Collegio Carlo Alberto

Trading Favors: Optimal Exchange and Forgiveness

Christine Hauser
Hugo Hopenhayn

Working Paper No. 88
December 2008
www.carloalberto.org

Trading Favors: Optimal Exchange and Forgiveness

Christine Hauser[1] Hugo Hopenhayn [2]

September 2008[3]

[1] Collegio Carlo Alberto and CHILD
[2] UCLA

Abstract

How is cooperation without immediate reciprocity sustained in a long term relationship? We study the case of two players in continuous time who have privately observable opportunities to provide favors, and where the arrival of these opportunities is a poisson process. Favors provided by a player give her an entitlement to future favors from her partner. As opposed to a "chips mechanism" where the rate of exchange of favors is one, we allow for two features: first, for the rate of exchange to depend on current entitlements, and second, for the possibility of depreciation or appreciation of entitlements. We show that these two features allow for considerably higher payoffs. We characterize and solve for the Pareto frontier of Public Perfect Equilibria (PPE) and show that it is self-generating. This guarantees that the equilibrium is renegotiation proof. We also find that optimal PPE have two key characteristics: 1) the relative price of favors decreases with a player's entitlement and 2) the disadvantaged player's utility increases over time during periods of no trade, so in the optimal equilibria there is forgiveness.

JEL Classification: C73, C63, D81
Keywords: repeated games, jump process, continuous time

1 Introduction

How is cooperation without immediate reciprocity sustained in a long term relationship? Consider the following example: Two firms are engaged in a joint venture. At random times one of them finds a discovery; if disclosed, the discovering firm's payoffs will be lower but total payoffs higher. In the absence of immediate reciprocity, will the expectation of future cooperation induce disclosure? And how much cooperation can be supported? In the case of perfect monitoring -where the arrival of the discovery is jointly observed- full cooperation can be supported with trigger strategies when discounting is not too strong. But in the case of imperfect monitoring -where the arrival of the discovery is privately observed- the first best cannot be supported and the above questions have no definite answer. This paper attempts to fill this gap.

The setup is as follows.[1] Two players interact indefinitely in continuous time. At random arrival times (following a Poisson processes with arrival rate α), one player has the possibility of providing a benefit b to the other player at a cost $c < b$.[2] Following Mobius (2001) we call this a *favor.* This opportunity is privately observed, so a player will be willing to do this favor only if this gives her an entitlement to future favors from the other player. Formally, this model is a repeated game with incomplete monitoring with random time intervals (given by the arrival of favors). We characterize and solve for the Pareto frontier of Public Perfect Equilibria (PPE) of that game.

Mobius (2001) considers a simple class of PPE, which we will call the *simple chips mechanism.*[3] Both players start with K chips each. Whenever one receives a favor, she gives the other player a chip. If a player runs out of chips (so the other one has $2K$), she receives no favor until she grants one to the other player and obtains a chip in exchange. Given that the arrival of a favor is private information, giving favors must be voluntary, and motivated by the claim to future favors. Incentive compatibility and discounting then put a limit on the number of chips that can exist in the economy. This is obviously a convenient and straightforward mechanism. However, it has two special features which suggest that there is room for improvement. First, the exchange rate is always one (current) for one (future) favor, so it is independent of the distribution of chips. Due to discounting, a player that is entitled to many future favors will value a marginal favor less. This suggests that the rate of exchange (or relative price) of favors should depend on current entitlements. Secondly, entitlements do not change unless a favor is granted (chips do not change hands when no favor has been observed). This is a special feature that rules out the possibility of appreciation or depreciation of claims. As we show in the paper, relaxing these two features allows for higher payoffs.

Our analysis proceeds in several steps. As usual in the literature, the recursive approach introduced by Abreu, Pearce and Stacchetti (1990) is used. Typically, in a discrete time

[1] Our setup is identical to Mobius (2001), with very minor modifications.

[2] The motivating example can be easily accommodated by letting b denote the value to firm 2 when firm 1 shares information and c the decrease in firm 1's profits when it discloses the information instead of keeping it secret.

[3] This type of mechanism is used in Skrzypacz and Hopenhayn (2004) in repeated auctions with incomplete information.

setting, payoffs are factorized into current period strategies and continuation values for each public signal observed. We adapt this approach to our continuous time framework by first predefining a horizon and determining, for any random arrival time of a favor within that horizon, the size of the favor to be done (for each player) and continuation values in such events, along with continuation values at the end of the horizon in case no favor has been exchanged. This can be done for any time horizon, and in particular, for one which shrinks to 0.

We first establish that the set of Pareto optimal PPE payoffs is self-generating. This is not generally true in games of private information and relies on some special features of our formulation which we discuss in the text. Intuitively, the threat of a punishment *when the good signal is not observed* (here, the non-occurrence of a favor) secures greater incentives for providing favors (and smaller necessary compensations for doing so). However, the gains from such a threat are quickly offset by the error risk when the probability of a no-arrival becomes too large relative to that of an arrival (typical in the case of a Poisson arrival process). Moreover, self-generation also guarantees that the equilibrium is renegotiation proof. As a result, the recursive formulation reduces to a one dimensional dynamic programming problem which is solved by a simple algorithm.

Optimal PPE have two key features: 1) The relative price of favors decreases with a player's entitlement. So starting from an initial symmetric point (the analogue of players having an equal number of chips), if a player receives any number of consecutive favors, he must give back a considerably greater number of favors before returning to the initial point. 2) The entitlements change over time even in periods with no trade. This is obtained through a *drift* in players' values while no favors are done, as opposed to discrete *jumps* when favors do occur.

We solve the model numerically for a large set of parameter values and find that the gains relative to the chips mechanism can be quite large (in some cases over 30% higher expected values). Interestingly, in all our numerical simulations the disadvantaged player's utility increases over time during periods of no trade, so in the optimal equilibrium there is forgiveness. We relate this observation to existing theories on the optimality of positive inflation in some monetary systems (Levine (1991), Green and Zhou (2005), Kehoe, Levine and Woodford (1992)).

Our model is a continuous time, repeated game with imperfect monitoring. This is a class of games that has not been widely analyzed in the past. Sannikov (2007) studies games where the stochastic component follows independent diffusion processes and provides a differential equation that characterizes the boundary of the set of PPE and Sannikov (2008) solves for the optimal principal-agent contract in continuous time, where the agent's output is a Brownian motion with a drift which depends on her effort. Our model does not fit exactly in that class since our stochastic process is a jump process (Poisson arrivals), yet we also derive a differential formulation to characterize the boundary of the set of PPE. In addition we establish that the set of Pareto optimal PPE is self-generating, which as far as we know is a new result in the literature on repeated games with imperfect monitoring. Kalesnik (2005) develops a similar methodology to Sannikov (2007) for Poisson signals, and characterizes the optimal strategies and payoffs in the limiting case where the signal becomes increasingly frequent and informative.

In our model there is a lack of double coincidence of needs, as players cannot instantaneously return the favors received. As suggested in a related paper by Abdulkadiroglu and Bagwell (2004), players give favors trusting that the receiver will have an incentive to reciprocate. Their game is set in discrete time, and agents can reciprocate favors either within the same period or in the future. The authors focus on symmetric self-generating lines along which total payoffs are constant and find similar properties to ours. However, the strategies and achieved payoffs remain suboptimal due to the strong restriction on the class of equilibria. In our setting, the lack of observability of opportunities for exchange is a complicating factor that limits the possibilities of exchange. Still, it can be easily shown that as the discount rate goes to zero (or the frequency of trading opportunities goes to infinity) the cost of this informational friction disappears.

Finally, our model can be reinterpreted as a moral hazard problem. Agents may choose to exert effort or not at a cost $c\alpha$. If they choose to do so, the other agent may receive a reward of value b with a Poisson arrival rate α. If no effort is exerted the arrival rate is zero. This is a special case of the *good news* scenario of Abreu, Milgrom and Pearce (1991) in a bilateral game.

The paper is organized as follows. Section 2 describes the model. Section 3 describes in more detail the simple chips mechanism. Section 4 develops the recursive formulation. Section 5 describes the solution algorithm and provides numerical results.

2 Model

We analyze an infinite horizon, two-agent partnership. Time is continuous and agents discount future utility at a rate r. There are two symmetric and independent Poisson processes -one for each agent- with arrival rate α representing the opportunity of producing a favor. We assume that favors are perfectly divisible, so partners can provide fractional favors.[4] Agents' utilities and costs are linear in the amount of favors exchanged. The cost per unit of a favor is c and the corresponding benefit to the other player $b > c$. Letting x_i represent a favor granted by player i and x_j a favor received, the utility for player i is given by:

$$U_i(x_i, x_j) = -cx_i + bx_j.$$

Since arrivals are Poisson and independent, only one player is able to grant a favor at a point in time, so $x_i(t) > 0$ implies $x_j(t) = 0$. Arrivals are privately observed by each player, so the ability to provide a favor is private information. Since the cost of providing a favor is less than the generated benefit, it is socially optimal for agents to always grant favors. Indeed, in the absence of informational constraints, a public perfect equilibrium would exist that achieves this optimum through a simple Nash reversion strategy: an agent grants favors whenever she can, as long as her partner has done so in the past, and stops granting favors

[4] Alternatively, given our assumption of linear utilities, we can assume there is public randomization for the provision of favors.

whenever her partner has defected. This equilibrium can be supported for

$$c < \left(\frac{\alpha}{r+\alpha}\right) b$$

The problem with private information is that an agent only observes whether her partner has provided a favor or not, but is unable to detect a deviation where her partner has passed the opportunity to do a favor. The question then becomes: how to ensure the maximum cooperation and exchange of favors between agents given these informational constraints?

3 A Simple Debt Accounting Mechanism

In his paper, Mobius (2001) considers equilibria of a simple class. The equilibrium proposed is Markov perfect, where the state variable is the difference k between the number of favors granted by agent 1, and those granted by agent 2. We call this a simple chips mechanism (SCM). For $-K \leq k \leq K$, the agents obey the following strategies:

- Agent 1 grants favors if $k < K$, and stops granting favors if $k = K$
- Agent 2 grants favors if $k > -K$, and stops granting favors is $k = -K$

It is obvious that the magnitude of K is crucial in determining the expected payoffs of players. Since it is efficient to have favors done whenever possible, an efficiency loss occurs when agents reach the boundaries K and $-K$, where only the indebted agent is granting favors. The larger K is, the lower is the incidence of this situation, and the larger are the expected payoffs of the agents. On the other hand, since one favor done today is rewarded by the promise of exactly one favor in the future, K cannot be infinitely large. To understand how K is determined, note that because of discounting the marginal value of the right to an extra favor diminishes with the current entitlement of favors the player has; $2K$ is the largest number such that this marginal value exceeds the cost c.

The SCM is very easy to implement. Moreover it is asymptotically efficient (as $\alpha/r \rightarrow \infty$). To see this, note that in the long run the distribution over the states $k = \{-K, -K+1, ..., 0, 1, ..., K-1, K\}$ is uniform, so the probability that a favor is not granted is $1/K$. It is easy to verify that $K \rightarrow \infty$ as $\alpha/r \rightarrow \infty$, so this probability goes to zero.

There are two special features of this scheme which suggest that there is room for improvement. The first special feature is that the rate of exchange of current for future favors is the same (equal to one) regardless of entitlements. Relaxing this constraint could reduce the region of inefficiency for two reasons. First, consider the case where the state is K so agent one has the maximum entitlement of favors. As we argued before, at this point agent one's marginal value to an entitlement of an extra favor is lower than c. But there is still room for incentives if agent two were to promise more than one favor in exchange. Moreover, in the SCM the incentive constraints only bind at the extremes, but are slack in between, where the marginal value to future favors exceeds c. A lower rate of exchange could allow to expand the number of possible favors.

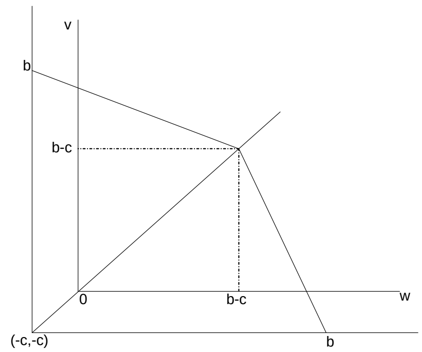

Figure 1: Feasible and Individually Rational Payoffs

The second special feature is that agents' continuation values do not change unless an agent does a favor. This is restrictive, and rules out the possibilities of appreciation (charging interest) or depreciation (forgiveness) of entitlements and punishment in case "not enough" cooperation is observed.

In the following section, strategies are not limited to a particular scheme. Instead, we characterize the optimal Perfect Public equilibria.

4 Characterizing the Optimal Perfect Public Equilibrium

4.1 Definition

The game described above falls in the class of repeated games with imperfect monitoring. As usual in this literature, we restrict our analysis to Public Perfect Equilibria (PPE), where strategies are functions of the public history only and equilibrium is perfect Bayesian.

A public history up to time t, denoted by h^t, consists of agents' past favors including size and date. A strategy $x_{it} : h^t \to [0,1]$ for player i specifies for every history and time period the size of favor the agent grants if the opportunity to do so arises. A public perfect equilibrium is a pair of strategies $\{x_{1t}, x_{2t}\}_{t\geq 0}$ that constitute a perfect Bayesian equilibrium. To analyze this game, we consider a recursive representation following the formulation of Abreu, Pearce and Stacchetti (APS) (1990). Let

$$\mathbf{V}^{*}= \{(v,w) \,|\exists \text{ PPE that achieves these values}\},$$

be the set of PPE values, where v and w denote player 1 and 2's values respectively. Given the linearity of payoffs and convexity of the strategy sets, it follows that V^* is convex. The set V^* is a subset of the set of feasible and individually rational payoffs, which is given in Figure 1.

The set of feasible payoffs is given by the outer x-y axis. Values have been multiplied by r/α so they are expressed as flow-equivalents. The extreme points are given by the vectors $(-c,b)$, $(b,-c)$, $(b-c,b-c)$, $(0,0)$. The first two correspond to the situation where only one player is giving favors and doing so whenever possible; the third vector corresponds to full cooperation by both players and the last one to no cooperation. The set is obtained by convex combinations of these points and is derived by combining the endpoints over time. Individually rational payoffs are found by restricting this set to the positive orthant. The set V^* is a proper subset of this set.

4.2 Factorization

To characterize the set V^* we follow the general idea of Abreu, Pearce and Stacchetti (APS), which decompose (factorize) equilibrium values into strategies for the current periods and continuation values for each possible end of period public signal. The difficulty in our case is that there is no *current period* in our model.

Let $\mathbf{W}$ denote an initial set of vectors of continuation values for the players (v, w). Adapting the approach in APS, equilibrium payoffs can be factorized in the following way. Let t denote the (random) time at which the next favor occurs. Consider any $T > 0$. Factorization is given by functions $x_1(t), x_2(t), v_1(t), v_2(t), w_1(t), w_2(t)$ and values v_T, w_T with the following interpretation. If the first favor occurs at time t and is given by player i, then $x_i(t) \in [0, 1]$ specifies the size of the favor, $v_i(t)$ the continuation value for player one and $w_i(t)$ the continuation value for player two, where for each t the vector $(v_i(t), w_i(t)) \in \mathbf{W}$. If no favor occurs until time T, the respective continuation values are v_T, w_T. The strategies and continuation values give the following utility to player one:

$$v = \int_0^T e^{-rt} \left\{ \frac{x_2(t) b + v_2(t) - x_1(t) c + v_1(t)}{2} \right\} p(t)\, dt + e^{-(r+2\alpha)t} v_T$$

where $p(t)$ denotes the density of the first arrival occurring at time t. This is the density of an exponential distribution with coefficient 2α (the total arrival rate). Letting $\beta = e^{-(r+2\alpha)}$ and $z_1(t) = x_2(t) b - x_1(t) c + v_1(t) + v_2(t)$, the above equation simplifies to:

$$v = \alpha \int_0^T \beta^t z_1(t)\, dt + \beta^T v_T. \tag{1}$$

Similarly, letting $z_2(t) = x_1(t) b - x_2(t) c + w_1(t) + w_2(t)$ one obtains the value w for player two:

$$w = \alpha \int_0^T \beta^t z_2(t)\, dt + \beta^T w_T. \tag{2}$$

The incentive compatibility condition requires that for all t,

$$v_1(t) - x_1(t) c \geq \alpha \int_t^T \beta^{s-t} z_1(s)\, ds + \beta^{T-t} v_T \tag{3}$$

$$w_2(t) - x_2(t) c \geq \alpha \int_t^T \beta^{s-t} z_2(s)\, ds + \beta^{T-t} w_T \tag{4}$$

The left hand side represents the net utility of giving a favor at time t and the right hand side the continuation utility if the agent passes this opportunity.

Starting with a set $\mathbf{W} \subset \Re_+^2$ of values, this factorization gives a new set of values $B_T(\mathbf{W})$ given by all pairs (v, w) such that there exist functions $x_1(t), x_2(t), v_1(t), v_2(t), w_1(t), w_2(t)$, where $(v_i(t), w_i(t)) \in \mathbf{W}$ and (1), (2), (3), (4) are satisfied. Following APS, a set of values $\mathbf{W}$ is *self generating* if $\mathbf{W} \subset B_T(\mathbf{W})$. If $(v, w) \in \mathbf{W}$, then there exists a PPE that gives the players initial payoffs (v, w). The set of PPE V^* is the largest set $\mathbf{V}$ such that $\mathbf{V} = B_T(\mathbf{V})$.The corresponding Pareto frontier of values can be characterized by the following program:

$$\begin{aligned} W(V) &= \max \alpha \int_0^T \beta^t \{x_1(t) b - x_2(t) c + w_1(t) + w_2(t)\} dt + \beta^T w_T \\ &\quad \text{subject to (1), (3), (4) , } (v_i(t), w_i(t)) \in \mathbf{V} \text{ and } (v_T, w_T) \in \mathbf{V} \end{aligned}$$

In general, payoffs in the Pareto frontier may require *inefficient equilibria* (i.e. equilibria with dominated payoffs) after some histories. One might expect, for example, that if a certain amount of time elapses and no favors are observed, cheating would be suspected and both payers would be punished with continuation values that are below the Pareto frontier. The following proposition shows that this is not needed in our repeated game.

Proposition 1 *1) The Pareto set of values* $\{(v, w) \in \mathbf{V}^* \text{ such that } w = W(v)\}$ *is self-generating. 2) The Pareto frontier is concave and its slope lies in the interval* $[-b/c, -c/b]$ *. 3) The domain is given by an interval* $[0, v_h]$ *where* $v_h \leq \bar{v}$.

Proof. In appendix A. ■

Self-generation implies that any point in the Pareto frontier can be obtained by relying on continuation values that are also in the frontier. This implies that PPE supporting the Pareto frontier are renegotiation proof. This result is not generally true in games of private information where inefficient punishments may be necessary in order to achieve the highest level of cooperation. Take a similar model in discrete time, and assume that the arrival of a favor for each player in any given period is α and the probability of no arrival is $(1 - 2\alpha)$. Then we can write players' values as

$$\begin{aligned} v &= \alpha(b - c) + \beta(\alpha v_1 + \alpha v_2 + (1 - 2\alpha) v_0) \\ w &= \alpha(b - c) + \beta(\alpha w_1 + \alpha w_2 + (1 - 2\alpha) w_0) \end{aligned}$$

where β is the discount rate. Depending on the probability α, it may be optimal to use interior continuation values in case no favor is observed since that would minimize the cost of giving incentives to players for doing favors. Figure 2 is an illustration of such a case. The large blue triangle connects the points $(v, w), (v_1, w_1), (v_2, w_2)$ which are all on the Pareto frontier, and where (v, w) are the continuation values if no favors are observed and v_1 and w_2 are such that $v_1 = v_0 + c$ and $w_2 = w_0 + c$. Here, we are assuming that the incentive constraint binds for the agent doing a favor, a claim which will be proved formally in the next section. Then $v_{t+1} = \alpha(v_1 + v_2) + (1 - 2\alpha) v_0$ and $w_{t+1} = \alpha(w_1 + w_2) + (1 - 2\alpha) w_0$ are the players' expected values (minus $\alpha(b - c)$) from such a mechanism. If α is large enough, (v_{t+1}, w_{t+1}) will be "far enough" from the Pareto frontier. Take instead the small red triangle where the continuation values after the event of no favor is interior to the Pareto

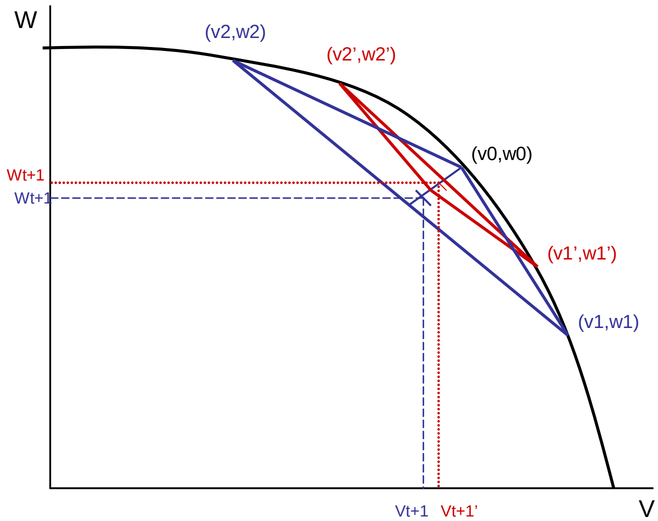

Figure 2: Discrete Time Punishments

frontier. Again, v_1' and w_2' are chosen such that the incentive compatibility constraints are exactly satisfied. One can easily see that the second scheme dominates the first one. So why doesn't this reasoning apply to the continuous time equivalent? Why doesn't an inward drift help with players' incentives there too? Intuitively, when the time period shrinks (basically to zero in continuous time), the information contained in not observing a signal becomes too imprecise to make the punishment contingent upon it. Since the probability of not observing a favor is not infinitely more probable when a player has cheated than when not, the loss from inefficient punishments becomes geater than the gain from improved incentives.

In order to understand better the tradeoff, take the discrete time version again and start with the symmetric point $(v^*, W(v^*))$, where each player's expected value is

$$\begin{aligned} & \alpha\,(b-c) + \beta\alpha\,(v^* + c/\beta + \alpha v_2) + \beta(1-2\alpha)v^* \\ = \;& \alpha b + \beta\,(\alpha v_2 + (1-\alpha)v^*) \end{aligned}$$

Now imagine an equal inward "punishment" in case no favor is observed, to some value pair $(v^* - \delta, W(v^*) - \delta)$. The incentive constraints binding, the continuation value of a player in case he does a favor will decrease by the same amount δ. Consequently, v_2 will increase by some corresponding ε which depends on the magnitude of δ and the curvature of the frontier. The new expected value is

$$\alpha b + \beta\,(\alpha v_2 + (1-\alpha)v^*) - (1-\alpha)\delta + \alpha\varepsilon$$

The difference in the expected values is $\alpha\varepsilon + (\alpha - 1)\,\delta$. Clearly, the larger is α, the greater the potential gains from the punishment. For a small enough α, there will be no gain from initiating such an inward movement and players are better off staying at $(v^*, W(v^*))$ in case no favor is observed.[5]

This is reminiscent of the result in Abreu, Milgrom and Pearce (1991), who point out that the timing of information is critical in designing effective punishments and achieving cooperation in moral hazard problems with continuous time. In a recent paper, Sannikov and Skrzypacz (2007) look at the use of information for providing incentives in games with frequent actions (basically as the period between players' actions shrinks to 0). They find that the information type (continuous or jump process or a mix of the two) is crucial in establishing which signals to exploit at all, and if so, for determining the effective use of incentives: for example, whether to use payoff transfers or money burning. Their analysis is contingent on keeping the information flows and players' payoffs independent from the change in the frequency of actions, whereas in our case, the arrival of information is actually *contingent* on players' taking action.

As in APS, an algorithm of successive approximations can be defined by iterating on the operator B_T, starting from a set containing V^* (such as the set of feasible and individually rational payoffs defined above). This procedure converges monotonically (by set inclusion) to V^*. The algorithm can be simplified in our case, restricting it to iterations on a value function defined by the frontier of values. It can be shown that starting from the frontier of the set of feasible and individually rational payoffs, convergence to the frontier of V^* is monotonic. There is a difficulty with this algorithm, since for each iteration the optimal strategies are the solution to an optimal control problem rather than a simple optimization problem, as in APS. Appendix B provides an alternative simplified algorithm, where incentive constraints are relaxed, that gives monotone convergence (from above) to the Pareto frontier by relying on an elementary optimization problem.[6]

The following section develops an alternative procedure to characterize the Pareto frontier that relies on the differentiable structure of the game. [7]

5 A Differential Approach

5.1 Recursive Formulation

We follow a heuristic approach. For small T the equation

$$W\left(v\left(0\right)\right) = \alpha \int_0^T \beta^t \left\{x_1\left(t\right)b - x_2\left(t\right)c + W\left(v_1\left(t\right)\right) + W\left(v_2\left(t\right)\right)\right\} dt + \beta^T W(v\left(T\right))$$

can be approximated by:

$$W\left(v\left(0\right)\right) - W\left(v\left(T\right)\right) = \alpha T \left\{x_1 b - x_2 c + W\left(v_1\right) + W\left(v_2\right)\right\} + \left(\beta^T - 1\right) W\left(v\left(T\right)\right)$$

[5] Evidently, once players' values shift away from the symmetric point, inward moves which favor more the player with higher value become more desirable if looking at a one-shot Pareto improving move. This is not necessarily the case when considering simultaneously the symmetric change when the values are reversed.

[6] In appendix B we show that the largest self-generating set V of the operator B_T is independent of T.

[7] A related procedure was developed by Sannikov (2007) for a continuous time game with stochastic diffusion processes.

Dividing by T, taking limits as $T \to 0$ and letting $v(0) = v$,

$$W'(v)\dot{v} = -\alpha\{x_1 b - x_2 c + W(v_1) + W(v_2)\} + (r + 2\alpha) W(v)$$

Following a similar procedure in equation (1) gives:

$$\dot{v} = -\alpha\{x_2 b - x_1 c + v_1 + v_2\} + (r + 2\alpha) v$$

Finally, the incentive constraints read:

$$\begin{aligned} v_1 - x_1 c &\geq v \\ W(v_2) - x_2 c &\geq W(v). \end{aligned}$$

In this continuous time problem, the choice variables are x_1, x_2, v_1, v_2 and $\dot{v}$, where the first four variables are the analogue of the controls $x_1(t)$, $x_2(t)$, $v_1(t)$, $v_2(t)$ in the previous problem, and $\dot{v}$ is the analogue of choosing $v(T)$. The optimization problem can be rewritten as:

$$\begin{aligned} rW(v) &= \max_{x_1, x_2, v_1, v_2, \dot{v}} \alpha(x_1 b - x_2 c) && (5) \\ &\quad +\alpha(W(v_1) - W(v) + W(v_2) - W(v)) + W'(v)\dot{v} && (6) \\ \text{subject to} &: \\ rv &= \alpha(x_2 b - x_1 c + v_1 - v + v_2 - v) + \dot{v} && (7) \\ v &\leq v_1 - x_1 c && (8) \\ W(v) &\leq W(v_2) - x_2 c. && (9) \end{aligned}$$

The following properties can be established as a result of the concavity of the function W.

Proposition 2 *The solution to the optimization problem defined by (5)-(9) has the following properties:*

1. *Both incentive constraints bind.*
2. $x_1 = \min((v_h - v)/c)$; $x_2 = \min((W(0) - W(v))/c)$.

Proof. In appendix A. ■

As in the SCM, favors are done while the players' values are away from the boundary. Full favors are granted unless there is not enough utility in the set to compensate the provider for for the cost of the favor. In that case, the size of the favor is limited by the distance to the boundary (divided by the unit cost). In contrast to the SCM where the incentive constraint sets the limit to the exchange of favors, here the lower bound on player values is actually zero, so the individual rationality constraint binds. Since the contributor of a favor is always justly compensated (his value icreases by excatly c), he would be willing otherwise to provide favors unlimitedly.

Also, the relative price of favors depends on v, i.e. on the entitlements of the players. The value gain a player receives when he does a favor represents a promise, from his partner, of compensating him with a certain number of favors in the future. Since the expected benefits from these claims are discounted, an agent is expected to repay each favor with an increasing number of favors, the closer his value gets to 0. The rate of exchange is approximately $|W'(v)| \in [c/b, b/c]$ which is true because of the binding incentive constraint.[8] In the simulations reported below, the extremes of the set $[c/b, b/c]$ are attained (at least approximately), giving rise to a range of relative prices of the order of $(b/c)^2$. Both of these features, i.e. the larger domain of values and the variable relative prices of favors, can potentially accommodate a considerably larger number of favors before reaching the boundaries.

Our game can be reinterpreted as a bargaining game where the protocol is for the buyer (or receiver of the favor) to make a take it or leave it offer to the seller (or grantor of the favor) and reaps all the gains from trade. In the SCM, since the incentive constraint of the seller only binds at the last favor, both agents share the surplus generated by all transactions before the last one. Could we imagine a scenario where the seller always moves first by offering the buyer her reservation value? Since all payments are made in terms of benefits from future favors, this gain can never be delivered if the seller today will also face a no surplus situation in the future, when he's in the buyer situation. Hence, the incentive constraint of the receiver could not bind all the time.

The following Proposition gives properties of the optimal $\dot{v}$.

Proposition 3 *In the optimal policy,*

1. $\dot{v}(0) = \dot{v}(v^*) = \dot{v}(v_h) = 0$, *where* v^* *is the unique point satisfying* $v^* = W(v^*)$;
2. *In the optimal mechanism,* $\dot{v}(v)$ *cannot be equal to* 0 *everywhere.*
3. *In the first best,* $\frac{\dot{v}}{v} = r > 0$ *for all values* v *such that* $W(v_1) \leq v^*$ *and* $\frac{\dot{v}}{W(v)} = -r\frac{c}{b} < 0$ *for all values* v *such that* $W(v_2) \geq v^*$.

Proof. In appendix A. ■

The second part of this proposition shows that, by virtue of the incentive constraint always binding, there needs to be a slack in the recursive equation for v at any point in the frontier. In general, taking any interval between two values of player 1 where he receives some number of favors, there will be at least one value v for which $\dot{v}(v)$ is different from 0.

The third part of the proposition shows that in the first best, $\dot{v}/v$ is positive constant and equal to the rate of interest along an sub-interval of values v in $[0, v^*]$, and is negative constant along the symmetric portion of the frontier. In the constrained problem, as $r \longrightarrow 0$, the Pareto frontier of PPE payoffs converges to the first best frontier, and the following corollary ensues.

[8] A favor given today by agent 1 in return for an increase of $v_1 - v = c$ in his value, will cost his partner exactly $W(v_1) - W(v)$ or roughly $W'(v)\,c$ in terms of future favors.

Corollary 1 *As* $r \longrightarrow 0$, $\frac{\dot{v}}{v} \longrightarrow r > 0$ *for values in the neighborhood of 0 and* $\frac{\dot{v}}{W(v)} \longrightarrow -r\frac{c}{b} < 0$ *for values in the neighborhood of* $\bar{v}$.

Conjecture 1 $\dot{v}'(0) > 0$, $\dot{v}'(v_h) < 0$.

Proof. Sketch of proof in appendix A. ■

We don't offer a formal proof of our conjecture, but argue informally why it is a credible and probable claim. Our simulations reported below suggest that this property indeed holds in the whole region (except the extremes), so that $\dot{v}$ drifts towards the *equal treatment* point v^*: agents' values are constantly changing over time even when no favors are observed, and the player with a lower entitlement is gradually rewarded. This is what we call forgiveness. In section 7 below, we give an interpretation of $\dot{v}$ in the context of some insurance models where an inflationary monetary policy is optimal.

5.2 Expected Favors Accounting

In the equilibrium described above, the player's values provide an accounting device of past history and current entitlements. An alternative equivalent accounting is given below. Let T_i represent the expected discounted number of favors that player i will give in the rest of the game. It obviously follows that:

$$\begin{aligned} v &= T_2 b - T_1 c \\ w &= T_1 b - T_2 c \end{aligned}$$

which solving, gives:

$$\begin{aligned} T_1 &= \frac{vc + wb}{b^2 - c^2} \\ T_2 &= \frac{wc + vc}{b^2 - c^2} \end{aligned}$$

and

$$T_1 - T_2 = \frac{v - w}{b + c}.$$

This difference can be understood as a net balance between the player's assets and liabilities. When $T_1 < T_2$, the player is in a net debt position. As we find, $\dot{v} > 0$ in this region, which has the interpretation of debt forgiveness.

6 Numerical Results

The optimal strategies differ in several dimensions from the very simple strategies proposed by the SCM of Mobius. How important is this? This section provides some numerical computations to examine this question.

There are 4 parameters in the model: r, α, c, b. In comparing the performance of different alternatives, two normalizations can be made where all that matters in these comparisons are the values of c/b and r/α. In the next tables, the following normalizations are used: $b = 1$ and $r = 0.01$.

The following tables give a measure of how far each alternative scheme is from the first best at the symmetric point of the boundary where players get equal utilities. The first column gives the percentage difference between the first best values and values for the optimal scheme described above; the second column gives the percentage difference with values for the optimal scheme with the added restriction that $\dot{v} = 0$; the third column gives the difference with SCM values.

%difference with optimum			
α		0.8	
c	$\dot{v} \neq 0$	$\dot{v} = 0$	SCM
0.4	2.7	3	6.4
0.5	3.1	3.3	7.6
0.65	3.4	3.5	10.5
0.75	3.5	3.6	13.5
0.95	10.3	10.3	32.3

%difference with optimum			
c		0.5	
α	$\dot{v} \neq 0$	$\dot{v} = 0$	SCM
0.2	6	6.5	15.5
0.3	4.9	5.3	12.9
0.4	4.3	4.6	11
0.5	3.8	4.1	9.8
0.6	3.5	3.7	9.1

The performance of all these schemes decreases with c and increases with the arrival rate α. There can be substantial improvements over the SCM: e.g. for $\alpha = 0.8$ and $c = 0.95$, the second best is 10% within the first best, while the SCM is 30% apart. It is also interesting to observe that restricting $\dot{v} = 0$ does not have a substantial impact on performance.

The following two tables give the maximum number of consecutive favors starting from the midpoint and leading to the boundary. The optimal scheme can accommodate a much larger number of favors (between 5 to over 10 times more). In part this is due to flexible relative prices, which depart significantly from one. A natural question is, when do fixed prices become a good enough substitute for flexible prices? Our computations show that, as expected, when trade opportunities are very frequent (or equivalently, when dicounting is very low), fixed prices approach in their performance flexible prices.

Number of favors			Average price left of midpoint
α	0.8		
c	$\dot{v} \neq 0$	SCM	
0.4	167.1	13.2	2.32
0.5	119	11.2	1.87
0.65	70.1	8.1	1.46
0.75	45.7	6.2	1.29
0.95	6.5	2.1	1.39

Number of favors			Average price left of midpoint
c	0.5		
α	$\dot{v} \neq 0$	SCM	
0.2	29	5.1	1.87
0.3	44	6.2	1.87
0.5	74	8.5	1.87
0.6	89	9.2	1.87

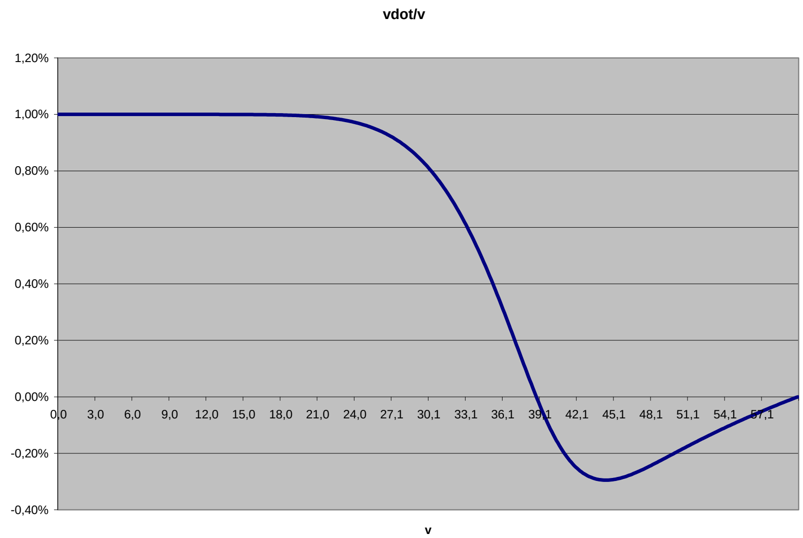

Figure 3: Forgiveness

In all the simulations, $\dot{v}$ is positive for values of v under the symmetric point v^* so the equilibrium displays forgiveness. Figure (3) illustrates this for the benchmark case. It is interesting to note that $\dot{v}/v$ is monotonically decreasing and that it equals the interest rate $r = 1\%$ in the lower section of its domain. Using (7) and the incentive constraint for player one it follows that:

$$r - \frac{\alpha}{v}\left(x_2 b + v_2 - v\right) = \frac{\dot{v}}{v}$$

so that $v_2 - v \approx x_2 b$. Using the incentive constraint for player two this implies that in this range of values $W'(v)$ must be approximately equal to $-c/b$ so that player one is almost indifferent between receiving an extra favor or not.

Figure (4) provides the decomposition of values in terms of *favor entitlements* indicated above. Note that for lower values of v most of the increase in value for player one is the result of an increased entitlement to favors of player two, with basically no change in the favors owed by player one.

7 Implementation with Chips and Relation to Money

This section examines implementation of our equilibrium with a *chips mechanism.* We first define such a mechanism and provide a mapping from values to *chips.* We then discuss the connection between forgiveness and *inflation* generated by the injection of chips.

Definition 1 *A chips mechanism induced by the equilibrium of the game is a strictly increasing mapping $c(v)$ from agent one's value to the interval $[0, 1]$ with the symmetry property that: $c(v) = 1 - c(W(v))$.*

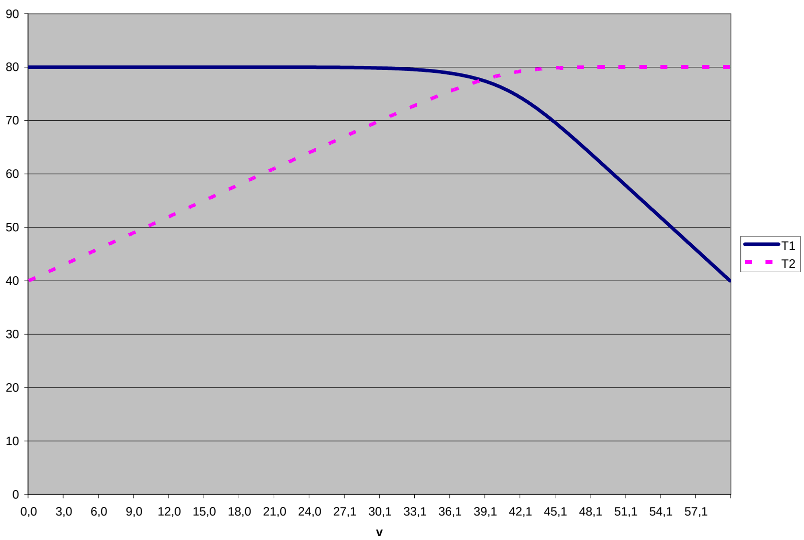

Figure 4: Favors Owed

The value $c(v)$ is interpreted as the share of chips held by agent one. The first property says it is an accounting device sufficient for determining the values of the players. The second property implies that the share of chips fully determines the utility of a player independently of its identity. These are two conditions satisfied by Mobius' scheme.

Let $c(v) = v/(v + W(v))$. It is easy to verify that both properties of the definition are satisfied. In our equilibrium prices depend on the shares of chips as follows:

1. Player one does a favor: $p_1(c) = c(v_1) - c(v)$ where $c = c(v)$.

2. Player two does a favor: $p_2(c) = c(v) - c(v_2)$ where $c = c(v)$.

The share of chips also changes independently of favors over time: $\dot{c}(c) = c'(v)\dot{v}(v)$ where $c = c(v)$. Subject to this *forgiveness rule*, the equilibrium can be implemented by a chips mechanism where the receiver of a favor has all the bargaining power to determine the price (in terms of chips) of a transaction. This follows immediately from the fact that in our equilibrium the incentive compatibility constraint binds for the agent doing the favor.

Forgiveness suggests negative real interest rates. These can be obtained by a specific injection of chips. Let $m(t)$ denotes the total number of chips at time t and $m_i(t)$ the number held by player i. Then $m_1(t)/m(t) = c(t)$. Let $\dot{m}(t)$ denote the injection of chips at time t and suppose that $\dot{m}_1(t) = \dot{m}_2(t) = \dot{m}(t)/2$ so that both players receive the same

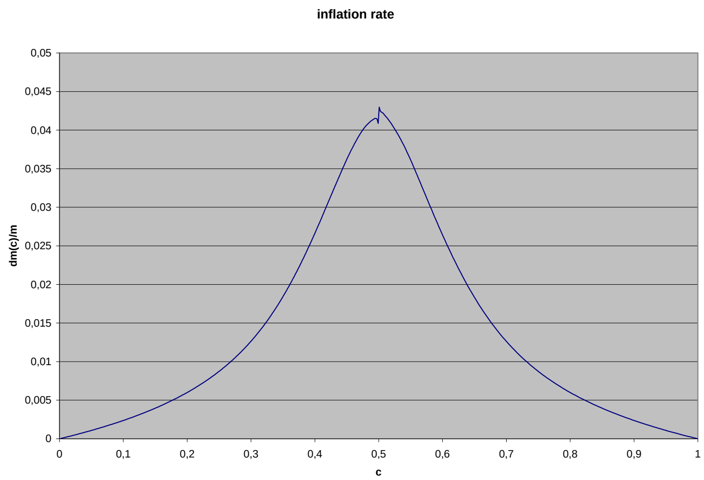

Figure 5: Rate of Expansion of Chips

number of chips. All prices grow at the same rate as the stock of chips. Note that

$$\begin{aligned}\dot{c}(t) &= \frac{d}{dt}\frac{m_1(t)}{m(t)} = \frac{m(t)\,\dot{m}(t)/2 + m_1(t)\,\dot{m}(t)}{m(t)^2} \\ &= \frac{\dot{m}(t)}{m(t)}\left(\frac{1}{2} - c(t)\right)\end{aligned}$$

so the rate of expansion of chips

$$\frac{\dot{m}(t)}{m(t)} = \frac{\dot{c}(t)}{\frac{1}{2} - c(t)},$$

which is positive whenever the sign of $\dot{c}(t)$ equals the sign of $\frac{1}{2} - c(t)$, as occurs in our computations. Figure 4 the implied rates of chips expansion (i.e. inflation) implied by the equilibrium levels of forgiveness. The peak occurs at the symmetric point reaching a value of 4% (four times larger than the discount rate $r = 1\%$.). Since the rate of inflation is equivalent to the negative interest rate implied, an alternative implementation is to have approximately 4% constant rate of expansion of chips and charge an interest to fill the gaps which according to the figure would increase with the size of the debt $\left(\frac{1}{2} - c\right)$, from zero to approximately 4% as it approaches its maximum at the extremes of the interval.

Relation of $\dot{v}$ to monetary inflation:

A number of papers have addressed the potential optimality of expansionary monetray policy in economies with uncertainty. In a nutshell, inflation can act as an insurance device which tends to rebalance agents' wealth in favor of the least "lucky", in order to avoid hitting extreme distributions where some agents are too poor or too wealthy, which could lead to inefficiencies. The two most relevant papers for us are by Green and Zhou (2005) and Levine (1991).

Green and Zhou (2005) study an economy where agents' preferences and endowments are private information and random. The trading mechanism is one where every period, each individual makes a donation to a common fund, and the proceeds are redistributed over the agents by a social planner. An agent's wealth depends directly on her past donations and received transfers and can be interpreted as a future claim on the common fund. In the case where endowments are constant and equal across agents, utility functions are linear and marginal utility of consumption varies between two levels, the authors show that an expansionary monetary mechanism Pareto dominates the mechanisms with constant or contractionary monetary policy. Moreover, under some parameter restrictions, the first best can be achieved in equilibrium through a stationary inflationary mechanism.

The reason why inflation is part of the optimal mechanism is that inefficiencies occur when low marginal utility agents consume their endowment instead of contributing to the common fund. By constantly inflating money balances, agents' wealths remain bounded from above, so they can never become wealthy enough to fully insure themselves against future fluctuations in their preferences. The two condition for ensuring that the first best be achievable are that utility be linear and that the total endowment in the economy can cover only part of the high types' consumption needs, in a way that consumption inequality among them is both unavoidable and not detrimental to efficiency. In our two-agent model, instead, anything below a full favor is inefficient and probable.

Levine (1991) gives an example with two types of infinitely lived agents that shift randomly between having high and low valuation of consumption. The state of the world is publicly observed, but not the agents' types. The author shows that if the difference in valuation is large enough, there will be an expansionary policy that dominates all contractionary policies.[9] The intuition is that a sufficiently long sequence of realizations where a group is the high type will result in that group running out of money, hence not being able to consume when their valuation is high. In that case, an expansionary policy will tend to insure unlucky buyers against that event. Here as well the first best is achievable, by conditioning agents' allocations on the observed state along with their asset holdings and messages.

The difference between the above two cases is subtle, but Levine's argument is analogous to the rationale for having $\dot{V}$ in our mechanism. In our model, one can think of agents valuing a favor differently whether they are givers or receivers, and the giver is always compensated with higher future claims on favors. Since the Individual Rationality constraint puts the bound on the exchange, there is no risk of an agent keeping a favor to herself, but rather, that her partner hits the zero value limit and has nothing to trade for a favor. $\dot{V}$, by constantly working against the agent with a higher value, has a similar function to an inflationary policy which partially insures agents against reaching that point. Still, in our model the informational frictions put a limit on the gains which can be achieved and the first best is in general not attained. As seen in the computations, this cost decreases significantly with the frequency of trading opportunities α.

[9] This difference in valuation is needed for agents to still have an incentive to delay their consumption in spite of the inflation.

8 Final remarks

Our model can be equally understood as a game with "market share favors" where c represents the opportunity cost of producing when a firm has a high cost, and b is the profit for a low cost firm. When a firm has a high production cost, it foregoes c in favor of the low cost firm which receives b in profits, and continues with a higher value. This is a reinterpretation of the scenario in Athey and Bagwell (2001), who study a model of collusion between two firms with private cost realization. Every period, a firm's cost can be high or low, so productive efficiency calls for only the low cost firm producing when the other firm has high cost. Firms are rewarded for truthfully revealing their costs through higher continuation values, which represent a higher market share in the future. When firms are sufficiently patient, these promises can be delivered merely during periods when both firms have the same cost, by allowing the firm with the high continuation value to produce more (or exclusively). In this case, the set of PPE values corresponds to the first best frontier, which is simply a line segment of slope -1, along with the threat values in case one firm deviates from the agreement.

There is one main difference between the above setting and our case: only one agent at a time has the opportunity to do a favor, and the only event where agents are similar is when none of them can do a favor (one can think about a state where marginal cost is greater than the buyer's reservation value), which makes it impossible to use these symmetric states as a basis for rewarding past behavior. Consequently, the higher continuation values can only be delivered through asymmetric states, by allowing a high value firm to take over the market even when it has a high cost: this is the case of a player reaching v_h and not doing any favors. So similarly to Athey and Bagwell (2001), we should observe future market shares to be negatively correlated with present market shares. On the other hand, symmetric states serve a totally different purpose than in their case, by indeed rewarding the firm with the lower value (through the drift $\dot{v}$) and keeping it from reaching the inefficiency point $\underline{v}$.

Finally, this paper considers cooperation in the absence of immediate reciprocity. The lack of double coincidence of needs has been the subject of many papers in monetary theory (Townsend (1980), Kiyotaki and Wright (1989)). In most of that literature, the interaction between agents is either a one-time encounter (Townsend (1980)), or is defined by repeated meetings, between agents who fail to match in their production abilities and consumption desires. In Kiyotaki and Wright (1989) for example, there are three agents who each specialize in the production of one good and the consumption of a second good, but no two agent can mutually satisfy their consumption desires. When storage is excessively costly or impossible, money (or a chips mechanism) arises as a necessary medium of exchange in trade. A second problem with barter, aside from the lack of double coincidence of needs, is the rate of exchange of goods. In the presence of many players and goods, it becomes excessively complicated to quote each good in terms of every other good, instead of having one unit of accounting to which to refer.

Money is then essentially a record keeping device which can be seen as a "primitive form of memory" (Kocherlakota (1998)), and is especially practical when players are numerous and do not act repeatedly. In our model the repeated interaction of players makes exchange possible as an equilibrium outcome by relying simply on continuation values. The exclusive partnership between the two agents and the random incidence of trade allow relative prices to naturally depend on previous trades, a feature which disappears in large economies with numerous agents and goods and frequent, routine interaction. Although informational frictions put a limit on the gains which can be achieved, as seen in the computations, this cost decreases significantly with the frequency of trading opportunities α.

In many organizations, internal exchange is not mediated with money. Obvious examples are households and other partnerships such as co-authorship or co-workship. Our analysis suggests that the lack of use of money should be related to the frequency of trade opportunities. Casual observation suggests that in many of these organizations, the multiple dimensions of exchanges enhance the frequency of trading opportunities, thus reducing the value of mediating trade with monetary payments.

References

[1] Abdulkadiroglu, A and K Bagwell (2004), "Trust, reciprocity and favors in cooperative relationships, " *Working Paper*.

[2] Abreu, D., P. Milgrom and D. Pearce (1991), "Information and Timing in Repeated Partnerships," *Econometrica*, 59, 1713-1733.

[3] Abreu, D., D. Pearce and E. Stachetti (1986), "Optimal Cartel Equilibria with Imperfect Monitoring," *Journal of Economic Theory*, 39, 251-269.

[4] Abreu, D., D. Pearce and E. Stachetti (1990), "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica*, 58, 1041-1063.

[5] Athey, S. and K. Bagwell (2001), "Optimal Collusion With Private Information," *The RAND Journal of Economics*, 32:3, 428-465.

[6] Atkeson, A. and R.E. Lucas (1992), "On Efficient Distribution with Private Information," *Review of Economic Studies*, 59, 427-453.

[7] Bewely, T. (1980), "The Optimum Quantity of Money," in J. Karaken and N. Wallace, eds., *Models of Monetary Economics*, Federal Reserve Bank of Minneapolis

[8] Bewely, T., (1983), "A Difficulty with the Optimum Quantity of Money," *Econometrica*, 51, 1485-1504.

[9] Fudenberg, D., Levine, D. and E. Maskin (1994), "The Folk Theorem with Imperfect Public Information," *Econometrica*, 62:5, 997-1039.

[10] Fudenberg, D. and K. Levine (2007), "Continuous Time Models of Repeated Games with Imperfect Public Monitoring," *Review of Economic Dynamics*, 10:2, 173-192.

[11] Green, E. and R. Zhou (2005), "Money as a Mechanism in a Bewely Economy," *The International Economic Review*, 46:2, 351-371.

[12] Hopenhayn, H., and A. Skrzypacz (2004), "Tacit Collusion in Repeated Auctions," *Journal of Economic Theory,* 114:1, 153-169.

[13] Kehoe, T., D. Levine and M. Woodford (1992), "The Optimum Quantity of Money Revisited," in P. Dasgupta, D. Gale, O. Hart, and E. Maskin, eds., *The Economic Analysis of Markets and Games: Essays in Honor of Frank Hahn*, MIT Press, 501-26.

[14] Kalesnik, V. (2005), "Repeated partnership with jump process signal in continuous time", *Working Paper.*

[15] Kocherlakota, N. (1998), "Money Is Memory," *Journal of Economic Theory,* 81:2, 232-251.

[16] Kandori, M. (1992), "The Use of Information in Repeated Games with Imperfect Monitoring," *Review of Economic Studies*, 59:3, 581-593.

[17] Levine, D. (1991), "Asset Trading Mechanism and Expansionary Policy," *Journal of Economic Theory,* 54:1, 148-164.

[18] Mobius, M. (2001), "Trading favors ", *Working Paper.*

[19] Kiyotaki, N. and R. Wright, (1989), "On Money as a Medium of Exchange," *Journal of Political Economy*, 97:4, 927-54.

[20] Kiyotaki, N. and R. Wright, (1993), "A Search-Theoretic Approach to Monetary Economics," *American Economic Review*, 83:1, 63-77.

[21] Rocheteau, G. and R. Wright, (2005), "Money in Search Equilibrium, in Competitive Equilibrium, and in Competitive Search Equilibrium," *Econometrica*, 73:1, 175-202.

[22] Sannikov, Y. (2007), "Games with Imperfectly Observable Actions in Continuous Time", *Econometrica*, 75:5, 1285-1329.

[23] Sannikov, Y. (2008), "A Continuous-Time Version of the Principal-Agent Problem," forthcoming in *Review of Economic Studies*

[24] Sannikov, Y. and A. Skrzypacz (2007),"Impossibility of Collusion under Imperfect Monitoring with Flexible Production," forthcoming in *American Economic Review.*

[25] Sannikov, Y. and A. Skrzypacz (2007),"The role of information in repeated games with frequent actions," *Working Paper.*

[26] Townsend, R. (1980), "Models of Money with Spatially Separated Agents," in J. Karaken and N. Wallace, eds., *Models of Monetary Economics*, Federal Reserve Bank of Minneapolis.

Appendix A: Proofs of Propositions:

Proof of proposition 1:

1. The Pareto frontier is self generating:

Take a point (v, w) in the Pareto frontier of V^*. For any T, this can be factorized by strategies and continuation values $\{x_i(t), v_i(t), w_i(t), v_T, w_T\}$. Suppose (v_T, w_T) is not in the Pareto frontier, so there exists $\varepsilon > 0$ such that $(v_T + \varepsilon, w_T + \varepsilon) \in V^*$. Define new paths $\tilde{x}_i(t) = \max(0, x_i(t) - d(t))$, where $d(t) = \frac{\varepsilon}{c} e^{\alpha(T-t)} \beta^{T-t}$. We prove that these paths together with

$$\{v_i(t), w_i(t), v_T + \varepsilon, w_T + \varepsilon\}$$

are admissible with respect to V. First we show incentive compatibility. For $x_i(t) = 0$, the incentive constraint is trivially satisfied since agent i is not required to rpovide any favor. For $x_i(t) > 0$, letting

$$z_1(s) = x_2(s) b - x_1(s) c + v_1(s) + v_2(s),$$

$$\begin{aligned}
& \alpha \int_t^T \beta^{s-t} \{\tilde{x}_2(s) b - \tilde{x}_1(s) c + v_1(s) + v_2(s)\} ds + \beta^{T-t}(v_T + \varepsilon) \\
\leq \; & \alpha \int_t^T \beta^{s-t} \{x_2(s) b - x_1(s) c + v_1(s) + v_2(s) + d(s)c\} ds + \beta^{T-t}(v_T + \varepsilon) \\
= \; & \alpha \int_t^T \beta^{s-t} \{x_2(s) b - x_1(s) c + v_1(s) + v_2(s)\} ds + \beta^{T-t}(v_T + \varepsilon) + \alpha c \int_t^T \beta^{s-t} d(s) ds \\
= \; & \alpha \int_t^T \beta^{s-t} z_1(s) ds + \beta^{T-t}(v_T + \varepsilon) + \alpha \varepsilon \beta^{T-t} \left(\frac{-1 + e^{\alpha(T-t)}}{\alpha} \right) \\
= \; & \alpha \int_t^T \beta^{s-t} z_1(s) ds + \beta^{T-t} v_T + d(t) c \\
\leq \; & v_1(t) - x_1(t) c + d(t) c = v_1(t) - \tilde{x}_1(t) c
\end{aligned}$$

so the incentive constraint for player one is satisfied. A similar argument shows that the same holds for player two. Let $(\tilde{v}, \tilde{w})$ denote the values associated to the new admissible path. We know that $\tilde{x}_2(t) \geq x_2(t) - d(t)$ and $\tilde{x}_1(t) \leq x_1(t)$ Then,

$$\begin{aligned}
\tilde{v} & \geq \alpha \int_0^T \beta^t z(t) dt + \beta^T (v_T + \varepsilon) - \alpha b \int_0^T \beta^t d(t) dt \\
& = v + \beta^T \varepsilon - \beta^T \varepsilon \frac{b}{c} (e^{\alpha T} - 1) \\
& = v + \beta^T \varepsilon \left(1 - \frac{b}{c} (e^{\alpha T} - 1) \right)
\end{aligned}$$

For small T, the right hand side exceeds v. A similar argument shows that $\tilde{w} > w$,contradicting the hypothesis that the value pair (v, w) is in the Pareto frontier of V^*. This proves that (v_T, w_T) must belong to the Pareto frontier of V^*.

We now show that $(v_i(t), w_i(t))$ also belong to the frontier. Suppose towards a contradiction that this was not true. Let

$$d = \min\{v + w | (v, w) \text{ are in the Pareto frontier of } V^*\}.$$

From the concavity and symmetry of the boundary of V^*, $d = W(0) > 0$. It follows that $v_T + w_T \geq b$. Let T_i be the set of time periods such that $v_i(t), w_i(t)$ are not in the Pareto frontier of V^*. We construct an alternative admissible path that improves on the given one as follows. Let $\hat{v}_i(t)$ be the value such that the pair $(\hat{v}_i(t), w_i(t))$ is in the frontier. Similarly define $\hat{w}_i(t)$. Without loss of generality (by choice of T),

$$\int_0^T \beta^t (\hat{v}_i(t) - v_i(t))\,dt + \int_0^T \beta^t (\hat{w}_i(t) - w_i(t))\,dt < \beta^T d.$$

It is thus possible to construct paths $\{\tilde{v}_i(t), \tilde{w}_i(t)\}$ where for each t and i either $(\tilde{v}_i(t), \tilde{w}_i(t))$ equals $(v_i(t), \hat{w}_i(t))$ or it equals $(\hat{v}_i(t), w_i(t))$ such that

$$\begin{aligned} \varepsilon_1 &= \alpha \int_0^T \beta^t [\tilde{v}_1(t) - v_1(t) + \tilde{v}_2(t) - v_2(t)]\,dt < v_T \\ \text{and} \quad \varepsilon_2 &= \alpha \int_0^T \beta^t [\tilde{w}_1(t) - w_1(t) + \tilde{w}_2(t) - w_2(t)]\,dt < w_T \end{aligned}$$

Define $\tilde{v}_T = v_T - \varepsilon_1$ and $\tilde{w}_T = w_T - \varepsilon_2$. Note that $(\tilde{v}_T, \tilde{w}_T) \in int(V^*)$. We now show that the path $\{x_i(t), \tilde{v}_i(t), \tilde{w}_i(t), \tilde{v}_T, \tilde{w}_T\}$ is admissible. Let

$$\begin{aligned} \delta_1(t) &= \tilde{v}_1(t) - v_1(t) + \tilde{v}_2(t) - v_2(t) \\ \delta_2(t) &= \tilde{w}_1(t) - w_1(t) + \tilde{w}_2(t) - w_2(t) \\ \tilde{z}_1(t) &= x_2(t)\,b - x_1(t)\,c + \tilde{v}_1(t) + \tilde{v}_2(t) \end{aligned}$$

So we have:

$$\begin{aligned} & \alpha \int_t^T \beta^{s-t} \tilde{z}_1(s)\,ds + \beta^{T-t} \tilde{v}_T \\ =\; & \alpha \int_t^T \beta^{s-t} z_1(s)\,ds + \beta^{T-t} v_T + \alpha \int_t^T \beta^{s-t} \delta_1(s)\,ds - \beta^{T-t} \varepsilon_1 \\ =\; & \alpha \int_t^T \beta^{s-t} z_1(s)\,ds + \beta^{T-t} v_T + \beta^{-t} \left(\alpha \int_t^T \beta^s \delta_1(s)\,ds - \beta^T \varepsilon_1 \right) \\ \leq\; & \alpha \int_t^T \beta^{s-t} z_1(s)\,ds + \beta^{T-t} v_T \\ \leq\; & v_1(t) - x_1(t)\,c \\ \leq\; & \tilde{v}_1(t) - x_1(t)\,c \end{aligned}$$

This holds for any t, and in particular for $t = 0$, so player one's incentive constraint is satisfied at all times. A similar argument can be used to verify incentive compatibility for player two. We have constructed paths where $(\tilde{v}_i(t), \tilde{w}_i(t))$ are in the Pareto frontier for all t. To end the proof, note that since $(\tilde{v}_T, \tilde{w}_T)$ are not in the frontier, by our previous argument the path can be further improved by an alternative one that takes values in the Pareto frontier (without modifying $(\tilde{v}_i(t), \tilde{w}_i(t))$.

2. The boundary of V^* has slope in the set $[-b/c, -c/b]$.

We show that if the set V satisfies this property, so will the set $U = B_T(V)$ for all T. Let (v, w) be a point in the boundary of U where $v > 0$. If there is a set of positive Lebesgue measure in $[0, T]$ where $x_2(t) > 0$, then for any $0 < \varepsilon < \int \beta^t x_2(t)\,dt$ there is a new path $\hat{x}_2(t) \leq x_2(t)$ with $\int \beta^t \hat{x}_2(t)\,dt = \int \beta^t x_2(t)\,dt - \varepsilon$. This gives rise to the points $(v - \alpha\varepsilon b, w + \alpha\varepsilon c)$ in $B_T(V)$. On the contrary suppose that $x_2(t) = 0$ for all $t \in [0, T]$. Since $v > 0$, either $v_T > 0$ or there exists a set of positive Lebesgue measure in $[0, T]$ where $v_1(t)$ or $v_2(t)$ is strictly positive. Since $x_2(t) = 0$ for all t, the points $(v_T, w_T), (v_1(t), w_1(t))$ and $(v_2(t), w_2(t))$ must be in the Pareto frontier of V. Suppose $v_T > 0$. Then for any $0 < \varepsilon < v_T$ there exists values $(\hat{v}_T, \hat{w}_T)$ in the frontier of V where $\hat{v}_T = v_T - \varepsilon$ and $\hat{w}_T \geq w_T + \varepsilon \cdot c/b$. These terminal values together with the original path are self-generating (recall that $x_2(t) = 0$ for $t \in [0, T]$) and give rise to a point $(\hat{v}, \hat{w}) \in B_T(V)$ where $(\hat{w} - w) / (\hat{v} - v) \leq -c/b$. A similar argument can be used in case there exists a set of positive Lebesgue measure in $[0, T]$ where $v_1(t)$ or $v_2(t)$ is strictly positive. This proves that the boundary of $B_T(V)$ has slope less than or equal to $-c/b$. A symmetric argument shows that the slope is greater than or equal to $-b/c$.

3. The domain is given by an interval $[0, v_h]$ where $v_h \leq \bar{v}$.

Let $\underline{v}$ be the lowest value player 1 can reach in the Pareto set and $w(\underline{v})$ be the corresponding value for player 2 in the Pareto frontier. Suppose that $\underline{v} > 0$. Since the slope of the frontier at that point is greater than or equal to $-c/b$, we can make an incentive compatible and Pareto improving change by extending the Pareto frontier along that slope to fit a fractional favor.

Since $\underline{v}$ be the lowest value player 1 can reach, then (by incentive compatibility of player 2) at $\underline{v}$, then there must be an interval $[0, T]$ such that $x_2(t) = 0$ for $t \in [0, T]$. Then for any $0 < \varepsilon < \underline{v}$ there is a new path $\hat{x}_2(t)$ such that $\int \beta^t \hat{x}_2(t)\,dt = \varepsilon$, and let $\int \beta^t \hat{w}_2(t) = \int \beta^t w_2(t) + \varepsilon c$. Then the incentive constraint of player 2 is satisfied, and his value $w(\underline{v})$ is unchanged. On the other hand, player 1's new value is

$$\begin{aligned} \underline{\hat{v}} &= \alpha \int_0^T \beta^t \{\tilde{x}_2(t) b - x_1(t) c + v_1(t) + \hat{v}_2(t)\}\,dt + \beta^T v_T \\ &= \underline{v} + \alpha\varepsilon b + \alpha \int_0^T \beta^t \{\hat{v}_2(t) - v_2(t)\}\,dt \end{aligned}$$

For T small enough, the value of the integral is $\geq -\varepsilon b$, meaning that either $(\underline{v}, w(\underline{v}))$ could not have been in the Pareto frontier (in the case of a strict inequality), or that $(\underline{v}, w(\underline{v}))$ is in the Pareto frontier, but the domain can be further extended to include $\underline{\hat{v}}$ and eventually 0.

Proof of Proposition 2:

1. Both incentive constraints bind:

Suppose that player 1's incentive constraint (8) was slack, so

$$\begin{aligned} v_1 - x_1 c &= v + \delta \\ &= \frac{\alpha}{r+\alpha}(x_2 b + v_2) + \frac{1}{r+\alpha}\dot{v} + \frac{r+2\alpha}{r+\alpha}\delta \end{aligned}$$

Consider the following change: decrease v_1 by an $\varepsilon > 0$ and increase $\dot{v}$ by $\alpha\varepsilon$, so player 1's value is unchanged (and his promise keeping constraint is satisfied). Then as long as $\varepsilon < \frac{(r+2\alpha)}{\alpha}\delta$, the incentive constraint of player 1 will be satisfied. The resulting change in $W(v)$ is: $\Delta(r+2\alpha)W(v) = \alpha(W(v_1 - \varepsilon) - W(v_1)) + W'(v)\alpha\varepsilon$. For ε small, we can then write

$$\text{sign } \Delta W(v) = \text{sign } \varepsilon(-W'(v_1) + W'(v)) > 0$$

Hence, the new implementation is Pareto improving, which means that the previous one was not optimal. A similar argument shows that player 2's incentive constraint must also bind for values on the Pareto frontier.

2. $x_1 = \min\left(1, (v_h - v)/c\right); x_2 = \min\left(1, (W(0) - W(v))/c\right)$

This part of the proposition says that it is optimal for players to always do the largest favor that would "fit" within the boundaries of exchange. Suppose that $\min\left(1, (v_h - v)/c\right) = 1$ but $x_1 < 1$. Then increase x_1 by a small ε and v_1 by εc, which would leave player 1's incentive constraint and the promise keeping constraint intact. On the other hand, it is straightforward to obtain

$$\text{sign } \Delta W(v) = \text{sign } b + W'(v_1)$$

which is positive, since $W'(v_1) \in [-b/c, -c/b]$. If $(v_h - v)/c < 1$, it is clear that we cannot give enough incentive to player 1 for a full favor, and instead x_1 will be equal $(v_h - v)/c$ (from part 1 of the proposition, $v_h = v + x_1 c$).

Proof of Proposition 3:

1. $\dot{v}(0) = \dot{v}(v^*) = \dot{v}(v_h) = 0$, where v^* is the unique point satisfying $v^* = W(v^*)$

First, $\dot{v}(v^*) = 0$, by symmetry.The fact that $\dot{v}(0) = 0$ follows from the incentive constraint binding and the fact that at value 0, the player expects to receive no favor. $\dot{v}(v_h) = 0$ is again obtained by symmetry.

2. In the optimal mechanism, $\dot{v}(v)$ cannot be equal to 0 everywhere.

To see this, notice that any value v stands for two meanings: as the value accumulated from doing a number of favors in the past, and as the entitlement to future favors, and hence can be calculated recursively along both margins. Holding the incentive constraints strict, the two calculations can only be equal if we allow for the drift $\dot{v}(v)$ to be different from 0.

Suppose wlog that the number of favors that can fit in the interval $[0, \bar{v}]$ is an integer m. Then by the incentive constraints binding, we can write $\bar{v} = mc$. Similarly, $\bar{v}$ can be calculated as the expected utility from unilaterally receiving favors before reaching 0, which by symmetry, is the expected discounted sum of utilities from m consecutive favors.

Let v_i be the value of agent 1 when he is entitled to i consecutive favors, starting from $\bar{v}$. Suppose that $\dot{v}(v) = 0$ for all v, then we can write recursively $rv_i = \alpha(v_{i-1} - v_i + b)$, with $v_0 = 0$. Letting $\beta = \alpha/(r+\alpha)$,

$$\bar{v} = \beta \left(\frac{1-\beta^m}{1-\beta} \right) b = mc$$

which defines m as a function of β, b and c. Once player 1 receives a favor (and his stock of favors goes down to $m-1$), his value can be written as

$$v_{m-1} = \beta \left(\frac{1-\beta^{m-1}}{1-\beta} \right) b = (m-\epsilon)\, c$$

for $\epsilon < 1$, where the last equality is due to the rate of exchange of favors being less than one for the player with the higher value. By taking the difference $\bar{v} - v_{m-1}$, we have a condition on the parameters: $\epsilon = \beta^m b/c$. On the other hand, because of symmetry, we know that $W(v_{m-1}) = c = \beta \epsilon b$, so $\epsilon = c/\beta b$, and $c^2 = \beta^{m+1} b^2$. This supposes a particular set β, b and c which in general will not be satisfied. In order to allow for a slack in the three equations, $\dot{v}(v)$ will have to be different from 0 for some v on the interval $[v_{m-1}, \bar{v}]$.

If one attempts the same exercise for the second favor received by player 1, assuming $\dot{v}(v) = 0$ for all v, the corresponding values will be:

$$\begin{aligned} v_{m-2} &= \beta \left(\frac{1-\beta^{m-2}}{1-\beta} \right) b = (m-\delta)\, c \\ \text{and} \quad W(v_{m-2}) &= \beta \delta b = 2c \end{aligned}$$

if the rate of exchange of favors is less than one after two favors, or

$$\begin{aligned} v_{m-2} &= \beta \left(\frac{1-\beta^{m-2}}{1-\beta} \right) b = (m-1-\delta)\, c \\ \text{and} \quad W(v_{m-2}) &= \beta b + \beta^2 \delta b = 2c \end{aligned}$$

otherwise. Similarly here, $\dot{v}(v)$ will have to be different from 0 for some v on the interval $[v_{m-2}, v_{m-1}]$ in order to "fill in" some slack in the equations. In general, taking any interval between two values of player 1 where he receives some number of favors, there will be at least one value v for which $\dot{v}(v)$ is different from 0.

3. In the first best, $\frac{\dot{v}}{v} = r > 0$ for all values v such that $W(v_1) \leq v^*$ and $\frac{\dot{v}}{W(v)} = -r\frac{c}{b} < 0$ for all values v such that $W(v_2) \geq v^*$.

By the incentive compatibility constraint always binding, we have: $W(v_2) - W(v) = c$ and $rv = \alpha(v_2 - v + b) + \dot{v}$. Along values of player 1 in the interval $[0, v^*]$, the slope of the first best Pareto frontier is $-c/b$, which implies that along that interval, for all values v such that $W(v_1) \leq v^*$, $(v_2 - v) = -b$, and $\frac{\dot{v}}{v} = r$. A similar argument shows that along the interval $[v^*, v_h]$ where the slope of the frontier is $-b/c$, $\frac{\dot{v}}{W(v)} = -r\frac{c}{b}$ for all values such that $W(v_2) \geq v^*$.So in the first best, the drift is always in favor of the player with the lower value.

Sketch of proof of conjecture: $\dot{v}'(0) > 0$, $\dot{v}'(v_h) < 0$.

First case: If the slope of the Pareto frontier is equal to $-c/b$ as $v \longrightarrow 0$:

This would correspond to the case where the individual rationality constraint binds for the last fraction of favor. If the range of values over which $W'(v) = -c/b$ is large enough, we can apply the same argument as in the first best case to show that $\dot{v}(v)$ has to be > 0.

Second case: If the lowest value that a player can attain $\underline{v} > 0$, where $\underline{v}$ is defined as the value below which no favor is received. Clearly, from the incentive constraint binding, this can only be true if $\dot{v} > 0$. Moreover, if $\frac{\dot{v}}{\underline{v}} = r$.

Third case: If the slope of the Pareto frontier $> -c/b$ and $\underline{v} = 0$:

Let $[0, \check{v}]$ be the set of values for player 1 such that $v_2(v) = \underline{v}$, so whenever player 1 receives a full or a fraction of a favor, his continuation value is $\underline{v}$. Suppose that $\dot{v}(v) = 0$ over that set. Then for all $v \in [0, \check{v}]$,

$$v = \frac{\alpha}{r+\alpha} x_2 b$$

and

$$\frac{W(0) - W(v)}{v} = \frac{x_2 c}{\frac{\alpha}{r+\alpha} x_2 b} = \frac{r+\alpha}{\alpha} \frac{c}{b}$$

Hence, the slope is a constant along that interval. By discounting, for player 1, going from value 0 to $\check{v}$ requires him to do more than one favor. Let $[0, \hat{v}] = \{v | v_1(v) \in [0, \check{v}]\}$, so when player 1 does a favor, his continuation value remains in the interval $[0, \check{v}]$. For those values,

$$\begin{aligned} rW(v) &= \alpha(W(v^+) - W(v) + x_1 b) \\ W(v) &= \frac{r+\alpha}{r} \frac{c^2}{b} + \frac{\alpha}{r} b \end{aligned}$$

which is a constant. Hence, in order to generate a range of values, it cannot be be that $\dot{v}(v) = 0$ all along that interval.

By extension, by taking any smaller neighborhood to the right of 0, if $\dot{v}(v) = 0$ over that small interval, then the slope of the Pareto frontier will be constant.

Appendix B: Factorization and Algorithm:

In this appendix, we show that values in V^* can be factorized according to the formulation we used without loss of generality, and provide a simple algorithm which approximates the Pareto frontier from above.

Let $B(V)$ denote the APS operator associated to $T = \infty$.

Lemma 1 *If $V \subset B_T(V)$, then $V \subset B(V)$.*

Proof: Let $(v, w) \in B_T(V)$. By monotonicity, (v, w) is also in $B_T^n(V)$ for all n with factorization $(x_i(t), v_i(t), w_i(t), v(nT), w(nT))_{t=0}^{nT}$ and associated values $z_i(t)$ such that:

$$\begin{aligned} v &= \alpha \int_0^{nT} \beta^t z_1(t)\, dt + \beta^{nT} v_{nT} \\ w &= \alpha \int_0^{nT} \beta^t z_2(t)\, dt + \beta^{nT} w_{nT}, \end{aligned}$$

such that as n is increased all previous $z_i(t)$ terms are maintained. Taking the limit of $(x_i(t), v_i(t), w_i(t))$ as $n \to \infty$ delivers a factorization of (v, w) for B. (This is like in a standard dynamic programming problem iterating forward the optimal policy.)

Lemma 2 *If $V = B(V)$ then $V \subset B_T(V)$.*

Proof: Take $(v, w) \in B(V)$ with factorization $\{x_i(t), v_i(t), w_i(t)\}$ with corresponding values $\{z_i(t)\}$. Let $v_T = \alpha \int \beta^t z_1(t+T)\, dt$ and $w_T = \alpha \int \beta^t z_2(t+T)\, dt$. By definition the associated values $(v_i(t+T), w_i(t+T)) \in V$, so $(v_T, w_T) \in B(V) = V$. It follows immediately that $\{x_i(t), v_i(t), w_i(t), v(T), w(T)\}$ factorizes (v, w) for B_T.

Corollary 2 *The largest fixed points of B and B_T are the same.*

Proof: Let V be a fixed point of B. By Lemma 2, $V \subset B_T(V)$, so the largest fixed point of B_T contains V. Let V be a fixed point of B_T. By Lemma 1, $V \subset B(V)$, so the largest fixed point of B contains V.

Algorithm:

We consider a relaxed problem. More precisely, we will say that

$$\left(\{x_i(t), v_i(t), w_i(t)\}_{t=0}^{T}, v_T, w_T\right)$$

is weakly admissible with respect to V if the above conditions hold with the weaker incentive constraints:

$$\begin{aligned} v_1(t) - x_1(t)\, c &\geq \beta^T v_T \\ w_2(t) - x_2(t)\, c &\geq \beta^T w_T \end{aligned}$$

Lemma 3 *Suppose that $\left(\{x_i(t), v_i(t), w_i(t)\}_{t=0}^{T}, v_T, w_T\right)$ is weakly admissible with respect to convex V. Then the constant paths*

$$(x_1, x_2, v_1, v_2, w_1, w_2, v_T, w_T)$$

defined by:

$$\begin{aligned} x_i &= \frac{1}{\int_0^T \beta^t dt} \int_0^T \beta^t x_i(t)\, dt \\ v_i &= \frac{1}{\int_0^T \beta^t dt} \int_0^T \beta^t v_i(t)\, dt \\ w_i &= \frac{1}{\int_0^T \beta^t dt} \int_0^T \beta^t w_i(t)\, dt \end{aligned}$$

are weakly admissible with respect to V.

Proof: By convexity of V, the continuation values lie in V and it is also obvious that $0 \leq x_i \leq 1$. So we only need to verify (weak) incentive compatibility. The incentive constraints verify immediately.

Denote by $\hat{B}_T$ the associated APS operator. Note that by definition, the *averaged* path gives rise to the same values v, w. So without loss of generality in defining $\hat{B}_T$ we can restrict to constant paths. Because the incentive constraints defining $\hat{B}_T$ are weaker than those defining B_T, its largest fixed point contains the set of PPE values.