Implementation by decent mechanisms*

Jernej Čopič and Clara Ponsati

April 2003

Abstract

We address the design of optimal mechanisms for bargaining problems subject to incomplete information on reservation values. We characterize decent rules, those that are Pareto Optimal in the constrained set of rules satisfying strategy profness, individual rationality and weak efficiency - a mild requirement of ex-post effciciency. We prove that decent rules are implemented by the Filtered Demands game, a natural procedure for dynamic bargaining under incomplete information where players submit their claims over time to a passive agent who minimizes transmission of information between the players.

*Čopič is at HSS Division, California Institute of Technology (jernej@caltech.edu) and Ponsatí is at Institut d'Anàlisi Economica -CSIC and CODE-Universtiat Autònoma de Barcelona (clara.ponsati@uab.es). This research was conducted in part while Ponsati was at the University of Toronto and Čopič was at the Universitat Autònoma de Barcelona. Financial support to Ponsati from DURSI grant SGR2001-00162 and DGCYT grant PB98-870; and to Čopič from the US-Spain cooperation Committee is gratefully acknowledged. We benefited from comments and discussions with J.Bruna, M. Jackson, J. Ledyard, V. Skreta and seminar audiences at Cambridge, Caltech, Edinburgh, UAB, UCSD, UPF and WSC-Berlin.

1 Introduction

It is well known that in environments where agents have private information and act strategically, achieving non-dictatorial first best allocations is often impossible. This is true in dominant strategy settings, as well as in Bayesian ones.[1] How efficient can a given incentive compatible *social choice rule* be? When addressing this question, it is important to distinguish between the rules that do well in the *ex-ante* sense,[2] those that do well in the *interim*, and the ones that do well *ex-post*. A rule that does well *ex-ante*, might do very poorly *ex-post*, and *vice-versa*.

In this paper we investigate rules that satisfy an *ex-post* efficiency property which we call *weak efficiency (WE).* We take a simple environment: a bargaining problem where two agents must share a unit of surplus and have private reservation shares (agents' types). This problem is easily transformed into bilateral trade or into the problem of sharing the cost of a public good.[3] In the bargaining context, possible outcomes are defined as all feasible divisions of the surplus union the disagreement point. The agents preferences over outcomes are determined by concave utilities over net surplus - the excess between an agents' share and her (privately known) reservation share.[4] A rule is a social choice function, assigning shares of the surplus to the agents,

[1] For dominant strategy environments see Hurwicz [1972], Gibbard]1973], Satterthwaite [1975], Green and Laffont [1977]. For Bayesian environments see Myerson and Satterthwaite [1983]. Corchon [1996] presents a unified treatment. In economic environments with quasi-linear preferences it has even been shown that requiring only Bayesian incentive compatibility doesn't provide any improvement in efficiency, as compared to the dominant strategy incentive compatibility. See Williams [1999] and Mookherjee and Reichelstein [1992].

[2] A non-cooperative game of incomplete information can be divided into three temporal stages. At the *ex-ante* stage each agent knows only the distributions of types of all agents, including himself. At the *interim* stage each agent knows her own type but still knows only the distribution of types of her opponents. At the *ex-post* stage the types of all agents are common knowledge.

[3] More precisely, in the bilateral trade context, the problem can be translated into one where the buyer has a valuation of the good and the seller has a cost of producing the good. Similarly, in the cost-sharing of a public good, the problem is translated into one where agents have private valuations of the public good and have to share its cost.

[4] In the language of social choice, the agents preferences lie in the restricted domain determined by each pair of concave utilities over net surplus. Thus agents private information on preferences is equivalent to private information on reservation values.

possibly in a probabilistic way, for each pair of reports on reservation shares. Weak efficiency imposes that whenever agents' reservation shares are compatible, the probability that all the surplus is allocated is strictly positive. We focus on rules satisfying weak efficiency and the strongest incentive requirements, strategy proofness and *ex-post* individual rationality. We call the rules that are Pareto optimal in this constrained set the *decent bargaining rules*.[5]

We fully characterize decent rules for all environments where agents have concave utilities; thus we do not require risk-neutrality. Decent rules are always probabilistic, the probability of implementing the disagreement point being the tool to elicit truthful revelation.[6] An important characteristic of the decent rules is that the outcomes depend non-trivially on agents' reports. Our characterization implies that, whenever it exists, the decent rule is always unique. We prove that a decent *net surplus rule*, a rule in which shares and probabilities depend only on $1-$(the sum of the revealed types), is necessarily one where agents obtain their reservation plus a fix portion of the revealed net surplus. Furthermore, we completely identify the class of utilities for which such decent rule exists: it includes utilities with constant relative risk aversion (CRRA), but does not include for example, exponential utilities.

For the domain of risk-neutral preferences, Hagerty and Rogerson [1987] prove in the context of bilateral trade, that any strategy proof mechanism yields the same *ex-ante* total surplus than a mechanism whose operation is independent of agents' reports. Such mechanisms, in the bargaining context we call them *posted split* mechanisms,[7] operate in the following way. First, a distribution function is announced, and a split of pie is selected according to a random draw from that distribution. The split is then announced publicly, and if both agents agree to it, it is implemented. Otherwise the agents get their disagreement payoffs. It is easy to see that to maximize *ex-ante* payoffs the posted split should be selected with a degenerate distribution; for any given distribution of agents' types, the mechanism must pick the split that

[5]This is unrelated to decency as a constraint on the behavior of agents as in Herrero and Corchon [2003].

[6]This follows from inexistence of efficient strategy-proof mechanisms which is implied by Myerson and Satterthwaite [1982].

[7]In the context of bilateral trade Hagerty and Rogerson [1987] call these mechanisms posted-price mechanisms. The revelation principle for dominant strategies (see Corchon [1996]) allows us to identify rules with direct mechanisms that implement them.

maximizes expected gains from trade.[8] Because the decent mechanism is necessarily probabilistic, under risk neutrality it must yield the same *ex-ante* payoffs as a posted split rule with a non-degenerate distribution. This payoff equivalence no longer holds under risk aversion. In fact, if agents are sufficiently risk averse, the ex-ante payoffs of the decent rule dominate those of any posted split rule.

In general, the decent rule does not maximize the sum of expected payoffs. However, designing ex-ante optimal rules requires knowledge of the distribution of agents' types. Without this information, the efficiency of rules that are not weakly efficient - posted split rules, for example - will depend on luck alone, and their inefficiency can be severe. A decent mechanism, although possibly not optimal *ex-ante*, will always be decent in terms of efficiency, regardless of the distribution of agents' types.[9]

Another reason we may be interested in the decent rule is that implementing a social choice rule requires strong commitment by the agents. They have to obey the mechanism, and abstain from renegotiating the outcomes, even if it is revealed that mutually agreeable improvements were possible.[10] Posted split rules do not account for such renegotiation considerations because they operate at the *ex-ante* stage. In other words, in a posted split mechanism it may well happen that the two agents have mutually compatible shares which are not compatible with the announced division. Thus, the agents might be inclined to renegotiate. But if the outcomes can be renegotiated *ex-post*, and the agents know it *ex-ante*, this will change the agents' incentives. Weak efficiency can be interpreted as a weak requirement on renegotiation-proofness.

The motivation for weak efficiency is most apparent when a bargaining rules are interpreted as the direct revelation implementation of equilibria in dynamic bargaining games among impatient agents. In these settings, the probability of disagreement is equivalent to delay. Weak efficiency is

[8]The characterizations of strategy proof rules by Barberà and Jackson [1995] and Sprumont [1991] are also related. Barberà and Jackson [1995] show that in two-person economies strategy proofness and individual rationality imply that trade is at fixed proportions. For the division problems with single peaked preferences, when the disagreement is not an alternative, Sprumont [1991] proves that the uniform rule is the unique rule that is efficient, symmetric and strategy proof.

[9]See Section 7 for concrete examples.

[10]See Hurwicz [1983], Jackson and Palfrey [2001] and Rubinstein and Wolinsky [1992] for discussions on the problem of enforceability of mechanisms. On renegotiation see for instance Maskin and Moore[1999].

then the property that an agreement occurs, sooner or later, when types are compatible and never otherwise. In a decent mechanism, agents agree as soon as possible, while the correct incentives are preserved.[11] We formalize this interpretation in the second part of the paper, where we provide a natural dynamic game implementing decent rules. We call this game the *Filtered Demands* (FD) game.

The FD game is the simplest possible bargaining game in continuous time, its main feature being that the information flow between the agents is minimized. It can be envisioned as a market with a completely closed order book where both agents keep posting their demands. In continuous time, the agents keep sending their changing demands (they have a common discount function) to a central agent, *the Filter*. The Filter's role is to record these messages secretly, making them public only when the agreement is possible. Then the game ends with the agents obtaining their agreed shares. Thus, the agents recognize how much net surplus is available only when they reach an agreement, and at that moment they share all the remaining surplus. Hence, at no moment of agreement can they renegotiate to a better outcome *ex-post*.

In the FD game we define an equilibrium in regular strategies as a Bayesian equilibrium which is undominated and in which the agents' strategies satisfy some smoothness requirements. We prove that the regular Bayesian equilibria of the FD game implement precisely all the decent rules. The proof is implied by our result that no common prior is needed to play a regular equilibrium in the FD game.[12] The implementation result that we obtain is quite strong. First, there is a one-to-one correspondence between the regular equilibria of the FD game and the decent rules. Second, the FD game is in a way more general than the decent mechanisms. To construct a decent mechanism, the designer (as well as both agents) has to know the restricted domain of preferences (i.e. the forms of the utilities of both agents, but not

[11] Another way to address this is the following. If an agent wished a strategy proof mechanism that gave her as much as possible for all her types and in any possible draw of the opponent's types, then a decent rule is what she should subscribe to. It should thus come as no surprise that the more risk averse the agents are, the more efficient are the decent rules. See also the remark in Section 6.

[12] By the first order condition, all the regular equilibria of the FD game are belief-independent, hence they are ex-post equilibria (see Ledyard[1978] or Bergemann and Morris [2001]). The informational assumption that we need is the conditional independence of beliefs, also called the spanning property.

their private types). On the other hand, in the FD game, only the agents need to know the forms of each other's utility functions. The implementation via the FD game proves that the best design for a dynamic game implementing a strategy-proof, weakly efficient rule is to completely restrict the communication between the agents.

The FD game also provides a strong link to the literature on non-cooperative bargaining games with incomplete information. While a great deal is known about the bargaining games where two agents alternate offers, and there is one-sided incomplete information, characterizing the set of equilibria with two-sided incomplete information has remained an elusive task.[13] When impatient agents bargain non-cooperatively over time, agreements are delayed because efficient equilibria are impossible under two-sided incomplete information. In such bargaining process, agents must learn what aspirations are reasonable before they are ready for an agreement; learning requires communication and time. But learning is a double-edged sword. On one hand, as agreements are more easily attained when the parties know well the limits of what is agreeable, it is important that the agents credibly communicate what they cannot accept. On the other hand, when an agent learns of her opponent's readiness for agreement, such knowledge may increase the agent's aspirations. In these circumstances, rational learning actually shrinks the room for agreement instead of widening it. Thus, when the agents bargain face to face, directly exchanging proposals and replies, the scope for useful credible communication is severely limited or inexistent. For instance, separating equilibria in stationary strategies exhibit the undesirable property that, in the limit as the time interval between the offers vanishes, the probability of agreement vanishes too.[14] Thus, when bargaining is face to face, a smooth learning process conducive to agreement is difficult, if not utterly impossible. Our results demonstrate what can be attained when face to face bargaining is either ruled out, or the agents are cognitively constrained and are cannot update their beliefs. Thanks to the Filter, bargainers learn only what their opponent cannot yield. Over time, learning smoothly decreases players' aspirations.

The idea of drastically filtering communication has previously been explored in Jarque, Ponsatí and Sákovics [2003]. There, instead of the present

[13] See Ausubel, Cramton and Denekere [2002] for an excellent survey and references.

[14] See Theorem 1 in Ausubel and Denekere [1992].

assumption that every division is possible and that concessions must be smooth, it is assumed that concessions must take place in discrete steps. Thus, only a few intermediate agreements can be reached. This characteristic is often a natural feature in realistic situations. However, the discretization of the set of partitions of the surplus comes at the expense of great technical problems. The set of equilibria is very large; they all depend on the distribution of types, and their existence and efficiency performance cannot be established without the detailed information about the distribution of types.[15]

The rest of the paper is organized as follows. In Section 2 we formally describe the mechanism design problem. We define and characterize decent rules. In Section 3 we present the Filtered Demands game. In Section 4 we characterize its equilibria. In Section 5 we show that the equilibria of the FD game implement decent rules and that all decent rules are attained via the FD game. For environments of CRRA utilities we explicitly compute the equilibrium. In Section 6 we illustrate our results with an application to bilateral trade. In Section 7 we provide some welfare comparisons. In Section 8 we conclude and discuss the extensions. Most of the proofs are in the Appendix.

2 Decent Bargaining Rules

Two agents, $i = 1, 2$, bargain over a unit of surplus. Denote by λ_i the share of the good that gets allocated to i. Index j will always indicate the agent other than i, i.e. $j \neq i$, for $i = 1, 2$. To denote a vector (x_i, x_j), we will write x.

The *set of bargaining alternatives*, A, is the set feasible divisions union the disagreement point. Formally, $A = \{\lambda | \lambda_1 + \lambda_2 \leq 1\} \cup d$. An agents type $s_i \in [0, 1]$ represents her reservation share, the share that leaves her indifferent to disagreement, and is her private information. The preferences of agent i over A are represented by a utility function $u_i(\lambda_i - s_i)$, wherewhere $u_i(.)$ is a twice differentiable, strictly increasing, and concave function, with

[15]If we drop the regularity requirement in the equilibria of the FD game, then the set of Bayesian equilibria also contains the equilibria which are similar to those in Jarque, Ponsatí, and Sakovics[2003]; if an agent believes that the opponent will only concede in discrete steps, then it only makes sense to concede at the complementary splits.

$u_i(0) = u_i(d - s_i) = 0$. The payoff from disagreement is normalized to 0. Agents' preferences over $\Delta(A)$, the set of lotteries over A, are represented by their expected utilities. From now on we fix $u = (u_1, u_2)$.

We define a *probabilistic bargaining rule* (Y, P) to be a direct revelation mechanism[16], mapping pairs of reports $z = (z_1, z_2)$ into two-point lotteries $P(z) \otimes Y(z) + (1 - P(z)) \otimes d$. Thus the rule (Y, P) prescribes disagreement with probability $(1 - P(z))$, and agreement at $Y(z) = (Y_1(z), Y_2(z))$ with probability $P(z)$, where $Y_i(z)$ is the share of the surplus allocated to agent i, $0 \leq Y_i(z) \leq 1$ and $Y_1(z) + Y_2(z) = 1$. Given (Y, P) the expected utility of agent i of type s_i upon reports z, $U_i(s_i, z)$, is

$$U_i(s_i, z) = u_i(Y_i(z) - s_i) P(z).$$

We consider the following properties of bargaining rules.

1. Strategy Proofness (SP): (Y, P) is *strategy-proof* if truthful reports constitute a dominant-strategy equilibrium. That is: $U_i(s_i, s) \geq U_i(s_i, z_i, s_j)$ for every $z_i \neq s_i$, all $s_i, s_j \in [0, 1]$, $i = 1, 2$.

2. Ex post Individual Rationality (IR): (Y, P) is *ex-post individually rational* if $P(s) > 0$ implies that $u_i(Y_i(s) - s_i) \geq 0$ for all $s \in [0, 1] \times [0, 1]$, $i = 1, 2$.

3. Weak Efficiency(WE): (Y, P) satisfies *weak efficiency* if

$$s_1 + s_2 < 1 \Rightarrow P(s_1, s_2) > 0.$$

Definition 1 *A quasi-decent bargaining rule is one that satisfies 1 to 3. A decent bargaining rule is a Pareto optimal rule among the quasi-decent bargaining rules.*

[16]This is a slight abuse of terminology. In a general framework, we should define a bargainig rule to be a social choice function, mapping pairs of utility functions into $\Delta(A)$. We will be dealing with the rules that are not manipulable in dominant strategies. Hence, we can appeal to the revelation principle for the dominant strategy environments and identify the set of non-manipulable social choice functions with the set of direct revelation mechanisms that implement them. Also note in our setting, the private information is restricted to be over a one-dimensional parameter, which is thus the only thing that an agent reports to the mechanism.

The first example shows that WE is not a trivial requirement.

Example 2 *Consider a rule which prescribes a split at fixed shares, a posted-split rule, and the agents divide the surplus if they both agree to such division. Thus,* $Y_1(z) = y_1$, $Y_2(z) = y_2 = 1 - y_1$, *where* $y_i \in [0,1]$. *Clearly, such a rule satisfies SP and IR, but does not satisfy the WE.*

WE imposes an open ended constraint. Perhaps one could design a WE mechanism as a combination of an *ex-ante* efficient mechanism that switches with a slight probability to a WE mechanism. In such case, decency would not be well defined. The next example hints that such switching mechanisms fail in either WE or IR, and sometimes also in SP. A formal proof of this is a simple consequence of Lemma 5. The fact that decency is a well defined concept is a trivial consequence of Theorem 7.

Example 3 *Take a rule prescribing a posted-split with probability* $\tilde{p} < 0$, *and giving the whole surplus to agent i with probability* p_i, *where* $p_1 + p_2 + \tilde{p} = 1$. *Such a rule satisfies SP, but doesn't satisfy the IR. To see that WE also fails observe that since the draw of how the surplus gets allocated is made ex-ante, it is not true that the probability of a Pareto-efficient outcome is positive whenever types are compatible. Consider now the rule where first a posted-split is announced, and if the agents don't both agree to it, then with probability* p_i *the whole surplus goes to agent* i, $p_1 + p_2 = 1$, $p_i \in [0,1]$. *This rule always satisfies WE, never satisfies IR, and sometimes satisfies SP, depending on utilities of the agents, on* p_i, *and on the posted split.*

Finally, we give a simple example of a decent mechanism, which provides most of the intuition for what we do in the rest of this section. Proposition 11 generalizes the next example.

Example 4 *Assume that the utilities of the agents are linear:* $u_i(\lambda_i - s_i) = \lambda_i - s_i$. *Then for each* $\pi \in [0,1]$ *a mechanism defined by*

$$P(s) = \begin{cases} \pi \cdot (1 - s_1 - s_2)\,;\ \textit{if } s_1 + s_2 < 1 \\ 0;\ \textit{otherwise} \end{cases},$$

$Y_i(s) = \frac{1}{2}(1 + s_i - s_j)$, $i = 1, 2$, *is quasi-decent. To see that compute*

$$U_i(z_i, s_j; s_i) = \pi(1 - z_i - s_j)\frac{1}{2}(1 + z_i - s_j - s_i)$$

It is immediate to see that this quadratic function has a maximum precisely at z_i s_i*. Theorem 7 below will show that these are all the quasi-decent mechanisms for this case. Hence the unique decent one is obtained by setting* π .

Lemma 5 *(i) If a bargaining rule* Y, P *satisfies (1) and (2) then, for* i , , Y_i s_1, s_2 *is monotonically increasing in* s_i*,* P s_1, s_2 *is monotonically decre*((i))3 1^{s_i} . (If)^{Y_i} 3 bargaining^{P} Rule(i)dictioa(l)-. (ly)3 ^{s_1,s_2} 1 .1(satisfies)-372 (ffi))3 7

u_i Y_i s_1, s_2 , s_i $>$

Y P

Proof.

ρ $- s_1 - s_2$

Y, P , α s

Y_1 s s_1 α s ρ Y_2 s

s_1 $- \alpha$ s ρ P_i α_i

P α α^1 s α s α^2 s $- \alpha$ s

α^i_j

Lemma 6 Y, P

s_1 $s_2 >$ P s .

s_1 $s_2 <$ Y_i s s_i α^i s ρ α^1 s α^2 s

α^i P

P_i s $-u_i'$ α^i s ρ $\left(\alpha^j \; s \; - \alpha^j_i \; s \; \rho\right)$ i , , j / i,

P s u_i α^i s ρ

$\leq \alpha^i$ s $\leq$, $< P$ s $\leq$ P , .

Proof. See Appendix. ■

In the main theorem of this section we summarize the full strength of quasi-decency.

Theorem 7 *Whenever a decent rule exists, it is unique. Moreover if* (Y, P) *is the decent rule for a given vector of utilities, then all the quasi-decent rules are given by* $(Y, \pi P)$*,* $\pi \in (0, 1]$*.*

Proof. See Appendix. ■

Remark 1 *Observe that quasi-decent rules are invariant with respect to multiplying utilities by constants. That is, if* (P, Y) *is a quasi-decent rule for utilities* (u_1, u_2) *then it is also a quasi decent rule for* $(C_1 u_1, C_2 u_1)$ *where* $C_i > 0, i = 1, 2$*.*

We already saw in Example 4 that the decent rule exists when utilities are linear. Does it exist for any other class of utilities? We focus our search for the decent rules among those that satisfy the following simplifying definitions.

Definition 8 *A net surplus rule is a rule* (Y^{ns}, P^{ns}) *where the probability of agreement and the share of the net surplus assigned to each agent depend only on the net surplus.*

Thus, with some abuse of notation, a net surplus rule can be expressed as

$$\begin{aligned} Y_i^{ns}(s) &= s_i + \alpha^i(\rho)\rho, i = 1, 2, \\ P^{ns}(s) &= \begin{cases} P(\rho), \rho > 0, \\ 0, \text{ otherwise.} \end{cases} \end{aligned}$$

Definition 9 *A fixed net surplus rule is a rule for which* $\alpha(s_1, s_2) = \alpha^*$ *for some constant* α^**,* $0 \leq \alpha^* \leq 1$*.*

We now show that among the quasi-decent rules, the net surplus rules are necessarily constant net surplus rules.

Lemma 10 *A quasi-decent rule that is net surplus must be a constant net surplus rule.*

Proof. See Appendix A. ■

In the next proposition we fully characterize the class of utilities for which decent fixed share mechanisms exist:

Proposition 11 *The decent rule exists and it is a constant net surplus rule if and only if either of the following holds:*

1. $u_1(x) = Cu_2(x)$, $C > 0$, *for every* $x \in [0,1]$. *In this case the unique decent rule is the constant net surplus mechanism, given by*

$$\begin{aligned} \alpha_1 &= \alpha_2 = \alpha^* = \frac{1}{2} \\ P^*(\rho) &= \frac{u_1\left(\frac{1}{2}\rho\right)}{u_1\left(\frac{1}{2}\right)} \end{aligned}$$

2. *For* $i = 1, 2$, *agent* i *has a utility of the form*[17]*:*

$$\begin{aligned} u_i(\lambda_i - s_i) &= C_i(\lambda_i - s_i)^{\gamma_i} e^{D_i(\lambda_i - s_i)},\ C_i > 0, \\ \gamma_i &\in (0,1],\ D_i \in [-\gamma_i - \sqrt{\gamma_i}, \sqrt{\gamma_i} - \gamma_i] \end{aligned}$$

   In this case a unique decent rule is the constant net surplus rule, given by

$$\begin{aligned} \alpha^* &= \frac{\gamma_1}{\gamma_1 + \sqrt{\gamma_2\gamma_1}} \\ P^*(\rho) &= \begin{cases} e^{-D_1\alpha_1 s_1 - D_2\alpha_2 s_2}\rho^{\sqrt{\gamma_1\gamma_2}};\ s_1 + s_2 \leq 1 \\ 0;\ \textit{otherwise.} \end{cases} \end{aligned}$$

Proof. See Appendix A. ■

[17] It is easy to see that u_i is increasing and concave on $[0,1]$ if and only if $\gamma_i \in (0,1]$ and $D_i \in \left[-\gamma_i - \sqrt{\gamma_i}, \sqrt{\gamma_i} - \gamma_i\right]$. Notice also that when $D_i = 0$ this utility function is $u_i(\lambda_i - s_i) = (\lambda_i - s_i)^{\gamma_i}$ which is the constant relative risk aversion (CRRA) utility. Also note that in this case, the mechanism is a net surplus mechanism.

3 The Filtered Demands Game

In this section we propose a dynamic bargaining game implementing decent bargaining rules.

THE GAME. The *Filtered Demands Game* (FD game) is a continuous-time game. The agents send private messages claiming some share of the good to the Filter. The Filter is a dummy player whose only role is to receive claims, keeping them secret while they are incompatible, and to announce the agreement as soon as it is reached. As time goes by, the agents can continuously decrease their demands at any moment. Thus the agents revise their claims until they become mutually compatible. Then the Filter announces that agreement has been reached, the agents receive the agreed shares, and the game ends.

STRATEGIES. A *strategy* $\lambda_i(.,.)$ *of player* i is a function mapping her type s_i and time t into a share,

$$\lambda_i(.,.) : [0,1] \times [0,\infty) \to [0,1] \ , \ i = 1,2$$

Thus $\lambda_i(s_i,t)$ is the share agent i of type s_i claims for herself at $t \geq 0$. Strictly speaking, a strategy is a function mapping each type and each history into a proposal at every moment. However, given her type s_i, the history at time t only depends on t, as the agent is not able to see the proposals of her opponent. The rules of the game are such that the following conditions have to be satisfied by the players' strategies:

1. $\lambda_i(s_i,.)$ is a right continuous function of time.

2. $\lambda_i(s_i,t)$ is non-increasing in t for all $t \in (0,\infty)$ and all $s_i \in [0,1]$.

The first condition assures that outcomes are well defined. The second condition is innocuous: an agent can only observe agreement (or disagreement) so it is plausible that she will either keep her initial demand or monotonically concede to her opponent[18].

OUTCOMES. If two strategies are such that $\lambda_1(s_1,0) + \lambda_2(s_2,0) < 1$, i.e. the demands are more than compatible at $t = 0$, then agreement between

[18] Suppose an agent decided to play a strategy that would at time t_1 require a share λ_1 for herself. Then it would make little sense to demand $\lambda_0 < \lambda_1$ at an earlier time $t_0 < t_1$. The agent might just as well stick to λ_1 at t_0 . Condition 2 allows the agents to forget about such considerations.

types (s_1, s_2) occurs at $t = 0$ at shares $\lambda_i(s_i, 0) + \frac{1-\lambda_1(s_1,0)-\lambda_2(s_2,0)}{2}$.[19] Given a pair of types s a strategy profile λ determines a unique outcome of the game denoted by $(x_1(\lambda, s), x_2(\lambda, s), \tau(\lambda, s))$, where x_1 and x_2 are the shares of agents 1 and 2, and τ is the time of agreement, i.e.

$$\lambda_i(s_i, \tau(s)) + \lambda_j(s_j, \tau(s)) \leq 1.$$

Inter-temporal Utilities. The static utility of agent i is given by the function $u_i(\lambda_i - s_i)$, satisfying the same requirements as in Section 2. Agents discount the future exponentially. Thus upon agreement at $t \geq 0$, the payoff of agent i is given by

$$U_i(\lambda_i, s_i, t) = e^{-t} u_i(\lambda_i - s_i).$$

Clearly, in the event of perpetual disagreement the payoffs are zero. This is for convenience and can be relaxed to a general class of discounting criteria $\delta_i(t)$, where $\delta_i(.)$ is a strictly pi

δ_i $_{t\to\infty}\delta_i$ t

Information and Beliefs. It is knowledge that pairs of reserva]Tonalues $s$are drawn from a continuous joint distribu]Ton function F, with f , $\times$, [20] i
type, and has beliefs over the distribu]Ton of the opponent's types. As noted above, the histories depend only on t. Thus, given a strategy profile λ, the beliefs of a player about the opponent are updated only as a function of time. We denote by $f_j(s_j|s_i, t)$ the density of the belief of agent i of type s_i
j t t
economize the nota]Ton write only (f_j) $s_j$whenever convenient. Note that if the strategies of both players are differentiable with respect to both parameters, the beliefs will be differentiable with respect to time.

equilibrium. We are interested in the Perfect Bayesian Equilibria (PBE) of the FD game. In our setup, the set of PBE is equivalent to the set of Bayesian Equilibria (BE) since off-equilibrium histories either end the game

[19] Our results are independent of the excess sharing rule, as long as it gives positive shares to both agents.

[20] Our results hold for any F with positive density f on a square $[\underline{s}, \overline{s}] \times [\underline{s}, \overline{s}]$, $\underline{s} < 1/2 < \overline{s}$. This is equivalent to the requirement that F has support on $[\underline{s}, \overline{s}] \times [\underline{s}, \overline{s}]$, and the conditional beliefs of agents are independent. In the literature, this condition also appears as the "spanning condition" (see for instance Mookherjee and Reichelstein[1992], p395).

or are unobservable to the opponent; own deviation from a BE cannot be optimal at any t. Hence, a formal definition of BE will suffice.

Let $EU_i^{s_i}(\lambda, F)$ denote the expected payoff of player i of type s_i at the strategy profile λ when types are distributed according to F:

$$EU_i^{s_i}(\lambda, F) = \int_0^1 u_i(x_i(\lambda, s_i, u) - s_i)\, e^{-\tau(\lambda, s_i, z)} dF(s_i, u).$$

Denote by Λ_i the set of strategies for player i. A strategy profile λ constitutes a *Bayesian equilibrium* if and only if

$$EU_i^{s_i}(\lambda, F) \geq EU_i^{s_i}\left(\lambda_i', \lambda_j, F\right), \forall \lambda_i' \in \Lambda_i,$$

for all $s_i \in [0,1]$, $i = 1, 2$, $j \neq i$.

Observe that for each BE profile λ, a profile λ' constructed by adding a *stand still interval* $[0, T)$, i.e. $\lambda_i'(s_i, t+T) = \lambda_i(s_i, t)$, is a BE as well, for any $T < \infty$. As the opponent does not concede any positive amount until T, no concession prior to T is useful[21]. Regardless of T, such strategy profiles λ' are weakly dominated. In the next section we will show that the type of player i who makes the earliest relevant offer is $s_i = 0$. Type $s_i = 0$ has nothing to lose if she starts moving at 0, since she has no reason to expect some other type to start moving any earlier. This, in turn, would provoke other types to start moving as well. We say that a BE is *undominated* if it does not have a stand still interval.

Since types and dates take values in a continuum, and the range of strategies is also a continuum, natural patterns of behavior should rule out dramatic changes when types change only marginally. We say that a strategy is *regular* provided that:

1. $\frac{\partial \lambda_i(s_i,t)}{\partial t}$ exists and is continuous for all $t \in [0, \infty)$, $s_i \in [0,1]$;
2. $\frac{\partial \lambda_i(s_i,t)}{\partial s_i}$ exists and is continuous for all $t \in [0, \infty)$, $s_i \in [0,1]$;
3. $\lim_{t\to\infty} \lambda_i(s_i, t)$ is a left-continuous function of s_i.

[21] When the support of types is $[s_{L,} s_H]$, $s_L < 0$, the delay that a negative type is ready to endure, rather than agree to 0 and obtain at least $-s_L > 0$, is bounded above. In these case stand- still PBE can be sustained only for $T \in [0, T^L)$, $T^L < \infty$.

The first condition means that players do not change her demand by a positive amount in 0 time (recall that she is only allowed to increase what she is willing to offer to her opponent). This condition eliminates strategies are as those described by Jarque et. al [2002], where demands are step functions taking only finitely many values. When condition 1 does not hold, the agents might in a sequentially rational way believe that the opponent will almost surely only bid finitely many intermediate agreement points between the two extreme agreements. Best response to such a strategy is to bid only the complementary intermediate agreement points (since any other bid is essentially irrelevant). The second condition requires smoothness with respect to types. In equilibrium, it will imply that player i's strategy is fully separating. The last condition is roughly an indifference breaking rule: if an agent of some type is at the horizon indifferent between two concessions to the opponent, she will concede more (see also the footnote in the proof of Lemma 14 in the Apendix). This condition is enough to assure that the continuity of the demands with respect to types is preserved at the time horizon.

Definition 12 *From now on, an equilibrium of the FD game is a BE in undominated and regular strategies.*

4 Equilibria in the FD Game

We will show that there is a one to one correspondence between the set of equilibria of the FD game and the set of decent rules. First we introduce some preliminary results that will be useful in characterizing equilibria of the FD game. The first obvious observation is that agents prefer disagreement to negative payoffs at every moment.

Lemma 13 Ex-post Individual Rationality: *In equilibrium $\lambda_i(s_i, t) \geq s_i$ for all t and all s_i.*

In the next lemma we state that claimed shares asymptotically approach the reservation values. The intuition is clear: if the agent of a given type doesn't reach agreement in a very long time, the opponent was probably of a relatively high type. Thus the agent should lower her demand, and she would only keep lowering it until her type[22].

Lemma 14 ASYMPTOTIC DEMANDS: *In equilibrium* $\lim_{t\to\infty} \lambda_i(s_i, t) = s_i$ *for all* $s_i \in [0, 1]$.

Proof. See Appendix. ■

We next assert that agents with high reservation values, "tougher" agents, never demand less than "softer" ones. Moreover, individuals with different types never make at the same time the same relevant concession (one that instantaneously leads to an agreement with some of the opponent's types).

Lemma 15 TYPE MONOTONICITY: *In equilibrium,* $\frac{\partial \lambda_i(s_i,t)}{\partial s_i} \geq 0$, $t \in (0, \infty)$, $s_i \in [0, 1]$. *Moreover, for each type* s_i *and* $t > 0$, *if* $\lambda_j(s_j, t) = 1 - \lambda_i(s_i, t)$ *for some* $s_j \in [0, 1]$ *then* $\frac{\partial \lambda_i(s_i,t)}{\partial s_i} > 0$.

Proof. See Appendix. ■

We now discuss the optimization problem of the agents when their opponent uses a strategy that is regular and strictly increasing in types. After two lemmas deriving the initial conditions for the optimal strategies of the agents, we state the dynamic optimization program that the agents are facing. The main proposition of this section follows. There we derive the first order condition, which turns out to be belief independent.

We first focus on the initial conditions for the agents' strategies. From Lemma 15, it follows that any s_i starts participating in the negotiations once her demand becomes feasible with the demand of $s_j = 0$. Before that moment the agent must know that she is demanding too much to agree even with the lowest type of the opponent. Therefore the question is: should an agent enter in the game already at $t = 0$ (and with what demand), or should she wait until the field softens up a bit. The answer is provided by the following lemma. We denote by $g_i(s_i)$ the starting point of the demand of type s_i: $g_i(s_i) = \lim_{t\searrow 0} \lambda_i(s_i, t)$.

Lemma 16 INITIAL CONDITION: *In any equilibrium it must hold that* $g_1(0) + g_2(0) = 1$.

[22]This argument is only valid if the reservation demand of the "toughest type" is very high - that is if $s_H \geq 1$.

Proof. In an equilibrium the type $s_i = 0$ at time 0 demands a share that will give her a positive probability of agreement in at least a very short time - otherwise each type of every agent would know that there was some dead delay at the start where the only thing that would happen would be that agents would lower their demands up to the point where the lowest types could agree. On the other hand, it could not be that she would demand a share which would meet the demand of some type $s_j^0 > 0$ of player j - meaning that $g_i(0) + g_j\left(s_j^0\right) = 1$. This follows from the excess profit sharing rule since then an agent $s_i = 0$ could profitably deviate by starting with a demand that met type $s_j = 0$. Then she would "rip off" all the excess agreement profits by lowering her demand very rapidly to $1 - g_j\left(s_j^0\right)$. By making her move fast enough it is clear that such deviation could be profitable. ■

Thus for all types except the lowest type it is in equilibrium optimal to wait with a high demand for a while. It means that there will necessarily be delays with probability 1.

Now we define the *entry time of* i of type s_i, $t_i^E(s_i)$, as the first moment that agent i makes a realistic proposal. That is

$$t_i^E(s_i) = \inf\{t > 0 | \, \lambda_i(s_i, t) + \lambda_j(s_j, t) \leq 1 \text{ for some } s_j \in [0,1]\}.$$

The following corollary to the above lemma establishes that $t_i^E(s_i)$ is the moment when the demand of type s_i is compatible exactly with the lowest type of the opponent. The proof is exactly the same as the proof of Lemma 16. The remark that follows is equally simple.

Corollary 17 Timing of Entry: *In equilibrium,* $\lambda_i\left(s_i, t_i^E(s_i)\right) = 1 - \lambda_j\left(0, t_i^E(s_i)\right)$, *for* $i = 1, 2$, $j \neq i$, *and all* $s_i \in [0,1]$.

Remark 2 *Notice that in any equilibrium* $t_i^E(s_i) < \infty$ *if and only if* $s_i < 1$. *Otherwise the strategy of* s_i *would be strictly dominated.*

We are now ready to turn attention to the dynamic optimization. In equilibrium, agents select a strategy aiming at the highest possible payoff, given the type-contingent strategies of the other player. Thus agents are picking optimal functions $\lambda_i(s_i, \cdot)$, $i = 1, 2$. This means that agent i of type s_i decides how her concessions of the good to the other side should optimally change with time.

From now on, let $\sigma_j(s_i, t; \lambda)$ be the function giving the type of agent j with whom agent i of type s_i enters in agreement at moment t if profile λ is played. We will omit λ in the arguments of $\sigma_j(.)$. Formally, for any $(s_i, t) \in [0,1] \times [t_i^E(s_i), \infty)$, $\sigma_j(s_i, t)$ is the solution of the equation

$$1 = \lambda_j(\sigma_j(s_i, t), t) + \lambda_i(s_i, t). \tag{2}$$

A consequence of Lemma 15 and the implicit function theorem is that such σ_j is well defined. This can also be seen from the proof of the following proposition.

Proposition 18 OPTIMIZATION PROGRAM: *In equilibrium, agent i of type s_i solves the following optimization program*

$$\underset{\lambda_i(s_i,\cdot)\in\Lambda_i}{Max} \int\limits_{[t_i^E(s_i),\infty)} e^{-t} u_i(\lambda_i(s_i,t) - s_i) f_j(\sigma_j(s_i,t)) \frac{\partial \sigma_j(s_i,t)}{\partial t} dt,$$

$$s.t.\ (2)\ and\ \sigma_j\left(s_i, t_i^E(s_i)\right) = 0\ defines\ t_i^E(s_i).$$

Proof. Fix the type of agent i to be s_i. When entering into negotiations at $t_i^E(s_i)$, she decides her optimal concession plan $\lambda_i(s_i,t)$, $t > t_i^E(s_i)$, in order to maximize her expected discounted future payoff. Denote by $P(t)$ the probability of type s_i reaching agreement up to time t (for simplicity we omit the parameter s_i in $P(t)$). Agent i is solving the following program

$$\underset{\lambda_i(s_i,\cdot)\in\Lambda_i}{Max} \int\limits_{[t_i^E(s_i),\infty)} e^{-t} u_i(\lambda_i(s_i,t) - s_i) dP(t)$$

But the possibility of reaching an agreement at some $t > t_i^E(s_i)$ is exactly the possibility that agent i will at t meet the demand of some type of agent j. For any $t \geq t_i^E(s_i)$, recall that $\sigma_j(s_i,t)$ is the type of agent j with whom i reaches agreement at moment t. Thus $\sigma_j(s_i,t)$ is implicitly defined from the relation

$$\lambda_j(\sigma_j(s_i,t), t) + \lambda_i(s_i,t) = 1.$$

By definition and Lemma 15, [23] $\sigma_j\left(s_i, t_i^E(s_i)\right) = 0$, and by Lemma 14 $\lim_{t\to\infty} \sigma_j(s_i,t) = 1 - s_i$. Taking the derivative with respect to t, we can

[23] The corollary implies that at every instant there will be only one type reaching an agreement with any particular type of the other agent.

express

$$\frac{\partial \sigma_j\left(s_i,t\right)}{\partial t} = -\frac{\frac{\partial \lambda_j(\sigma_j(s_i,t),t)}{\partial t}+\frac{\partial \lambda_i(s_i,t)}{\partial t}}{\frac{\partial \lambda_j(\sigma_j(s_i,t),t)}{\partial s_j}}.$$

By assumption, $\frac{\partial \lambda_i}{\partial t}$ and $\frac{\partial \lambda_j}{\partial t}$ are both finite and non-positive. Hence we see from Lemma 15, and the implicit function theorem, that for any $t \geq t_i^E\left(s_i\right)$, $\sigma_j\left(s_i,t\right)$ is a well defined differentiable function of time, with $0 \leq \frac{\partial \sigma_j(s_i,t)}{\partial t} < \infty$. In other words, at any $t \geq t_i^E\left(s_i\right)$ there exists exactly one type $\sigma_j\left(s_i,t\right)$ of player j, with whom s_i would reach agreement at that moment. These facts have two consequences. First, the probability of reaching an agreement by t, $P\left(t\right)$, has no mass points because the distribution of types of player j has no mass points. Second, the marginal increase in $P\left(t\right)$, i.e. $dP\left(t\right)$, is equal to the marginal increase of the mass of types of player j, that player i would agree with by moment t. Also, agent i knows that before $t_i^E\left(s_i\right)$ her proposals were unrealistic, so she cannot update her beliefs until that moment. Since σ_j is differentiable with respect to time, the beliefs are updated continuously and differentiably from $t_i^E\left(s_i\right)$ on. In other words, we have established that at $t_i^E\left(s_i\right)$ the belief of agent i is exactly $F_j(s_j)$, and at every moment $dP\left(t\right) = dF_j\left(\sigma_j\left(s_i,t\right)\right) = f_j\left(\sigma_j\left(s_i,t\right)\right)\frac{\partial \sigma_j(s_i,t)}{\partial t}dt$. This completes the proof. ■

The optimization problem stated in Proposition 18 can be best approached as a problem where i is choosing two unknown functions $\lambda_i(s_i,\cdot)$ and $\sigma_j\left(s_i,.\right)$ which are bound by the constraint (2), where $\lambda_j(\cdot,\cdot)$ is a given and fixed function (the strategies of all possible types of agent j).[24] The optimality condition at the lower boundary of optimization is given by definition of $t_i^E\left(s_i\right)$ - implicitly written it is $\sigma_j(s_i,t_i^E\left(s_i\right)) = 0$. In the following lemma we provide the first order condition of the optimization program of agent i, for $t > t_i^E\left(s_i\right)$. We are omitting most of the arguments in the functions. The arguments are: $\lambda_i = \lambda_i\left(s_i,t\right)$, except when differentiating with respect to λ_i, and $\sigma_j = \sigma_j\left(s_i,t\right)$.

Lemma 19 First order condition: *Fix an equilibrium and a type s_i . For $t > t_i^E\left(s_i\right)$ the function $\lambda_i(s_i,.)$, $i=1,2$, satisfies the following first order condition*

$$u_i\left(\lambda_i,s_i\right) = u_i'\left(\lambda_i - s_i\right)\left(\frac{\partial \lambda_j\left(\sigma_j,t\right)}{\partial s_j}\frac{\partial \sigma_j}{\partial t}+\frac{\partial \lambda_i}{\partial t}\right). \tag{3}$$

[24] Our manual for the calculus of variations is Elsgolts [1970].

Proof. See Appendix.

Lemma 19 yields a condition that is independent of the beliefs of player i about the types of player j. This remarkable property is of importance

j

λ_j

j *for any given type* s_j i s_i

λ_j s_j, t

i j

i *must be playing a best response to every type of the other player*

[25]

Proposition 20 General Discounting functions *Suppose that discounting is given by a general function* δ t *where* δ $.$ *is a strictly positive, monotonically decreasing function with* δ *, and* $_{t\to\infty}\delta$ t $.$ *Denote the equilibrium strategies for exponential discounting by* λ_i s_i, τ *. Then the equilibrium strategies for discounting* δ t *are given by* λ_i s_i, t

$\left\{ , \lambda_i \; s_i, - \quad \delta \; t \quad \right\}$.

Proof.

5 Existence of Equilibria and Implementation

Implemented Bargaining rule. τ s

λ_i s_i, τ s λ_j s_j, τ s $e^{-\tau(s)}$

λ *implements* Y, P *if and only if the outcome associated to* λ *is such that* Y_i s λ_i s_i, τ s *, and* P s $e^{-\tau(s)}$.

[25]This independence with respect to discounting could prove useful when designing experiments.

Since time plays a crucial role in the present setup, a natural interpretation is that rather than selecting outcomes stochastically, bargaining rules allocate agreements over time. We may thus view the FD game as a *dynamic bargaining rule (*and its associated *dynamic direct revelation mechanism*), (X, τ), where individuals report their type at $t = 0$ and are instructed to implement an agreement with shares $X_i(s)$ only at date $\tau(s) \in [0, \infty]$. If $\tau(s) = \infty$, then the prescribed outcome is disagreement.

The FD game in equilibrium implements decent bargaining rules. We will show that this is always true, independently of agents' utilities. Clearly, the question is whether any equilibria of FD game exist at all (the agents' strategy sets are non-compact). We will prove that there is a one-to-one correspondence between the set of equilibria of the FD game and the set of decent rules. Hence, for any utilities, an equilibrium of the FD game exist if and only if a decent mechanism exists. By the uniqueness theorem for decent mechanisms, we know that the equilibrium of the FD game will always be unique. All of this is summarized in the following lemmas and propositions.

Lemma 21 INDIVIDUAL RATIONALITY: *Equilibria of the FD game implement rules satisfying IR.*

Proof. This is a direct consequence of Lemma13. ■

Lemma 22 WEAK EFFICIENCY*: Equilibria of the FD game implement rules satisfying weak efficiency.*

Proof. Lemmas 14 and 15 imply that all pairs that produce a positive net surplus reach agreement at a finite date, which translates into WE. ■

Lemma 23 STRATEGY PROOFNESS: *Equilibria of the FD game implement rules satisfying SP.*

Proof. By Lemma 19 equilibria of the FD game are belief independent. This implies that the implemented rule must be strategy proof (see Ledyard[1978]). ■

Lemma 16 then implies that the FD game implements precisely decent rules. As we show in the next proposition, there is a one-to-one correspondence between the set of the equilibria of the FD game and the set of decent mechanisms.

Proposition 24 IMPLEMENTATION: *An equilibrium of the FD game implements a decent bargaining rule. Conversely, any decent bargaining rule is implementable as an equilibrium of the FD game.*

Proof. The previous three lemmas show that the rule implemented in the equilibrium of the FD game must be quasi-decent. Decency is then implied by the fact that equilibria of the FD game are by definition undominated, hence the probability of the Pareto-efficient outcome is always set to be maximal.

For the converse, see the Appendix. ■

Corollary 25 FD-UNIQUENESS: *Whenever the equilibrium in the FD game exists, it is unique.*

Proof. This is a direct consequence of Theorem 7 and the previous proposition. ■

These are the central results of this section. Proposition 6 shows how to *calculate* all decent rules, given one-parametric utility functions. Proposition 24 goes much further: regardless of the utilities, an equilibrium of the FD game implements a decent bargaining rule, as long as the two agents know each others utilities. *The designer does not need to know this information.*

Unless the set of equilibria of the FD game is empty, this game must implement a decent bargaining rule. Computing an equilibrium requires solving self-referential equations (2) and (3). For the environments where the decent rule is known, it is by the Proposition 24 easy to compute the equilibria of the FD game. In the next proposition we provide the strategies of the agents if they both have CRRA utilities. A similar exercise can be repeated for the other cases where the decent exists and it is a constant net surplus rule.

Proposition 26 EXISTENCE: *If agents have CRRA utilities, then the following type-contingent strategies are the unique equilibrium and they implement the α^*-constant net surplus rule*[26]*:*

$$\lambda_i\left(s_i, t\right) = \min\left\{1, s_i + \alpha_i^* e^{-\frac{t}{\sqrt{\gamma_1\gamma_2}}}\right\}, i = 1, 2,$$

[26] Notice that at the time when $1 = s_i + u_i^{-1}\left(u_i\left(\alpha^*\right)e^{-t}\right)$ this violates our assumption on differentiability of strategies, but the strategies are still differentiable a.e. Moreover, at that point, the demand of agent of type s_i is irrelevant, hence we can modify it slightly to make it smooth.

where

$$\alpha_1^* = \frac{\gamma_1}{\gamma_1 + \sqrt{\gamma_2\gamma_1}} \tag{4}$$

and $\alpha_2^* = 1 - \alpha_1^*$.

Proof. For a proof that these strategies satisfy the FOC of the FD game see Appendix A. The rest follows from Propositions 11 and 24, and the previous corollary. ■

6 Bilateral trade and ex-ante efficiency

In this section we evaluate the ex-ante performance of the decent rule in different scenarios, and we compare it to alternative mechanisms.

To carry out this exercise we focus attention to problems of bilateral trade: A seller can produce the good at a cost s, the buyer values the good at b. Upon agreement at date t, on a price p, the seller obtains $u_s(p-s)e^{-t}$ and the buyer obtains $u_b(b-p)e^{-t}$.[27]

In this context, the FD game is a closed-book dynamic double auction, agents continuosly submit ask and bid prices, $A(s,t)$ and $B(b,t)$, that are not displayed unless trade proceeds. The Filter can be interpreted as a computer with two input nodes (one for each agent) and two output screens where either "No trade yet" or "Trade" appear in the corresponding states of the world.

Risk neutrality: When both agents are risk neutral individually rational rules maximizing the ex-ante sum of payoffs subject to Bayesian incentive compatibility and strategy proofness are known thanks to Myerson and Satterthwaite [1983] and Hagerty and Rogerson [1987] respectively. We can thus compare the expected surplus attained under these two rules to the expected payoffs of the decent rule.

Example 27 Potential Welfare under uniform distributions. *Assume that both agents are risk neutral, and that costs and valuations are independent and uniformly distributed in* $[0,1]$.

[27] To relate to our previous notation simply identify player 1 with the seller and player 2 with the buyer and set $s = s_1$, $b = 1 - s_2$.

1. The unconstrained potential welfare is

$$W^U = \int_0^1 \int_s^1 (b-s)\, dbds = \frac{1}{6}.$$

2. Myerson and Satterthwaite [1983] prove that the optimal rule under Bayesian incentive compatibility and individual rationality is to trade for sure if $b \geq s + \frac{1}{4}$ and abstain from trade otherwise. This yields expected total surplus

$$W^{MS} = \int_0^1 \int_{s+\frac{1}{4}}^1 (b-s)\, dbds = \frac{9}{64}.$$

   Hence, approximately 85% of the potential expected gains W^U are attained.

3. Hagerty and Rogerson [1987] show that under strategy proofness, the ex-ante surplus that obtains with any given rule can be attained with a posted-price rule. It is immediate to check that expected surplus is maximized at the fixed price $p = \frac{1}{2}$. This yields

$$W^{HR} = \int_{\frac{1}{2}}^1 \int_0^{\frac{1}{2}} (b-s)\, dsdb = 0.125,$$

   which is 75% of W^U.

4. Finally, the decent rule, trade at price $\frac{b+c}{2}$ with probability $(b-c)$ if and only if $b >$, yields expected surplus

$$W^d = \int_0^1 \int_0^b \frac{b+c}{2} (b-c)\, dcdb = \frac{1}{12},$$

   which is only $\frac{1}{2}$ of W^U.

The next example compares decent mechanisms with posted prices when the designer is poorly informed on the distribution of costs and values.

Example 28 Misspecification of the Distribution: *Let agents be risk neutral. Assume that the designer had no knowledge about the distribution of costs and valuations and that by the principle of insufficient reason he assumed that it was symmetric for both agents. He would then use the* $p = \frac{1}{2}$ *posted-price rule. However, assume that the distributions of agents' reservation shares were in fact asymmetric. In particular, take an* $\varepsilon > 0$*, and assume that the reservation shares of the agents were independent and their densities had the following forms:*

$$f_s(s) = \begin{cases} \varepsilon \textit{ for } 0 \le s \le \frac{1}{2} \\ \left(\frac{1}{\varepsilon} - 1 + \varepsilon\right) \textit{ for } \frac{1}{2} < s \le \frac{1}{2} + \varepsilon \\ \varepsilon \textit{ for } \frac{1}{2} + \varepsilon \le s \le 1 \end{cases}$$

$$f_b(b) = \begin{cases} \varepsilon \textit{ for } b < 1 - \varepsilon \\ \left(\frac{1}{\varepsilon} - 1 + \varepsilon\right) \textit{ for } 1 - \varepsilon \le b \le 1 \end{cases}$$

Some tedious, straight-forward calculus shows that then the gains from trade under the posted-price rule are equal to $\frac{1}{8}\varepsilon\left(3 - 4\varepsilon + 2\varepsilon^2\right)$ *and the gains from trade under the decent rule are* $\frac{1}{24}\left(6 - 27\varepsilon + 60\varepsilon^2 - 53\varepsilon^3 + 12\varepsilon^4 + 4\varepsilon^5\right)$*. Clearly, for* ε *small enough, the decent rule extracts a sizeable portion of the possible gains from trade, whereas the posted-split rule extracts almost none.*

Risk Aversion. When both agents have CRRA utilities

$$u_s(p - s) = (p - s)^{\gamma_s}, u_b(b - p) = (b - p)^{\gamma_b},$$

the decent rule for bilateral trade implements trade at price $p(s,b) = \alpha^*(s+b)$, $\alpha^* = \frac{\gamma_b}{\gamma_b + \sqrt{\gamma_b\gamma_s}}$, with probability $P(s,b) = (b-s)^{\sqrt{\gamma_s\gamma_b}}$ if and only if $b > s$. Observe is that as the agents become more risk averse, i.e. γ_1 and γ_2 go to 0, the probability of trade when $b > s$ increases, approaching full effciency in the limit. Therefore, under sufficient risk aversion, the decent mechanism dominates the best posted price rule ex-ante.

Example 29 The decent rule dominates posted prices: *Assume that both agents are equaly risk averse,*

$$u_s(p - s) = (p - s)^{\gamma}, u_b(b - p) = (b - p)^{\gamma},$$

and that costs and valuations are independent and uniformly distributed in $[0, 1]$*. In the absence of incentive constraints the total surplus is maximized when agents trade at a price* $\frac{b+s}{2}$ *whenever* $b > s$*. This yields ex-ante payoffs*

$$W^U = \int_0^1 \int_s^1 2 \left(\frac{b-s}{2} \right)^\gamma dbds.$$

The ex-ante payoffs under the optimals posted price rule and under the decent rule are easily computed:

$$W^{pp} = \int_{\frac{1}{2}}^1 \int_0^{\frac{1}{2}} \left(b - \frac{1}{2} \right)^\gamma + \left(\frac{1}{2} - s \right)^\gamma dbds = \frac{2}{(\gamma+1)(\gamma+2)} \left(1 - \frac{1}{2^{\gamma+1}} \right)$$

$$W^d = \int_0^1 \int_s^1 2 \left(\frac{b-s}{2} \right)^\gamma (b-s)^\gamma dbds = \frac{2}{2^\gamma (2\gamma+1)(2\gamma+2)}$$

It is easily checked that if $\gamma < .29$ *then the decent mechanism performs better ex-ante than the optimal posted price mechanism.*

7 Conclusion and Extensions

We have addressed the design of mechanisms for the bargaining problem where the disagreement points are private information. For the environments with concave utilities, we have fully characterized bargaining rules that we call decent - those that are Pareto Optimal in the constrained set of rules satisfying individual rationality, weak efficiency, and strategy proofness. We have proved that when it exists, the decent rule is unique; by construction we have proved the existence for a large set of utilities. We have proposed a simple dynamic game, the FD game, which always implements the decent rule, regardless of the agents' utilities and discounting criterion. This implementation result is due to the fact that the equilibria of the FD game do not depend on agents' beliefs. The game protocol itself is simple. Neither a dictatorial principal designing complex contracts, nor strong commitments to assure the agents' obedience over time are required. The dynamic game thus provides a link between the weak efficiency and the renegotiation-proofness. It also provides a sharp prediction to the situations of bilateral bargaining under incomplete information when the agents' behavior is regular and their updating constrained.

Our characterization of the decent rules can be interpreted in the spirit of the classical axiomatic approach to bargaining. Taking as the starting point a bargaining problem in which the disagreement point is not common knowledge, we characterize the rule that induces agents to reveal their private information, and assigns with a positive probability a Pareto-optimal solution to the revealed problem. Given that we require strategy proofness, this probability has to non-trivially depend on the agents' private information.

The present work can be extended to address situations with more than two agents. In Copic and Ponsati [2002a] we address multilateral bargaining with private reservation shares. This generalization is appropriate to address the problems of when to supply, and how to share the cost of a public good when there are many agents[28]. In this case, the Filter can be envisioned as a central agent administering a public account. Individuals pledge their contributions towards the cost of the public good, and can increase their pledge at any time. The Filter assures that contributions are not publicly disclosed until the necessary amount has been pledged. Payments are made only if and when the project is carried out. In Copic and Ponsati [2002b] we discuss markets with many participants. There we generalize the FD game to a dynamic double auction with many sellers and buyers with private valuations of the object. In that case the FD game can be imagined as a market with continuous trading and a closed limit order book.

References

[1] Ausubel, L. M., P. Cramton, and R. J. Denekere [2002]: "Bargaining with Incomplete Information", in *Handbook of Game Theory* , edited by R. J. Aumann and S. Hart, Amsterdam: Elsevier Science B. V. , forthcoming.

[2] Ausubel, L. M., and R. J. Denekere [1992]: "Durable Goods Monopoly with Incomplete Information ", *Review of Economic Studies*, **59**, 795-812.

[28] Mailath and Postlewaite [1990] address the question of Bayesian incentive-compatible mechanisms for the environments with risk-neutral agents.

[3] Barberà, S. and M.O. Jackson [1995]: "Strategy-Proof Exchange", *Econometrica*, **63**, 51-87.

[4] Copic, J. and C. Ponsati [2002a]: "Dynamic Revelation of Preferences for Public Goods", mimeo.

[5] Copic, J. and C. Ponsati [2002b]: "Dynamic Double Auctions", mimeo.

[6] Corchon, L. [1996]: *The Theory of Implementation of Socially Optimal Decisions in Economics*, London, MacMillan Press.

[7] Corchon, L. and C. Herrero [2003]: *A Decent Proposal,* mimeo.

[8] Elsgolts, L.: *Differential Equations and the Calculus of Variations,* Mir Publishers, Moscow, 1973.

[9] Gibbard, A. [1973]: "Manipulation of voting schemes", *Econometrica*, **41**, 587-601.

[10] Green, J. and J.-J. Laffont [1977]: "Characterization of satisfactory mechanisms for the revelation of preferences for public goods", *Econometrica*, **41**, 617-631.

[11] Hagerty, K. M. and W. P. Rogerson [1987]: "Robust Trading Mechanisms", *Journal of Economic Theory*, **42**, 94-107.

[12] Hurwicz, L. [1972]: "On Informationally Decentralized Systems" *Decision and Organization: A Volume in honor of Jacob Marshak* , edited by R. Radner and C. B. McGuire, Amsterdam, North Holland.

[13] Hurwicz, L. [1994]: "Economic Design, Adjustment Processes, Mechanisms and Institutions" *Economic Design*, 1, 1-14.

[14] Jackson, M.O. and T. R. Palfrey [2001]: "Voluntary Implementation" *Journal of Economic Theory*, 98, 1-25.

[15] Jarque, X., C. Ponsati and J. Sakovics [2003]: "Mediation:Incomplete Information Barganing with Filtered Communication", *Journal of Mathematical Economics*, forthcoming.

[16] JOHN, FRITZ [1971] *Partial Differential Equations*, Springer -Verlag, New York, Heidelberg, Berlin.

[17] LEDYARD, J. O. [1978]: " Incentive Compatibility and Incomplete Information", *Journal of Economic Theory*, **18**, 171-189.

[18] MAILATH, G. J. AND A. POSTLEWAITE [1990]: "Asymmetric Information Bargaining Problems with Many Agents", *Review of Economic Studies*, **57**, 351-367.

[19] MOOKHERJEE, D. AND S. REICHELSTEIN [1992]: "Dominant strategy implementation of Bayesian incentive compatible allocation rules", *Journal of Economic Theory*, **56**, 378-399.

[20] MYERSON, R. B., M. A. SATTERTHWAITE [1983]: "Efficient Mechanism for Bilateral Trading", *Journal of Economic Theory*, **29** , 265-281.

[21] RUBINSTEIN, A. AND A. WOLINSKY [1992]: "Renegotiation-Proof Implementation and Time Preferences", *American Economic Review*, **83**, 600-614.

[22] RUDIN, W. *Real And Complex Analysis,* McGraw-Hill, Singapore, 1987.

[23] SATTERTHWAITE, M. [1975]: "Strategy proofness and Arrow's conditions: Existence and correspondence theorems for voting procedures and social welfare functions", *Journal of Economic Theory*, **10**, 187-217.

[24] SPRUMONT, Y. [1991]: "The division problem with single-peaked preferences: A characterization of the Uniform Allocation Rule", *Econometrica*, **59**, 509-519.

[25] WILLIAMS, S. [1999]: "A characterization of efficient, Bayesian incentive compatible mechanisms", *Economic Theory*, **14**, 155-180.

Appendix

PROOF OF LEMMA 5:

Proof. i) Assume (Y, P) satisfies IR and SP. We first prove monotonicity. By strategy proofness for all z_i, z_i' and z_j

$$P(z_i, z_j)\, u_i(Y_i(z_i, z_j) - z_i) \geq P(z_i', z_j)\, u_i(Y_i(z_i', z_j) - z_i)$$

and

$$P\left(z_i', z_j\right) u_i\left(Y_i\left(z_i', z_j\right) - z_i'\right) \geq P\left(z_i, z_j\right) u_i\left(Y_i\left(z_i, z_j\right) - z_i'\right)$$

Since u_i is strictly increasing these inequalities imply that for all $z_i > z_i'$, $P\left(z_i', z_j\right) \geq P\left(z_i, z_j\right)$ and $Y_i\left(z_i, z_j\right) \geq Y_i\left(z_i', z_j\right)$, so that (Y, P) must be monotone.

Continuity: IR and SP imply that Y_i and P must be continuous at all (s_1, s_2) such that $u_i\left(Y_i\left(s_1, s_2\right) - s_i\right) > 0$, for $i = 1, 2$.

We first show that if $Y_i\left(z_1, z_2\right) \geq z_1$, and Y_i is continuous at all (z_1, z_2) such that $u_i\left(Y_i\left(z_1, z_2\right) - z_i\right) > 0$, then so must be P, and vice-versa. Assume by way of contradiction that Y_i is continuous and P is discontinuous at some (z_1, z_2). Then there is an $\varepsilon > 0$ such that for all $\delta > 0$,there is a $z_1' \in (z_1 - \delta, z_1 + \delta)$ such that$\left|P\left(z_1', z_2\right) - P\left(z_1, z_2\right)\right| > \varepsilon$. Assume wlog that $P\left(z_1', z_2\right) \geq P\left(z_1, z_2\right) + \varepsilon$. Then for δ small enough, an agent of type z_1 must be better off reporting z_1' instead of her true type z_1: the continuity of Y_i implies that the possible loss in the allocated share is negligible, while a strictly positive gain in probability of agreement is attained. Proving that continuity of P implies continuity of Y_i is analogous.

Assume that both Y and P are discontinuous at some (z_1, z_2), such that $u_i\left(Y_i\left(z_1, z_2\right) - z_i\right) > 0$, and again wlog let $P\left(z_1', z_2\right) \geq P\left(z_1, z_2\right) + \varepsilon$. To assure that z_1 reports truthfully under such discontinuous P, the discontinuity in Y_i must be such that $Y_1\left(z_1, z_2\right) \geq Y_1\left(z_1', z_2\right) + \gamma$ for some $\gamma > 0$ in order to assure that agent one report truthfully (note that it is possible to find such γ by $u_i\left(Y_i\left(z_1, z_2\right) - z_i\right) > 0$ and IR). Since $Y_2\left(z_1, z_2\right) = 1 - Y_1\left(z_1, z_2\right)$ the discontinuity of P at (z_1, z_2) is such that z_2 cannot prefer to report truthfully when facing z_1, contradicting strategy proofness.

ii)If (Y, P) also satisfies weak efficiency then it is continuous everywhere. This follows from (i), since WE and strategy proofness imply continuity at points (z_1, z_2) s.t. $u_i\left(Y_i\left(z_1, z_2\right) - z_i\right) = 0$. To see it, observe first that by monotonicity, Y_i has to be continuous in the neighbourhood of the line $z_1 + z_2 = 1$. By IR, this line is the boundary of the set of points (z_1, z_2) s.t. $u_i\left(Y_i\left(z_1, z_2\right) - z_i\right) = 0$. If P were discontinuous anywhere on the line, then by same arguments as in (i) the types at the points of discontinuity could gain by mis-representing. ■

Proof of Lemma 6:

Proof. By the Lemma 5 quasi-decent rules are continuous and monotonic, hence they are differentiable almost everywhere.

Step 1: Necessity. Consider a decent rule (Y,P). By individual rationality and weak efficiency there is no loss of generality in requiring that $Y_1(z) = z_1 + \alpha(z)(1 - z_1 - z_2)$ where $0 \leq \alpha(z) \leq 1$. The condition $P(s) = 0$ for all s such that $s_1 + s_2 > 1$ is necessary for individual rationality. Also, $0 < P(s) \leq 1$ for $s_1 + s_2 < 0$ is necessary for weak effciency.

Strategy proofness is equivalent to the requirement that for every s_i and z_j the function $U_i(z_i; s_i, z_j) = P(z_i, z_j)\, u_i(Y_i(z_i, z_j) - s_i)$ has a global maximum at the point $z_i = s_i$. Since (Y,P) must be differentiable (a.e.) by Lemma 5, the necessary first order condition for a maximum is that

$$\frac{\partial U_i(s_i; s_i, z_j)}{\partial z_i} = 0 \text{ for all } s_i, z_j.$$

That is

$$P_i(s_i, z_j)\, u_i(Y_i(s_i, z_j) - s_i) + P(s_i, z_j)\, u_i'(Y_i(s_i, z_j) - s_i)\,\frac{\partial Y_i(s_i, z_j)}{\partial z_i} = 0$$

for all s_i, z_j, and $i = 1, 2,\ j \neq i$. Substituting $\alpha^i(s_1, s_2)$ in $Y_i(s_i, s_j)$ yields (1).

We now check that $P(0,0) = 1$ is necessary. Assume by way of contradiction that $P(0,0) = \pi < 1$. The monotonicity of the mechanism implies that $P(s_1, s_2) < P < 1$ for all (s_1, s_2). Take an alternative rule (Y', P') where for all (s_1, s_2) the sharing rules are the same $Y_i'(s_1, s_2) = Y_i(s_1, s_2)$ and the probabilities of agreement $P'(s_1, s_2) = \frac{P(s_1,s_2)}{\pi}$ are increased. It is immediate to check that the rule (Y', P') still satisfies (1). Moreover, it is strictly preferred by all types, contradicting decency of (P, Y).

Step 2: Sufficiency. Consider $U_1(z_1; s_1, z_2)$. It is enough to show that for all $z_1 > s_1$ the derivative of $U_1(z_1; s_1, z_2)$ w.r.t. z_1 is decreasing whenever $U_1(z_1; s_1, z_1) > 0$ (deviations that give negative expected utility cannot be profitable). Thus compute

$$\frac{dU_1(z_1; s_1, z_2)}{dz_1} = P_1(z_1, z_2)\, u_1(Y_1(z_1, z_2) - s_1) +$$

$$+P(z_1, z_2)\, u_1'(Y_1(z_1, z_2) - s_1)\left[1 + \alpha_1(z_1, z_2)(1 - z_1 - z_2) - \alpha(z_1, z_2)\right]$$

From the first order condition we can express

$$(1 + \alpha_1(z_1, z_2)(1 - z_1 - z_2) - \alpha(z_1, z_2)) = -\frac{u_1(Y_1(z) - z_1)}{u_1'(Y_1(z) - z_1)}\frac{P_1(z)}{P(z)}$$

Substituting this into the above expression we get

$$\frac{dU_1(z_1;s_1,z_2)}{dz_1}=P_1(z)\left[u_1(Y_1(z)-s_1)-\frac{u_1'(Y_1(z)-s_1)\,u_1(Y_1(z),z_1)}{u_1'(Y_1(z)-z_1)}\right]$$

From here we see that whenever $u_1(Y_1(z_1,z_2)-s_1)>0$, $\frac{dU_1(z_1;s_1,z_2)}{dz_1}$ is a decreasing function of z_1. This follows directly from the fact that u_1 is increasing and concave in λ_1. Thus the local maximum of U_1 is unique, and is also a global maximum. Similarly for U_2. ■

Proof of Theorem 7:

A decent rule requires a solution to (1) where α^i remain in $[0,1]$. Assuring such boundedness in a neighborhood of the line $s_1=1-s_2$ uniquely selects boundary conditions that our solution must satisfy. They are stated in the following lemma.

Lemma 30 *A solution to (1) characterizes a decent rule if and only if α solves*

$$\frac{((1-\rho)\alpha_{12}-\alpha_1-\alpha_2)}{(\alpha-1-(1-\rho)\alpha_1)((1-\rho)\alpha_2-\alpha)}=\frac{B_1-B_2}{A_1+A_2} \tag{5}$$

where $A_i=\frac{u^{i\prime}(x)}{u^i(x)}$ and $B_i=(A_i)'$; with the following boundary conditions on $\rho=0$:

$$\alpha=\frac{\sqrt{\beta_1}}{\sqrt{\beta_1}+\sqrt{\beta_2}},\ \beta_i=\lim_{x\to 0}x\frac{u^{i\prime}(x)}{u^i(x)},0<\beta_i\leq 1,$$

and

$$\alpha_1=\alpha_2=h,$$

where h is an arbitrary C^1 function defined on $\rho=0$.[29]

Proof. : Write (1) as

$$P_i=-P\left[w^i\left(\rho\alpha^i\right)\left(\alpha^j-\alpha_i^j\rho\right)\right]\ ,j\neq i. \tag{6}$$

Recall that $\alpha^1=\alpha$ and rewrite (6) as

$$\begin{aligned}P_1&=-w^1(\rho\alpha)\left((1-\alpha)-\rho\alpha_1\right)P,\\ P_2&=-w^2(\rho(1-\alpha))(\alpha-\rho\alpha_2)P.\end{aligned}$$

[29]Terms in w^1 are evaluated at $\alpha(s)(1-\rho)$ and terms in w^2 evaluated at $(1-\alpha(s))(1-\rho))$.

Since $P_{12} = P_{21}$ we obtain that α must solve the second order PDE (5).

Assume that a solution to (5) exists and it is $0 < \alpha(a,b) = \alpha_0 < 1$ for $a + b = 1$, then

$$\lim_{(s_1,s_2)\to(a,b)} \frac{B_1\left(\alpha(1-\rho)\right) - B_2\left((1-\alpha)(1-\rho)\right)}{A_1\left(\alpha(1-\rho)\right) + A_2\left((1-\rho)(1-(1-\alpha)\right)} = \alpha_0.$$

Since $\lim\limits_{(s_1,s_2)\to(a,b)} A_1 + A_2$ behaves as $\frac{B_1}{\alpha(1-\rho)} + \frac{B2}{(1-\alpha)(1-\rho)}$ and since $\lim\limits_{(s_1,s_2)\to(a,b)} A_i = \beta_i$ it is necessary that

$$\lim_{(s_1,s_2)\to(a,b)} \left(\frac{-\beta_1}{\left(\alpha\left(s_1,s_2\right)\right)^2} + \frac{-\beta_2}{\left(1-\alpha\left(s_1,s_2\right)\right)^2}\right)\frac{1}{1-\rho}$$

exists and it is finite. This implies the first boundary condition $\alpha(a,b) = \frac{\sqrt{\beta_1}}{\sqrt{\beta_1}+\sqrt{\beta_2}}$. Differentiating $\alpha(z, 1-x)$ yields the boundary conditions on α_1 and α_2. ■

Proof. (of the Theorem7) The proof of the theorem is now immediate. By Lemma 30 α must solve ((5) with the appropriate boundary conditions on the line $\rho = 0$, which is non-characteristic. Whenever it exists, a solution to such equation is unique[30]. Thus we obtain unique functions $\alpha_1(s)$ and $\alpha_2(s)$ and then the linear PDE system for $P(s)$ can be integrated by the construction of α. It is easy to see that $P(s)$ is then determined uniquely up to a multiplicative constant π, setting $\pi = 1$ yields a unique decent rule. ■

Proof of Lemma 10:

Proof. It is straightforward to check that if $\alpha(s_1,s_2) = \alpha(\rho)$ and $P(s_1,s_2) = P(\rho)$ then the (1) is equivalent to

$$\alpha'(\rho) = -\frac{1}{\rho}\left(\alpha(\rho) - \frac{1}{1+g(\alpha(\rho)\rho)}\right) \tag{7}$$

and

$$\frac{P'(\rho)}{P(\rho)} = -w^1\left(\alpha(\rho)\rho\right)\left(1-\alpha'(\rho)\rho - \alpha(\rho)\right) \tag{8}$$

where we write

$$g(\alpha(\rho)\rho) \equiv \frac{u_1\left(\alpha(\rho)\rho\right)u_2'((1-\alpha(\rho))\rho)}{u_1'\left(\alpha(\rho)\rho + s_1, s_1\right)u_2((1-\alpha(\rho))\rho)}.$$

[30]See John [1971].

Consider (7). First notice that in order for $\alpha\left(.\right)$ to be only a function of ρ then the implicit equation for $\alpha\left(\rho\right)$

$$\alpha\left(\rho\right)=\frac{1}{1+g(\alpha\left(\rho\right)\rho)}$$

either admits a constant solution, or no solution to (7) satisfies $0\leq\alpha\left(\rho\right)\leq1$.

In order for $\alpha\left(\rho\right)$ to be bounded at $\rho=0$ it has to be that $\alpha\left(\rho\right)$ is the singular solution of the above equation (7), i.e. the solution satisfying $\alpha\left(0\right)=\frac{1}{1+g(0)}$. The singular solution satisfies for every ρ the condition $\alpha\left(\rho\right)-\frac{1}{1+g(\alpha(\rho)\rho)}=0$. Thus the only case where a singular solution exists (meaning that such function indeed solves the differential equation) is the case where $\alpha\left(.\right)$ is the constant α^{*}, solving the equation $\alpha^{*}=\frac{1}{1+g(\alpha^{*}\rho)}$. ■

Proof of Proposition 11:

Proof. It is trivial to check that when $u_i=Cu_j$, where $C>0$, $\alpha=\frac{1}{2}$ solves equation(1). We plug $\alpha=\frac{1}{2}$ into the equation (1) of Lemma 6 to obtain

$$\frac{\frac{\partial P(s)}{\partial s_i}}{P\left(s\right)}=-\frac{u_i'\left(\frac{1}{2}\rho\right)}{2\,u_i\left(\frac{1}{2}\rho\right)};\ i=1,2.$$

By integrating this simple system we obtain the solution for $P\left(s\right)$

$$P\left(s\right)=\pi\frac{u_i\left(\frac{1}{2}\rho\right)}{u_i\left(\frac{1}{2}\right)},\ \pi\in(0,\pi].$$

Now we apply Theorem 7 which proves the part of the statement for equal utilities.

We could proceed similarly as above to show that the mechanism defined by

$$\begin{aligned}\alpha^{*}&=\frac{\gamma_1}{\gamma_1+\sqrt{\gamma_2\gamma_1}}\\ P^{*}\left(\rho\right)&=\begin{cases}e^{-D_1\alpha_1 s_1-D_2\alpha_2 s_2}\rho^{\sqrt{\gamma_1\gamma_2}};\ s_1+s_2\leq1\\ 0;\ \text{otherwise.}\end{cases},\end{aligned}$$

is decent when the utilities are of the form

$$\begin{aligned}u_i\left(\lambda_i,s_i\right)&=C_i\left(\lambda_i-s_i\right)^{\gamma_i}e^{D_i(\lambda_i-s_i)},\\ \gamma_i&\in(0,1],\ D_i\in\left[-\gamma_i-\sqrt{\gamma_i},\sqrt{\gamma_i}-\gamma_i\right].\end{aligned}$$

Now we prove that there for no other utilities there exists a constant net surplus rule. So assume that $u_1 \neq u_2$ and take $\alpha_i = \alpha_i^* = const$. Then the equation (1) simplifies to

$$\frac{\left(u_1'\left(x_1\right)\right)^2 - u_1''\left(x_1\right)u_1\left(x_1\right)}{\left[u_1\left(x_1\right)\right]^2} = \frac{\left(u_2'\left(x_2\right)\right)^2 - u_2''\left(x_2\right)u_2\left(x_2\right)}{\left[u_2\left(x_2\right)\right]^2},$$

which is in fact

$$\left(\frac{u_1'\left(x_1\right)}{u_1\left(x_1\right)}\right)' = \left(\frac{u_2'\left(x_2\right)}{u_2\left(x_2\right)}\right)'. \tag{9}$$

If $\alpha_i = \frac{1}{2}$ it is easily seen that the equation (9) implies $u_1 = Cu_2$, $C > 0$. So assume that $\alpha_i \neq \frac{1}{2}$. Then (9) implies that $\left(\frac{u_i'(x_i)}{u_i(x_i)}\right)' = \bar{h}_i\left(x_i\right)$ where $\bar{h}_i\left(x_i\right)$ has the property $\bar{h}_i\left(tx\right) = K_i g_i\left(t\right)\tilde{h}\left(x\right)$ - otherwise α_i would necessarily be a function of ρ. Reversing the roles of t and x this means that $\bar{h}_i\left(tx\right) = \bar{h}_j\left(tx\right) = \tilde{h}\left(tx\right) = \tilde{h}\left(t\right)\tilde{h}\left(x\right)$. It is easy to see that the only class of functions for which the last equality holds is the class $\tilde{h}\left(x\right) = K_i x^h$, where h is a constant. Next, we also get that in order for concavity of u to hold, it has to be that $\lim_{x\to 0} u_i\left(x\right) = O\left(x^{-\frac{h}{2}}\right)$. But such u will only be concave in the neighborhood of $x = 0$ if $h \in [-2, 0]$. Integrating

$$\left(\frac{u_i'\left(x\right)}{u_i\left(x\right)}\right)' = K_i x^h$$

we obtain for $h > -2$ utilities that cannot satisfy the requirement $u_i\left(0\right) = 0$. For $h = -2$ we obtain precisely the above class of utilities. ■

Proof of Lemma 14:

Proof. Denote $L_i\left(s_i\right) = \lim_{t\to\infty}\lambda_i\left(s_i, t\right)$. The proof is divided into three steps. In step 1 we show that $L_i\left(1\right) = 1$ (which holds trivially) and the continuity at 1 imply that $L_i\left(0\right) = 0$. In step 2 we show that $L_i\left(.\right)$ is a continuous function, hence it attains all values in the interval $[0, 1]$. Finally, in step 3 we show that the statement of the lemma is true.

Step 1: $L_i\left(0\right) = 0$. Suppose this did not hold, i.e. $L_i\left(0\right) = K > 0$ in equilibrium. Denote by $\lambda_i\left(0, t\right)$ such equilibrium strategy of player i, and by $\lambda_j\left(s_j, t\right)$ the equilibrium strategy of player j, when his type is s_j. By individual rationality we have that $L_j\left(1\right) = 1$. Also by individual rationality, we have that $L_j\left(s_j\right)$ is bounded below, i.e. $L_j\left(s_j\right) \geq s_j$. Since $L_j\left(s_j\right) \leq 1$, these imply that $L_j\left(s_j\right)$ is continuous at point $s_j = 1$. From continuity of

L_j around $s_j = 1$ we get that there is a positive mass of types $s_j \in [0,1]$ for which $L_j(s_j) > 1 - K$. But then type 0 of agent i could improve her expected payoff by playing λ_i until some large time t', and then lowering her demand to 0, according to some strategy λ_i'. To see this, notice that λ_i and λ_j are continuous and for all s_j, $\lambda_i(0,t)$ and $\lambda_j(s_j,t)$ are non-increasing in t . Thus the support of $f_j(s_j|t)$ is shrinking as time elapses. When t is very large, the support of $f_j(s_j|t)$ will be very close to the ex-post belief when no agreement has been reached. Hence t' is given as the moment when the expected continuation payoff of playing λ_i, conditional on $s_j \leq 1 - K$, is lower than the expected continuation payoff of playing λ_i', conditional on $s_j < 1$. This establishes the contradiction. The same argument shows that $L_i(s_i)$ is continuous in a neighbourhood of the point $s_i = 0$.

Step 2. Assume thus that $L_i(s_i)$ is discontinuous at $\bar{s}_i$, i.e. $L_i(\bar{s}_i) = \hat{l}$ and $\lim_{s_i \searrow \bar{s}_i} = \bar{l}$, where $\bar{l} > \hat{l}$. Then there must exist an $\bar{s}_j$ s.t. $L_j(\bar{s}_j) = 1 - \bar{l}$, and $\lim_{s_j \searrow \bar{s}_j} L_j(s_j) = 1 - \hat{l}$ (same argument as in Step 1, and left-continuity of L_i and L_j). Take any $\hat{s}_i > \bar{s}_i$. By continuity of λ_i in t, there exists an M_i s.t. $\lambda_i(\bar{s}_i, t) - \hat{l} < \varepsilon$ for all $t \geq M_i$. Also, notice that $\lambda_i(\hat{s}_i, t) \geq \bar{l}$. Now fix $\varepsilon = \frac{\bar{l} - \hat{l}}{4} > 0$ and take a $t \geq M_i$. Then at t, $\lambda_i(\bar{s}_i, t) < \hat{l} + \varepsilon$ while $\lambda_i(\hat{s}_i, t) \geq \bar{l}$ for all $\hat{s}_i > \bar{s}_i$, contradicting the continuity of λ_i in s_i. This proves that $L_i(s_i)$ has to be right-continuous. By assumption, $L_i(s_i)$ is left-continuous[31], hence it is continuous. In step 1 we proved that $L_i(1) = 1$ and $L_i(0) = 0$, so by Rolle's theorem it attains all values between 0 and 1.

Step 3: $L_i(s_i) = s_i$ for all $s_i \in [0,1]$. Take an $s_i \in (0,1)$. By steps 1 and 2, L_i takes all the values in the interval $[0,1]$ and is continuous (thus measurable), strictly positive on $(0,1]$. Thus we can define the measure μ_i

$$\mu_i(S) = \int_S L_i(s)\, dm(s) \text{ for any mesaurable } S \subset [0,1],$$

where $m(.)$ denotes the usual Lebesgue measure. By strict positivity, continuity, and boundedness of $L_i(s_i)$, μ_i is an equivalent measure to m. Now

[31] Type $\bar{s}_i$ is at $t = \infty$ indifferent between demanding $\hat{l}$ and $\bar{l}$; the former doesn't improve her probability of reaching an agreement since the mass of opposing types with demands between $1 - \bar{l}$ and $1 - \hat{l}$ is 0. However, by an argument similar to the proof of Step 1, we can argue, that she doesn't lose anything by bidding $\hat{l}$, which gives us left-continuity of L_i. Left-continuity of L_i is thus essentially an assumption on how agents resolve their indifference at the horizon.

suppose that $L_i(s_i) > s_i$. By equivalence of μ_i to m there exists a positive mass of types s_j s.t. $L_j(s_j) \in (1 - L_i(s_i), 1 - s_i)$. To see this define $B = \{s_j | L_j(s_j) \in (1 - L_i(s_i), 1 - s_i)\}$. Since μ_i and m are equivalent, $m(B) > 0$. Now repeat the same argument as in Step 1 to get a contradiction. Hence indeed $L_i(s_i) = s_i$. ■

PROOF OF LEMMA 15:

Proof. i) Let $t_i^E(s_i) = \inf\{t > 0 | \lambda_i(s_i, t) + \lambda_j(s_j, t) \leq 1 \text{ for some } s_j \in [0, 1]\}$. Then for no $t > t_i^E(s_i)$ there exists an interval $(\tilde{s}_i, \bar{s}_i)$s.t. $\lambda_i(s_i, t) = \lambda_i(s_i', t)$ for all $s_i' \in (\tilde{s}_i, \bar{s}_i)$.

Suppose that $\lambda_i(s_i, t)$ is strictly decreasing in s_i at t. By lemma14 it is enough to show that for no $t > 0$ s.t. $\lambda_j(s_j, t) = 1 - \lambda_i(s_i, t)$ for some s_j, it could hold that $\frac{\partial \lambda_i(s_i,t)}{\partial s_i} = 0$ for $s_i \in (\underline{s}_i, \overline{s}_i)$ for some $\underline{s}_i < \overline{s}_i$. This so because of the continuity of $\lambda_i(s_i, t)$ with respect to both parameters and because $L_i(s_i)$ is increasing in s_i. Fix $t > 0$ and assume that $\lambda_i(s, t) = \widehat{\lambda}$ for all $s \in [\underline{s}_i, \overline{s}_i], 0 < \widehat{\lambda} \leq \lambda_i(s_i, 0) \leq 1, \underline{s}_i < \overline{s}_i$. Since i demands $\widehat{\lambda}$ with positive probability at t, there is some $\mu > 0$ s.t. $\lambda_j(s_j, \tau) \geq 1 - \widehat{\lambda}$ for all s_j and $\tau \in (t - \mu, t)$: note that if $\lambda_j(s_j, \tau) < 1 - \widehat{\lambda}$ for all $\tau \in (t - \mu, t)$, then s_j is reaching an agreement worse than $1 - \widehat{\lambda}$ with all $s \in [\underline{s}_i, \overline{s}_i]$ while the better agreement $1 - \widehat{\lambda}$ is feasible with the same counterparts (that have positive measure) only at the (negligible) cost of delaying agreement with types $s_i < \underline{s}_i$ (whose measure is negligible as $\mu \to 0$). Now, since strategies are decreasing in t, $\lambda_j(s_j, \tau) \geq 1 - \widehat{\lambda}$ for all $\tau \geq 0$ and $\lambda_j(s_j^l, 0) \geq 1 - \widehat{\lambda}$. Hence any agreement in $[0, t)$ must give i at most $\widehat{\lambda}$, and $s \in [\underline{s}_i, \overline{s}_i]$ do not reach any agreement in $[0, t)$,contradicting that for each $s \in [\underline{s}_i, \overline{s}_i]$ there is s_j and $\tau \in [0, t)$ such that $\lambda_j(s_j, \tau) = 1 - \lambda_i(s_i, \tau)$.

ii) $\frac{\partial \lambda_j}{\partial s_j}(s_j, t) \geq 0$ for all s_j, all $t \geq 0$ implies $\frac{\partial \lambda_i}{\partial s_i} \geq 0$ for all s_i, all $t \geq 0$.

Recall that at time t, agent i maximizes

$$\max \int_t^\infty e^{-\tau} u_i (\lambda_i(s_i, \tau) - s_i) \, d\mu(s_i, \tau) \tag{10}$$

where $\mu(s_i, \tau)$ is the measure of types of player j with whom agent i reached agreement by time τ. We proceed by considering a specific deviation of player i at time t, from $\lambda_i(s_i, \tau)$ to $\bar{\lambda}_i^\delta(s_i, \tau)$. For any $\delta > 0$, we construct $\bar{\lambda}_i^\delta(s_i, \tau)$ as follows. Take $\bar{s}_i > s_i$, $t' > t$ and define $\bar{s}_i(\tau)$ as a differentiable

$i \quad ' \quad i \quad i \quad ' \quad i$

$$i \quad i \quad \begin{cases} i \quad i \quad \leq \quad ' \\ i \quad i \quad ' \leq \quad ' \\ i \quad i \quad ' \quad \leq \quad \infty \end{cases}$$

$\frac{\partial \lambda_j}{\partial s_j} \geq \quad \frac{\partial \lambda_j}{\partial s_j} \quad j \quad i$

$j \quad i$

$\frac{\partial \lambda_j}{\partial s_j}$

$\frac{\partial \lambda_j}{\partial s_j} \quad j \quad i \quad \frac{\partial}{\partial \tau} \quad i$

$i \quad i$

$$\int_t^\infty \quad -\tau \quad i \quad i \quad i \quad - \quad i \quad i \quad \geq \int_t^\infty \quad -\tau \quad i \Big(\quad i \quad i \quad - \quad i \Big) \quad i$$

i

i

$$\int_t^{t'+\delta} \quad -\tau \quad i \quad i \quad i \quad - \quad i \quad i \quad \geq \quad \int_t^{t'} \quad -\tau \quad i \quad i \quad i \quad - \quad i \quad i$$

$$\int_{t'}^{t'+\delta} \quad -\tau \quad i \quad i \quad i \quad - \quad i$$

$$\| \quad i \quad ' - \Big(\quad i \quad \Big) \| \quad \Rightarrow \| \quad {}^{i}_{i} \quad ' - \quad \Big(\quad i \quad \Big) \|$$

$\int_{t'}^{t'+\delta} \quad -\tau \quad i \quad i \quad i \quad - \quad i \quad \leq$

$-T' \quad i \quad i \quad ' \quad {}'_i - \quad i \quad \rightarrow$

$$\int_t^{t'} \quad -\tau \quad i \quad i \quad i \quad - \quad i \quad i \quad \geq \int_t^{t'} \quad -\tau \quad i \quad i \quad i \quad - \quad i \quad i$$

$' \quad i \quad i$

$' \rightarrow$

π

$$- \quad \geq \quad -$$

$$- \quad \geq \quad -$$

From these two inequalities it follows that $\lambda_i(\bar{s}_i, t) \geq \lambda_i(s_i, t)$ and $m(s_i, t) \geq m(s_i', t)$, which completes the proof. ■

Proof of Lemma 19:

Proof. We fix s_i and economize the notation to write $\sigma_j(s_i, t) = \sigma_j(t)$ and $\dfrac{\partial \sigma_j(s_i, t)}{\partial t} = \dot{\sigma_j}(t)$. We write the Hamiltonian

$$\begin{aligned} H_i(t) &= e^{-t} u_i(\lambda_i(s_i, t) - s_i) f_j(\sigma_j(t))\, \dot{\sigma_j}\,(t) - \\ &\quad -\mu(t)\,(1 - \lambda_j(\sigma_j(t), t) - \lambda_i(s_i, t)) \end{aligned}$$

and compute the Euler conditions for the unknown functions

$$\begin{aligned} \frac{\partial H_i}{\partial \sigma_j} &= e^{-t} u_i(\lambda_i(s_i, t) - s_i) f_j'(\sigma_j)\, \dot{\sigma_j} + \mu \frac{\partial \lambda_j(\sigma_j, t)}{\partial \sigma_j} \\ \frac{d}{dt}\frac{\partial H_i}{\partial\, \dot{\sigma_j}} &= e^{-t} u_i(\lambda_i(s_i, t) - s_i) f_j'(\sigma_j(t))\, \dot{\sigma_j} + e^{-t} u_i'(\lambda_i(s_i, t) - s_i) \frac{\partial \lambda_i(s_i, t)}{\partial t} f_j(\sigma_j) + \\ &\quad -e^{-t} u_i(\lambda_{i(}s_i, t) - s_i) f_j(\sigma_j) \\ \frac{\partial H_i}{\partial \lambda_i} &= e^{-t} u_i'(\lambda_i(s_i, t) - s_i) f_j(\sigma_j)\, \dot{\sigma_j} + \mu \\ \frac{\partial H_i}{\partial \dot{\lambda}_i} &= 0 \end{aligned}$$

Whence we have the two Euler equations

$$\begin{aligned} &\mu \tfrac{\partial \lambda_j(\sigma_j, t)}{\partial \sigma_j} - e^{-t} u_i'(\lambda_i(s_i, t) - s_i) \tfrac{\partial \lambda_i(s_i, t)}{\partial t} f_j(\sigma_j)) + e^{-t} u_i(\lambda_i(s_i, t) - s_i) f_j(\sigma_j) = 0 \\ &e^{-t} u_i'(\lambda_i(s_i, t) - s_i) f_j(\sigma_j)\, \dot{\sigma_j} + \mu = 0 \end{aligned}$$

From the second Euler equation we can eliminate μ and the density f_j also disappears from the first to obtain the final condition

$$(u_i'(\lambda_i(s_i, t) - s_i))\,(\frac{\partial \lambda_j(\sigma_j, t)}{\partial \sigma_j}\, \dot{\sigma_j} + \frac{\partial \lambda_i(s_i, t)}{\partial t}) + u_i(\lambda_{i(}s_i, t) - s_i) = 0$$

or equivalently

$$\begin{aligned} u_i(\lambda_i, s_i) &= u_i'(\lambda_i(s_i, t) - s_i) \left(\frac{\partial \lambda_j(\sigma_j, t)}{\partial s_j} \frac{d\sigma_j}{dt} + \frac{\partial \lambda_i}{\partial t} \right), \\ \text{for } t &\geq t_E(s_i). \end{aligned}$$

■

Proof of Proposition 20:

Proof. The strategy $\bar{\lambda}_i(s_i,\tau)$ maximizes the expected gain, that is

$$\bar{\lambda}_i(s_i,\tau) \in \arg\max \int_{t_i^E(s_i)}^{\infty} e^{-\tau} u_i(\bar{\lambda}_i(s_i,\tau)-s_i) f_j(\bar{\sigma}_j(s_i,\tau)) \frac{\partial \bar{\sigma}_j(s_i,\tau)}{\partial \tau} d\tau$$

$$\text{subject to } \bar{\sigma}_j\left(s_i, \tau_i^E(s_i)\right) = 0.$$

Now make a substitution $\delta(t) = e^{-\tau}$, so that $\dot{\delta}(t)\,dt = -e^{-\tau}d\tau$. Also, defining $\sigma_j(s_i,t) = \bar{\sigma}_j(s_i,\tau)$, we have $\frac{\partial \bar{\sigma}_j(s_i,\tau)}{\partial \tau} = \frac{\partial \sigma_j(s_i,t)}{\partial t}\frac{dt}{d\tau}$. Inserting all this into the above program we see that then $\bar{\lambda}_i(s_i,-\ln(\delta(t)))$ maximizes

$$\int_{t_i^E(s_i)}^{\infty} \delta(t)\, u_i(\bar{\lambda}_i(s_i,-\ln(\delta(t))) - s_i) f_j(\sigma_j(s_i,t)) \frac{\partial \sigma_j(s_i,t)}{\partial t} dt$$

$$\text{subject to } \sigma_j\left(s_i, t_i^E(s_i)\right) = 0;$$

which completes the proof. ■

Proof of Proposition 24:

Proof. We have already shown that the regular equilibria of the FD game implement decent rules. What we still need to show is that by taking all possible solutions of the first order condition for the FD game we get all possible decent mechanisms. We will show this by demonstrating that the equilibria of the FD game translate into decent rules via a simple substitution. So take strategies $\lambda_1(s_1,t)$ and $\lambda_2(s_2,t)$ that solve

$$\frac{\partial u_i(\lambda_i,s_i)}{\partial \lambda_i}\left(\frac{\partial \lambda_j(\sigma_j,t)}{\partial s_j}\frac{\partial \sigma_j}{\partial t} + \frac{\partial \lambda_i}{\partial t}\right) = u_i(\lambda_i,s_i) \tag{11}$$

$$\text{s.t.}\lambda_i(s_i,t) + \lambda_j(\sigma_j,t) = 1 \tag{12}$$

Implicitly derive the relationship (12) on t to get

$$\frac{\partial \lambda_j(\sigma_j,t)}{\partial s_j}\frac{\partial \sigma_j}{\partial t} + \frac{\partial \lambda_i}{\partial t} = -\frac{\partial \lambda_j}{\partial t} \tag{13}$$

Define $Y_j(s_j, s_i) = \lambda_j(s_j, t(s_i, s_j))$, where $t(s_i, s_j)$ is defined from the relationship $\lambda_i(s_i, t) + \lambda_j(s_j, t) = 1$. Thus $\frac{\partial \lambda_j}{\partial t}\frac{\partial t}{\partial s_i} = \frac{\partial Y_j}{\partial s_i}$, hence

$$\frac{\partial \lambda_j}{\partial t} = \frac{\partial Y_j}{\partial s_i}\frac{1}{\frac{\partial t}{\partial s_i}} \tag{14}$$

Now substitute (13), (12), and (14) into (11) to obtain

$$-\frac{\partial u_i(Y_i, s_i)}{\partial Y_i}\frac{\partial Y_j}{\partial s_i}\frac{1}{\frac{\partial t}{\partial s_i}} = u_i(Y_i, s_i) \tag{15}$$

Since $Y_j(s_j, s_i) + Y_i(s_i, s_j) = 1$, we have that $\frac{\partial Y_j}{\partial s_i} = -\frac{\partial Y_i}{\partial s_i}$. Now interpret the probability of implementation as the discount due to delay, that is $P(s_i, s_j) = e^{-t(s_i, s_j)}$. Hence

$$\frac{\partial P(s_i, s_j)}{\partial s_i} = -e^{t(s_i, s_j)}\frac{\partial t(s_i, s_j)}{\partial s_i} = -P(s_i, s_j)\frac{\partial t(s_i, s_j)}{\partial s_i}$$

Thus

$$\frac{\partial t(s_i, s_j)}{\partial s_i} = -\frac{\frac{\partial P(s_i, s_j)}{\partial s_i}}{P(s_i, s_j)}$$

Plugging all of this into (15) we get that $\lambda_i(.)$ and $\lambda_j(.)$ satisfy (11) if and only if $Y_i(.)$, $Y_j(.)$, and $P(s_i, s_j)$ satisfy the first order condition

$$-\frac{\partial u_i(Y_i, s_i)}{\partial Y_i}\frac{\partial Y_i}{\partial s_i} = \frac{\frac{\partial P(s_i, s_j)}{\partial s_i}}{P(s_i, s_j)}u_i(Y_i, s_i)$$

This completes the proof. ■

PROOF OF PROPOSITION 26:

Proof. Observe that the proposed strategies are regular and satisfy Lemmas 13 to 16. Therefore, to show that they constitute an equilibrium of the FD game, it suffices to check that they satisfy (3). Let α^* be the constant $\alpha^* = \frac{\gamma_1}{\gamma_1 + \sqrt{\gamma_1 \gamma_2}}$, and denote $h(t) = e^{-\frac{t}{\sqrt{\gamma_1 \gamma_2}}}$. We check (3) for λ_1, the calculus for λ_2 is analogous. Now

$$\begin{aligned}
\sigma_2(s_1, t) &= 1 - s_1 - h(t) \\
\frac{\partial \lambda_2(\sigma_2, t)}{\partial s_2} &= 1, \ \frac{\partial \sigma_2}{\partial t} = -\dot{h}(t) \\
\frac{\partial \lambda_1}{\partial t} &= \alpha^* \dot{h}(t)
\end{aligned}$$

so by substituting λ_1 into the first order condition, we get

$$\gamma_1 \left(\lambda_1 - s_1\right)^{\gamma_1 - 1} \left(-\dot{h} + \alpha^* \dot{h}\right) = \left(\lambda_1 - s_1\right)^{\gamma_1}$$

Noticing that $\lambda_1 - s_1 = \alpha^* h$. This simplifies into

$$\frac{\alpha^*}{\gamma_1 \left(1 - \alpha^*\right)} = -\frac{\dot{h}}{h}.$$

Deriving the analogous expression from (3) for agent 2 yields

$$\frac{\left(1 - \alpha^*\right)}{\gamma_2 \alpha^*} = -\frac{\dot{h}}{h}.$$

Substituting h and α^*, it is immediate that (3) holds. ■