). Substituting $s_H^* = \frac{\beta_H \omega}{1+\beta_H}$ into inequality (5.20) and simplifying yields condition (3.15). Clearly, if inequality (3.15) does not hold, Ω_{II}^{sep} is empty; and if inequality (3.15) holds with equality, $\Omega_{II}^{sep} = \left\{\frac{\beta_H \omega}{1+\beta_H}\right\} \cap \left[\frac{\alpha\beta_H\omega}{\gamma}, \omega\right]$. In the latter case, $\Omega_{II}^{sep} = \left\{\frac{\beta_H \omega}{1+\beta_H}\right\}$ if $\frac{\alpha\beta_H\omega}{\gamma} \leq \frac{\beta_H\omega}{1+\beta_H} \Leftrightarrow \alpha \leq \frac{\gamma}{1+\beta_H}$ (the necessary condition $\alpha\beta_H < \gamma$ is automatically taken care of since this is implied by $\alpha \leq \frac{\gamma}{1+\beta_H}$), and Ω_{II}^{sep} is empty if $\alpha > \frac{\gamma}{1+\beta_H}$. The former of the last two cases corresponds to condition 1 in the Proposition.

Now suppose that condition (3.15) holds with strict inequality. (This assumption is maintained throughout this pharagraph.) Then inequality (5.20) is satisfied for all $s_H^* \in \left[s_{II}', s_{II}''\right]$, where s_{II}' and s_{II}'' are defined in equations (3.17) and (3.18) and where accordingly $s_{II}' \in \left(0, \frac{\beta_H\omega}{1+\beta_H}\right)$ and $s_{II}'' \in \left(\frac{\beta_H\omega}{1+\beta_H}, \omega\right)$. This is true since the left-hand side of inequality (5.20) is stricly concave in s_H^* and obtains it maximum at $s_H^* = \frac{\beta_H\omega}{1+\beta_H}$. Hence, in this case, $\Omega_{II}^{sep} = \left[s_{II}', s_{II}''\right] \cap \left[\frac{\alpha\beta_H\omega}{\gamma}, \omega\right]$. Depending on how $\frac{\alpha\beta_H\omega}{\gamma}$ is related to s_{II}' and s_{II}'', there are three cases to consider. First, if $s_{II}' > \frac{\alpha\beta_H\omega}{\gamma}$, then $\Omega_{II}^{sep} = \left[s_{II}', s_{II}''\right]$. Second, if $\frac{\alpha\beta_H\omega}{\gamma} \in \left[s_{II}', s_{II}''\right]$, then $\Omega_{II}^{sep} = \left[\frac{\alpha\beta_H\omega}{\gamma}, s_{II}''\right]$. Third, if $s_{II}'' < \frac{\alpha\beta_H\omega}{\gamma}$, then Ω_{II}^{sep} is empty. If inequality (5.20) *does not* hold for $s_H^* = \frac{\alpha\beta_H\omega}{\gamma}$, then we have either the first case or the third (to verify this, the reader may want to draw a figure). Substituting $s_H^* = \frac{\alpha\beta_H\omega}{\gamma}$ into inequality (5.20), reverting the inequality sign (i.e., $<$ instead of $\geq$), and simplifying yield condition (3.16). Hence, the first case obtains if and only if inequality (3.16) holds and $\frac{\alpha\beta_H\omega}{\gamma} < \frac{\beta_H\omega}{1+\beta_H} \Leftrightarrow \alpha < \frac{\gamma}{1+\beta_H}$ (the condition $\alpha\beta_H < \gamma$ is again implied by $\alpha < \frac{\gamma}{1+\beta_H}$). This case corresponds to condition 2 in the Proposition. The second case obtains if and only if inequality (3.16) does not hold and $\alpha\beta_H < \gamma$. This case corresponds to condition 3 in the Proposition.

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

Cho, I.-K. and J. Sobel (1990), Strategic stability and uniqueness in signaling games, Journal of Economic Theory 50, 381-413.

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

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

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

Kreps, D. K. and Sobel (1994), Signalling, in "Handbook of Game Theory, Volume 2," edited by R.J. Aumann and S. Hart.

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

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

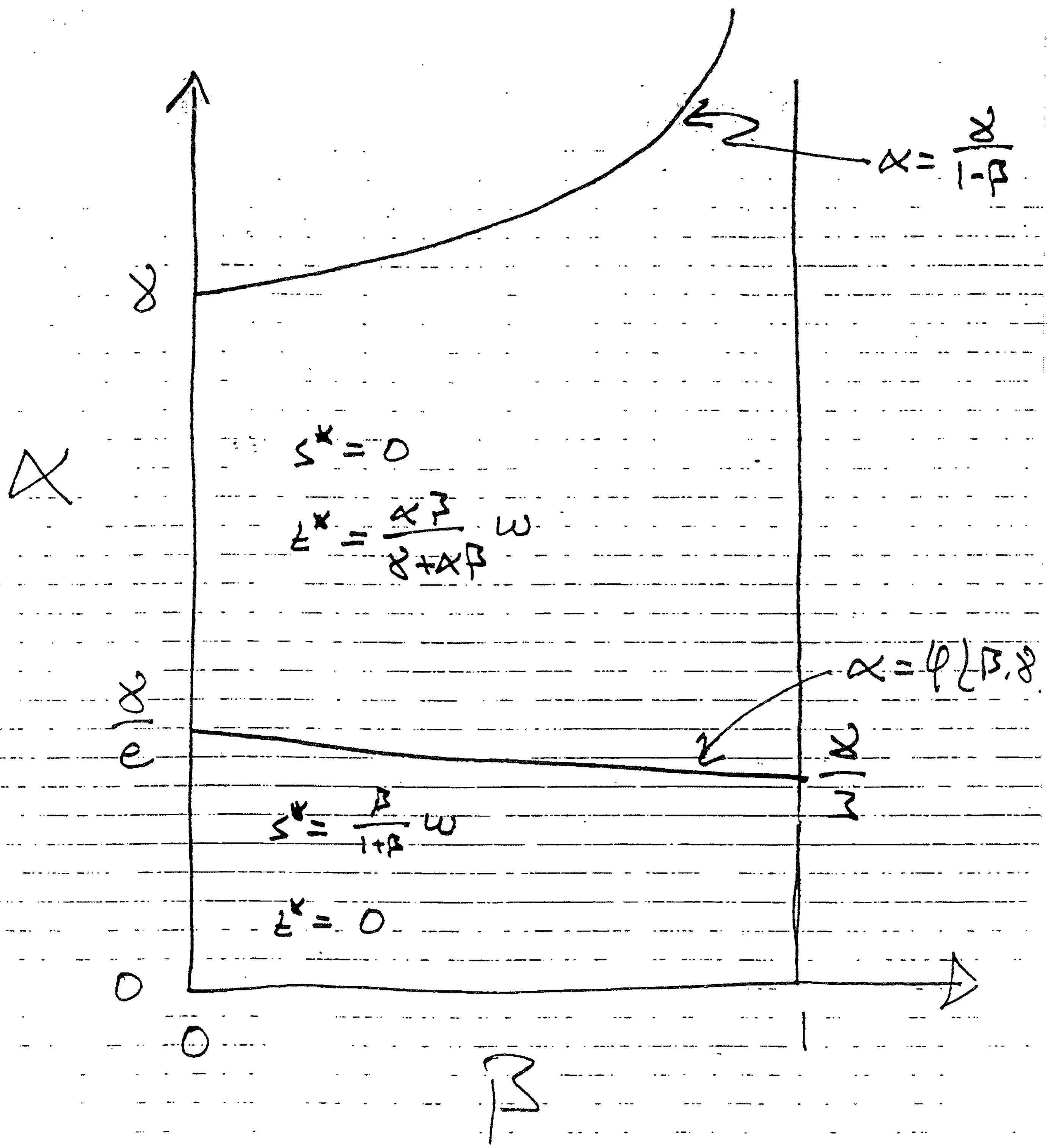
Spence, M. (1973), Job market signaling, Quarterly Journal of Economics 87, 355-374.

Spence, M. (1974), *Market signaling: Informational transfers in hiring and related screeing 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.

van Damme, E. (1991), Stability and perfection of Nash equilibria, Second, revised and enlarged edition, Berlin: Springer-Verlag.

Varian, H.R. (1994), Sequential contributions to public goods, Journal of Public Economics 53, 165-186.



$\alpha = \frac{\gamma}{1-\beta}$
γ
α
$s^* = 0$
$t^* = \frac{\alpha\beta}{\gamma+\alpha\beta} w$
$\alpha = \varphi(\beta, \gamma)$
$\frac{\gamma}{e}$
$\frac{\gamma}{3}$
$s^* = \frac{\beta}{1+\beta} w$
$t^* = 0$
0
0
1
β

Figure 1

Figure 2

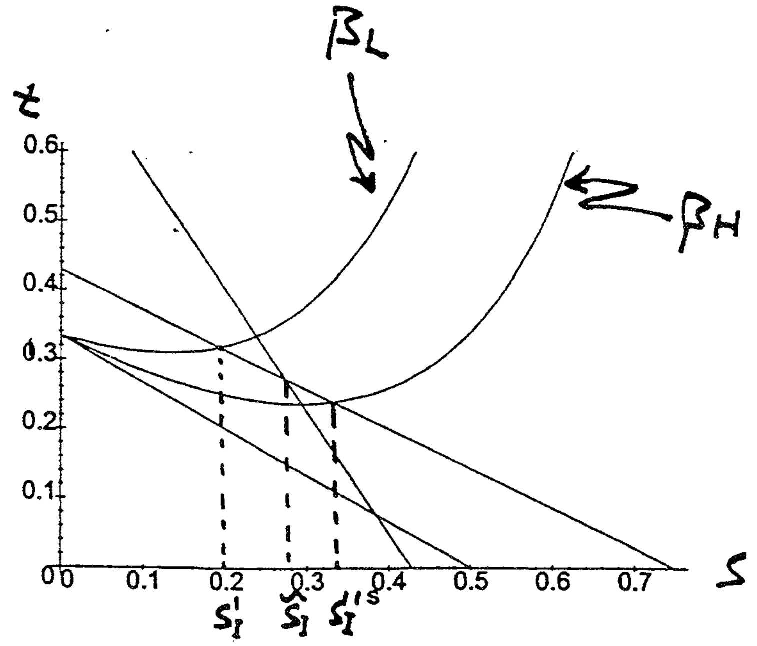


Figure 2: Separating equilibria.
$(\alpha = \gamma = \omega = 1, \beta_L = \frac{1}{2}, \beta_H = \frac{3}{4})$