On the Optimality of the Equality Matching Form of Sociality*

Guilherme Carmona

Universidade Nova de Lisboa

Faculdade de Economia

Campus de Campolide, 1099-032 Lisboa, Portugal

email: gcarmona@fe.unl.pt

telephone: (351) 21 380 1671

fax: (351) 21 388 6073

June 16, 2006

*This paper is a heavily revised version of Chapter 4 of my Ph.D. thesis. I wish to thank my advisors, Narayana Kocherlakota, and Marcel K. Richter, for their guidance, encouragement, and very helpful comments and suggestions. I also wish to thank Pedro Amaral, Kemal Badur, Kyle Bagwell, Mehmet Barlo, Ronald Findley, Larry Gray, Leonid Hurwicz, Max Jodeit, Frank Lewis, Luís Pinto, Hamid Sabourian, Manuel Santos and William Sudderth for very helpful comments. I thank also John Huffstot for editorial assistance. While they all improved the material in many ways, any remaining errors are my own. Financial support from the *Subprograma Ciência e Tecnologia do 2º Quadro Comunitário de Apoio* is gratefully acknowledged.

Abstract

We consider a two-player game in which one player can take a costly action (i.e., to provide a favor) that is beneficial to the other. The game is infinitely repeated and each player is equally likely to be the one who can provide the favor in each period. In this context, equality matching is defined as a strategy in which each player counts the number of times she has given in excess of received and she gives if and only if this number has not reached an upper bound.

We show that the equality matching strategy is simple, self-enforcing, symmetric, and irreducible. Furthermore, we show that the utility for each player is at least as high under equality matching as under any other simple, self-enforcing, symmetric, and irreducible strategy of the same complexity. Thus, we rationalize equality matching as being an efficient way to achieve those properties.

This result is applied to risk sharing in village economies and used to rationalize the observed correlations between individual consumption and individual income and between present and past transfers across individuals.

Journal of Economic Literature Classification Numbers: C73, D02, 012, Z13.

Keywords: Equality matching, positive reciprocity, negative reciprocity, social institutions, village economies.

Correspondent: Guilherme Carmona, Universidade Nova de Lisboa, Faculdade de Economia, Campus de Campolide, 1099-032 Lisboa, Portugal. Email: gcarmona@fe.unl.pt; telephone: (351) 21 380 1671; fax: (351) 21 388 6073.

1 Introduction

People that live in the villages of developing countries typically have a low and highly volatile income. In the absence of insurance and credit markets, informal institutions have developed there in order to allow for some risk sharing across individuals. In fact, people in village economies transfer a significant part of their income in order to assist those who have received a low income (see, for example, Fafchamps and Lund (2003)).

This practice of transferring part of one's income to assist others is an example of the equality matching form of sociality defined in Fiske (1992). In this form of behavior, each person maintains a balance, which increases one unit when she takes a costly action and decreases one unit when she benefits from a costly action taken by another person. This balance is then used to decide whether or not she should take a costly action again: she will take it if and only if the balance has not reached an upper bound. In the case of the village economies, people not only transfer part of their income to those in need (typically, referred to as a form of positive reciprocity), they also stop giving if the other never reciprocates, or does not reciprocate enough (a form of negative reciprocity). Indeed, as Fafchamps and Lund (2001, p. 28) have shown, there is a significant negative correlation between current and past transfers received by individuals in their sample.

Why do we observe equality matching? Is there a sense in which this form of behavior is optimal? While one can easily explain the positive reciprocity aspect of equality matching through repeated interaction, is it the case that we can understand both its positive and negative aspect as being simultaneously part of an optimal equilibrium behavior?

In this essay, we provide an answer to these questions. We consider an infinitely repeated two-player game with no discounting in which one player can take a costly action (i.e., provide a favor) that is beneficial to the other, and in which each player is equally likely to be the one who can provide the favor in every period.[1] Several authors have pointed out that many real life institutions are self-enforcing, treat individuals symmetrically, cannot be simplified, and their rules are simple to understand. Following their work, we define a social institution as a repeated game strategy with those properties. Then, we show that the equality matching strategy satisfies those properties in an optimal way: the welfare of each player is at least as high under the equality matching strategy as under any other social institution of the same complexity. Hence, in this sense, equality matching is an optimal social institution.

Most of the properties we focus on are standard. We follow Schotter (1981, p. 24) in defining a social institution by a repeated game strategy. Following the work of Aumann (1981), and Selten (1975) respectively, we say that a strategy is simple if it can be represented by a finite automaton and is self-enforcing if it is subgame perfect.[2] The complexity of an automaton

[1]Thus, this game is a symmetric repeated dictator game, where by "symmetric" we mean that in every period each player is equally likely to be the dictator. Naturally, we assume that the benefit is higher than the cost of the action.

[2]An automaton is described by a set of states (one of which is specified to be the initial state), by a transition function (which gives the next period's state as a function of the current period's state and actions), and by a behavior function (which prescribes behavior according to the state of the automaton). As Kalai and Stanford (1988) have shown, an automaton is an equivalent way of describing a strategy, and so throughout this paper the two terms will be used synonymously.

is defined as the size of the state space as in Rubinstein (1986). Finally, an automaton is symmetric if players use the same state space, initial state and transition function and if players in the same situation play the same action.[3] This notion is similar to that in Okuno-Fujiwara and Postlewaite (1995).

While the other properties we use require little comment, it is useful to discuss in some detail the concept of irreducibility. Formally, an automaton is irreducible if all its states can be reached from any other state. Note that in a reducible automaton, there is a state s' that is never reached from a state s. Therefore, the automaton can be simplified when it reaches state s by reducing state s' in a way that it produces the same outcome. Thus, only irreducible automata may be impossible to simplify.

Furthermore, we argue that irreducibility discards some strategies that are based on empty threats when players care about the complexity of the strategy they use. We illustrate this point using the grim-trigger strategy, which can be represented by a two-state automaton with a cooperative state and a punishment state (see Kalai (1990, p. 141)). Consider suggesting to the players that they use grim-trigger. Player 1 could then reason as follows: "Player 2 is using grim-trigger; if I did so as well then a favor is provided in every period. However, if I use grim-trigger, I have to study the history of the play in every period to determine whether someone has failed to provide a favor in the past. I do not like to do that. Fortunately, there is a better alternative: I do not look at the history, and I always provide

[3]We impose still an additional condition. As explained in footnote 8, an automaton with the above properties induces a Markov chain on the set of states; in our definition of symmetry, we require that its transition matrix be symmetric. See Section 2 for further discussion.

a favor when I can. This leads exactly to the same outcome, hence payoff, and I do not have to bother to check what has happened in the past. I will do this instead of grim-trigger." This argument shows that, if players care about complexity, then grim-trigger is not an outcome we should expect: each player simply does not have an incentive to play it because he has a less complex alternative way of obtaining the same payoff. This happens precisely because grim-trigger is not irreducible, since the punishment state is never reached from the cooperative state.[4]

In the particular context of village economies, a further reason for irreducibility is that people may not want those who fail to reciprocate to be severely punished. This may be the case since the people with whom any person interacts are typically family members or close, long-term friends. But what can "not too severely punished" mean? A possible meaning is that nothing unusual happens if one deviates from the equilibrium strategies. More precisely, that the actions taken after someone fails to reciprocate when he should (i.e., outside the equilibrium path) are also taken regularly in the regular course of the game (i.e., on the equilibrium path). But this is clearly implied by irreducibility.

In conclusion, if we accept that social institutions are represented by finite, subgame perfect, symmetric and irreducible automata, then we can rationalize equality matching as an optimal social institution. Furthermore, in the particular case in which the costly action consists of transferring part of an individual's endowment, equality matching implies a particular pattern

[4]Formally, irreducibility is a necessary condition for semi-perfection, an equilibrium concept developed by Rubinstein (1986).

of individual consumption and transfers that is consistent with observed correlations in village economies. In fact, it implies some risk sharing, which is not complete due to its negative reciprocity aspect. Moreover, it implies a positive correlation between individual consumption and current and lagged individual income (documented in Townsend (1994), among others) and a negative correlation between current and past transfers among individuals (reported, for instance, in Fafchamps and Lund (2003) and La Ferrara (2003)).

2 Related Literature

The rationale for equality matching is also analyzed by Abdulkadiroğlu and Bagwell (2005). They study a repeated trust game with incomplete information and, for most of the paper, they focus on payoffs that lie on a symmetric self-generating line. They show that equality matching can be specified by a symmetric self-generating line although not by the highest one. Nevertheless, they show that the highest symmetric self-generating line can be implemented in a way that reflects an intertemporal balancing of favors. The difference between equality matching and this implementation is that the size of the favor owed diminishes in neutral periods where no player can provide a favor, but other than that, this optimal implementation is consistent with equality matching (accordingly, it was named sophisticated equality matching by Abdulkadiroğlu and Bagwell (2005)).

Compared to the work of Abdulkadiroğlu and Bagwell (2005), our results have the advantage of rationalizing equality matching exactly as described in

Fiske (1992) and in a complete information environment. This latter feature is important since this seems to be a reasonable assumption in the context of village economies. However, this requires a strengthening of the equilibrium concept, namely the introduction of irreducibility.

The conclusion that it is reasonable that the size of the favor owed diminishes in neutral periods is also reached by Hauser and Hopenhayn (2004) in a model similar to the one in Abdulkadiroğlu and Bagwell (2005) except that time is continuous.[5] Furthermore, they show that players would improve their well-being if the rate of exchange between favors received and conceded were to change with players' balance (recall that in the equality matching form of behavior described in Fiske (1992) this rate is always one).

In our model there is always a player that can provide a favor. However, equality matching would still be optimal even if we relaxed this assumption. The reason why the size of the favor owed does not diminish in neutral periods in our model is due to no discounting. Hence, there is no contradiction between our results and those of Abdulkadiroğlu and Bagwell (2005) and Hauser and Hopenhayn (2004), which depend on discounting. We can also interpret this difference in the results as suggesting that, in certain circumstances, the no-discounting case provides a better description. This seems to be the case not only in the examples provided by Fiske (1992), but also for risk sharing in village economies: in fact, as Fafchamps and Lund (2003) have shown, in their sample of rural Filipino households, risk is shared through zero-interest informal loans with an open-ended repayment period. This is

[5]That equality matching is a (Markov perfect) equilibrium in such a model was established first by Möbius (2001).

just a form of equality matching in which there is no decrease in the amount owed.

The reason why the dependence of the rates of exchange on the balance is not part of the optimal equilibrium is due to our symmetry assumption. In particular, this assumption requires that the transition matrix of the Markov chain induced by an automaton be symmetric. If we drop this requirement, which is not an intuitive economic condition, then it is possible to obtain a strategy that yields a higher payoff to both players. For instance, let players' possible values for their balance be any integer from 0 to 4 and let favors provided increase the balance by 2 units if the balance of the player providing it is 0 and by 1 unit otherwise. This change in the rate of exchange makes favors more likely (since they take place unless the player that can provide it has a balance of 4), thus increasing players' payoffs.

However, the increase in the payoff obtained by dropping the above condition does not come without costs. In fact, this strategy is harder to sustain as a subgame perfect equilibrium and so one needs to add an extra assumption. Indeed, a player (say, player 1) could deviate by providing a favor only when the "price" is high. This implies that the probability of reaching a balance higher than 2 for player 1 is zero starting from a balance less than or equal to 2. Hence, in the long run, player 1 is providing favor at a price of 2 and receiving then at a price of 1. The disadvantage is that favors are less likely. However, if the difference between the benefit of receiving a favor and the cost of providing it is not sufficiently high, then the deviation is profitable. In contrast, the only assumption needed to support the equality matching strategy with a rate of exchange identically equal to one is simply that such difference be strictly positive.

Our results are also related to those regarding risk sharing in village economies. When applied to this problem, our framework is similar to that in Kocherlakota (1996), except for the following differences: we consider no-discounting, an indivisible good, no aggregate uncertainty (in fact, there are only two equally likely states and each player has a positive endowment in only one of them), and, of course, a stronger equilibrium concept. The advantage of our theory is that it has the potential to generate stronger correlations between individual consumption and individual income, current and lagged, and between current and past transfers. In particular, it generates non-zero correlations even if consumers are extremely patient. This is in contrast to the main results of Kocherlakota (1996) since: first, if players are sufficiently patient, they predict that those correlations equal zero; second, as Koeppl (2006) and Rincón-Zapatero and Santos (2006) have shown, this can still be the case even if players are sufficiently impatient.

3 The Model

There are two players that interact in every period $t \in \mathbb{N}$. In every period, one of them can provide a favor to the other; we assume that this is decided by nature, in a way that each player has in every period a $1/2$ probability of being the one who can provide the favor.

When a player provides a favor, he suffers a utility cost $d > 0$, and the player receiving it obtains a positive utility $u > 0$. If the favor is not provided, then both players receive zero utility. We assume that favors are efficient in the sense that their benefit exceeds their cost. That is, we assume that $u > d$.

Let $N = \{1, 2\}$ stand for the set of players, $\Omega = \{1, 2\}$ for the set of states of nature, and $A = \{P, NP\}$ for the set of possible actions. We make the convention that when the state of nature equals 1, only player 1 can provide a favor, and so he chooses an action from the set A; similarly, when the state of nature equals 2, player 2 is the one who can provide the favor. The payoffs, which players receive period-wise, and which depend on the state of nature and on the choice made by the player who can provide the favor, are summarized in the following table:

$\omega \backslash a$	P	NP
1	$-d, u$	0,0
2	$u, -d$	0,0

Table 1: Stage Game Payoffs

We denote the period-wise payoffs as $u_i(\omega, a)$.

We describe the behavior of each player in the repeated game by an automaton. An *automaton* for player i is a triple $I_i = ((S_i, \bar{s}_i), T_i, B_i)$ where: S_i is a set of *states*; $\bar{s}_i \in S_i$ is the *initial state*; $T_i : \Omega \times S_i \times A \rightarrow S_i$ is a *transition function*; and $B_i : S_i \rightarrow A$ is a *behavior function.*

A pair of individual automata $I = (I_1, I_2)$, or for short, an automaton, together with a sequence of states of nature $\boldsymbol{\omega} = \{\omega_k\}_{k=1}^{\infty} \subseteq \Omega$ induce a sequence of actions $\mathbf{a}(I, \boldsymbol{\omega}) = \{a_k\}_{k=1}^{\infty} \subseteq A$ in the following way: $a_1 = B_{\omega_1}(\bar{s}_{\omega_1})$, and $a_k = B_{\omega_k}(s_{\omega_k}^k)$, where $s_i^k = T_i(s_i^{k-1}, a_{k-1})$, for both $i = 1, 2$.[6]

[6]Recall that player i is the producer in period k if $\omega_k = i$, for all $i = 1, 2$.

For $\omega^n \in \Omega^n$, we define the $n-$dimensional vector $\mathbf{a}(I, \omega^n)$ in a similar way and let $\mathbf{a}_k(I, \omega^n)$ denote its kth coordinate for all $1 \leq k \leq n$.

Each player's payoff in the repeated game depends on the payoff he receives in all periods, in the following way: first, for $i = 1, 2$, and $n \in \mathbb{N}$, we define a function $U_i^n(I) : \Omega^n \to \mathbb{R}$ by defining

$$U_i^n(I)(\omega^n) = \frac{1}{n} \sum_{k=1}^{n} u_i(\omega_k, \mathbf{a}_k(I, \omega^n)), \tag{1}$$

and we define

$$U_i^n(I) = \frac{1}{2^n} \sum_{\omega^n \in \Omega^n} U_i^n(I)(\omega^n). \tag{2}$$

Then, payoff of an automaton I for player i, $i = 1, 2$, is

$$U_i(I) = \limsup_{n \to \infty} U_i^n(I). \tag{3}$$

By using the above payoff function, we are assuming that players are extremely patient. In fact, as shown below, the payoff of any irreducible automaton equals the limit, as the discount factor goes to one, of the payoffs computed using the discounted sum criterion. Since players incur a cost whenever they provide a favor, favors will occur in equilibrium only if players are sufficiently patient. By using the above payoff function, we are able to present our results in a clearer way, while allowing us to simplify some of the proofs.

4 Equality Matching

In our framework, it is natural that players choose to provide favors, at least at some times. However, since this is costly, the provision of favors will not be unconditional. One possible way of conditioning the provision of favors is described by the equality matching form of sociality.

Under equality matching, a player will provide a favor if and only if he has not given much more than he has received in the past. In other words, the player that can provide a favor bases his decision on the number of times he has given in excess of received, and chooses to provide it whenever this number is below a certain threshold.

Formally, an *equality matching automaton* $I_M = (I_1^M, I_2^M)$ *with a threshold* $M \in \mathbb{N}$ is defined as follows: the set of states is

$$S_1^M = S_2^M = S_M = \{0, \ldots, M\}, \tag{4}$$

and the initial state is $\bar{s}_M \in S_M$.[7] The transition function $T_1^M = T_2^M = T_M : \Omega \times S_M \times A \to S_M$ is defined by:

$$T_M(1, m, P) = \begin{cases} m+1 & \text{if } m \leq M-1, \\ M & \text{if } m = M \end{cases} \tag{5}$$

$$T_M(2, m, P) = \begin{cases} m-1 & \text{if } m \geq 1, \\ 0 & \text{if } m = 0 \end{cases} \tag{6}$$

$$T_M(\omega, m, NP) = m, \tag{7}$$

The interpretation is as follows: $s \in S_M$ represents the balance of player 1. Whenever player 1 provides a favor, her balance increases by 1 unit, except when this balance has reached the upper bound M. Similarly, whenever she receives a favor, her balance decreases by 1 unit, except when it has reached

[7] Two equality matching automata are distinct if and only if they differ in their initial state. Since players do not discount the future, the initial state does not play any role in our analysis.

0. Since player 2's balance is just $M - s$, the latter case occurs exactly when player 2's balance has reached the upper bound.

Player 1's behavior function is defined by:

$$B_1^M(m) = \begin{cases} P \text{ if } m < M, \\ NP \text{ otherwise;} \end{cases} \tag{8}$$

Similarly, Player 2's behavior function is defined as follows:

$$B_2^M(m) = \begin{cases} P \text{ if } m > 0, \\ NP \text{ otherwise;} \end{cases} \tag{9}$$

Intuitively, any player provides a favor if and only if the other player has a positive balance, which occurs if and only if his own balance has not reached the upper bound. The definition of I_M does describe equality matching in the sense that each player takes costly actions that benefit the other, but will stop doing so if this other player does not reciprocate enough.

Note that the equality matching automaton satisfies many symmetry properties. First, we have that players use a common state space, initial state, and transition function: $S_1^M = S_2^M$, $\bar{s}_1^M = \bar{s}_2^M$, and $T_1^M = T_2^M$. Second, some states can be associated in a natural way: if we define $\phi(m) = M - m$, we obtain a bijection $\phi : S^M \to S^M$, satisfying $B_1^M(m) = B_2^M(\phi(m))$. Third, the equality matching automaton induces a Markov chain on S^M, described by a symmetric transition matrix. In fact, if Π_M denotes such a matrix, one easily sees that the nonzero entries of Π_M are:

$$\begin{aligned} &\pi_{0,0} = \frac{1}{2},\ \pi_{0,1} = \frac{1}{2} \\ &\pi_{m,m-1} = \frac{1}{2},\ \pi_{m,m+1} = \frac{1}{2},\ \text{for all } 0 < m < M \\ &\pi_{M,M-1} = \frac{1}{2},\ \pi_{M,M} = \frac{1}{2}. \end{aligned} \tag{10}$$

Generalizing from the particular case of the equality matching automaton, we say that an automaton $I = (I_1, I_2)$ is *symmetric* if: (1) $S_1 = S_2$, $T_1 = T_2$, and $\bar{s}_1 = \bar{s}_2$; (2) there exists a bijection $\phi : S \to S$ such that $B_1(s) = B_2(\phi(s))$; and (3) I induces a Markov chain on S, described by a symmetric transition matrix.[8] Intuitively, the class of symmetric automata consists of those in which different individuals in the same situation determined by the realization of the uncertainty and with the same state are prescribed by the same action.

5 Equality Matching as an Optimal Social Institution

In our model the two players interact in every period of time. This interaction is described by an automaton I, which consist of a pair of individual automata: $I = (I_1, I_2)$. By changing each player's automaton, we obtain different outcomes, some of which may be unreasonable.

The first requirement we impose on the automaton that players use is that it is self-enforcing. More precisely, we will require that each player, given the other player's behavior, has an incentive to act in the way that the

[8] Note that any finite automaton I satisfying (1) induces a Markov chain on S defined by the following transition matrix Π:

$$\pi_{ss'} = \begin{cases} 1 & \text{if} \quad T\left(\omega, s, B_\omega(s)\right) = s' \text{ for all } \omega = 1, 2, \\ 0 & \text{if} \quad T\left(\omega, s, B_\omega(s)\right) \neq s' \text{ for all } \omega = 1, 2, \\ \frac{1}{2} & \text{otherwise.} \end{cases} \tag{11}$$

automaton prescribes for all possible contingencies. Formally, this amounts to requiring that the automaton is a subgame perfect equilibrium.

Before giving the formal definition, we need the following notation: given an automaton $I = ((S, \bar{s}), T, B)$, then $(I, s) = ((S, s), T, B)$ denotes the automaton that differs from I only in the initial state. We then say that an automaton I is a *subgame perfect equilibrium* if for all $i = 1, 2$, s in $S = S_1 \times S_2$, and any player i's automaton I'_i with initial state s', we have that

$$U_i(I, s) \geq U_i((I'_i, I_{-i}), (s', s_{-i})). \tag{12}$$

A second requirement we impose is that there are no obsolete states: all states should be used regularly in the regular course of the game. As Rubinstein pointed out "[these] considerations have some similarity to phenomena frequently observed in real life: social institutions, various types of organizations, and human abilities degenerate or are readily discarded if they are not used regularly." Formally, we say that a symmetric automaton I is *irreducible* if the Markov chain induced by I is irreducible.[9]

The view that we take here is that, in our framework, only automata that are finite, symmetric, subgame perfect, and irreducible can describe a social institution. For $N \in \mathbb{N}$, let $\mathcal{A}_N$ be the set of all symmetric, irreducible, subgame perfect automata with a state space having no more than N elements. Our main result is:

Proposition 1 *For all $N \in \mathbb{N}$, every equality matching automaton I_M, with*

[9] A Markov chain represented by a transition matrix Π is irreducible if for all states s, and s' there exists $K \in \mathbb{N}$ such that $\pi^{(K)}_{s,s'} > 0$.

$M = N - 1$, *solves*

$$\max_{I \in \mathcal{A}_N} U_1(I) + U_2(I).$$

Proposition 1 asserts that not only is the equality matching automaton a symmetric, irreducible, subgame perfect automaton for any possible upper bound M, but in fact, it is efficient within that class. In particular, any finite, symmetric, irreducible, subgame perfect automaton can be (weakly) dominated by the equality matching automaton of the same level of complexity. In this sense, equality matching is an optimal social institution.

6 Risk Sharing in Village Economies

As an application of the framework in Section 3, we consider the problem of risk sharing in village economies (see, among others, Kocherlakota (1996) and Ligon, Thomas, and Worrall (2002)). This application is particularly interesting since there is considerable empirical evidence that can be used to test our results. As Proposition 2 below states, if players use equality matching, then the correlation between current individual consumption and current and lagged income is positive. Furthermore, it implies that current individual net transfers are negatively correlated with the previous period's individual net transfers.

The model is like the one in Section 3. There are two players who interact in every period $n \in \mathbb{N}$. In every period, each person receives an endowment of a single perishable and indivisible consumption good. The pair of endowments belongs to $\{(0, 2), (2, 0)\}$, that is, one person receives 2 units of the good while the other receives 0. The endowments are determined by nature: the set $\Omega = \{1, 2\}$ denotes the set of states of nature. The relationship between endowments and the state of nature is as follows: if $\omega = 1$ then $y = (y^1_\omega, y^2_\omega) = (2, 0)$, while if $\omega = 2$ then $y = (0, 2)$. The values of $\omega \in \Omega$ are drawn independently and are equally likely, i.e., each occurs with probability equal to 1/2. Note that the aggregate endowment $Y = y^1 + y^2$ is always equal to 2, i.e., there is no aggregate uncertainty.

The two individuals interact in the following way. At the beginning of each period $n \in \mathbb{N}$ they are informed about the current value of ω. Then, the player with 2 units of the good can choose to transfer 1 unit of his current endowment to the other. We let $t^i_{\omega,t}$ denote the transfer made by individual i in state of nature ω in period n; it has to satisfy $0 \leq t^i_{\omega,n} \leq y^i_\omega$ and $t^i_{\omega,n} \in \{0, 1\}$. Once the decision regarding transfers is made, each individual consumes $c^i_{\omega,n} = y^i_\omega - t^i_{\omega,n} + t^{-i}_{\omega,n}$. The net transfer received by player 1 is $\theta_{\omega,n} = t^2_{\omega,n} - t^1_{\omega,n}$.

Individuals have the same period-wise utility function defined on consumption levels $u : \{0, 1, 2\} \to \mathbb{R}$. We let $u(0) = 0$, $u(1) = u$ and $u(2) = u+d$ and assume that $u > d > 0$. In this way u is strictly increasing and satisfies the following form of strict concavity:

$$\frac{u(0) + u(2)}{2} = \frac{u + d}{2} < u = u(1) = u\left(\frac{0 + 2}{2}\right). \tag{13}$$

For all $i = 1, 2$, we define $u_i : \Omega \times A \to \mathbb{R}$ as follows:

$$u_i(\omega, t) = u(y^i_\omega - t^i + t^{-i}). \tag{14}$$

This game is the same as the one in Section 3: only one player can choose to transfer; if she does, she loses d (since she consumes 1 unit instead of 2) and the other player gains u (he consumes 1 unit instead of 0). Thus, the player with a high endowment is, effectively, like the player who can provide a favor in the model of Section 3 and the costly action that she can take is to transfer one unit of her endowment to the other player. Therefore, we can define equality matching in the same way as before, and use Proposition 1 to conclude that for all $N \in \mathbb{N}$, every equality matching automaton I_M, with $M = N - 1$, solves $\max_{I \in \mathcal{A}_N} U_1(I) + U_2(I)$.

This behavior, in turn, implies that the pattern of consumption and transfers satisfies the following properties.

Proposition 2 *For every equality matching automation I_M, there exists $\alpha > 0$ such that*

1. $\text{cov}(c_t^1, y_{t-k}^1 | Y_{t-k}) > 0$ *and*

2. $\text{cov}(\theta_t, \theta_{t-1} | Y_{t-1}) < 0$

for all $t \geq \alpha$ and $0 \leq k \leq t - \alpha$.

It is easy to explain why this behavior leads to a positive correlation between current individual consumption and current and lagged income. If a consumer has a zero balance, she can consume if and only if she receives a positive endowment. Also, a consumer with a zero endowment today and a zero balance yesterday can consume today if and only if she received a positive endowment yesterday. Thus, the equality matching form of behavior can make current individual consumption and current and lagged individual income move together. A similar intuition holds for transfers.

7 Concluding Remarks

The main point of the paper is that the equality matching form of sociality can be regarded as an optimal social institution. When applied to a simple risk sharing problem, this conclusion provides an explanation for the observed correlations between individual consumption and individual income, current and lagged, and between current and past transfers in village economies.

As explained in Section 2, these correlations are stronger than those predicted in Kocherlakota (1996), at least when there is little or no discounting. Unfortunately, it may be hard to empirically distinguish these two theories. It can, nevertheless, be tested in the situation where there is evidence that players are sufficiently patient.

Furthermore, they can also be tested by using experiments to find out how players behave out of the equilibrium path. For instance, the experimental subjects could be told that there has been a history of plays before they start. Evidence on how players play out of the equilibrium path can be used to distinguish the two approaches since, in contrast to the equality matching strategy, the optimal strategy in Kocherlakota (1996) predicts that those who fail to transfer as prescribed by the equilibrium strategies will remain in autarky for at least a long period.

Just as experimental evidence can be useful to better understand our results, these can also be used to comment on some recent experimental studies on village economies. One such study was done by Henrich, Boyd, Bowles, Camerer, Fehr, Gintis, and McElreath (2001), who report experimental results on the one-shot ultimatum game played in fifteen village economies. They found that the mean offers were substantially above zero, ranging from

25 to 50 percent of the stake size, despite the fact that the game was played just once and anonymously, and players were randomly matched.

Perhaps more importantly, they have shown that this behavior is consistent with the typical everyday-life behavior in these economies. This observation suggests that repeated-game effects in experiments may be more subtle than what is generally considered. Even if players understand that they are playing a one-shot game, their behavior may be guided by the social norms of their society, which are designed for everyday, repeated interaction. Hence, their behavior will reflect not the optimal one-shot behavior, but rather the optimal choices in familiar, recurrent situations that are similar to the game being played.

Furthermore, as Kandori (1992) and Ellison (1994) have shown, we can expect deviations from optimal one-shot behavior even when a large population interacts in an anonymous, random matching way. Hence, the fact that players in experiments play the game anonymously and are randomly matched with other players may not be enough to test one-shot games.

If one accepts that the experimental evidence in Henrich et al. (2001) reflects the behavior induced by optimal institutions, then our results can be used to reconcile it with the canonical economic model of self-interested players. In fact, our results imply that we should expect that people in village economies transfer a considerable amount of their resources if they behave in the way predicted by the optimal equilibrium in our model of self-interested players.

However even if people are naive and play games by following their instinct and their emotional impulses, it is still likely that their choices reflect the social institutions and culture of the society they live in. If, in fact, the social institutions that endure are those that are simple, self-enforcing, symmetric, and irreducible in an efficient way, then, the behavior of any individual will be well described by the canonical model, even if he is naive and not completely self-interested.

A Appendix

A.1 Proof of Proposition 1

In this appendix, we prove Proposition 1. The first step of the proof is to show that any equality matching automaton is symmetric, irreducible, and subgame perfect. One easily sees that any equality matching automaton is symmetric. Since any state can lead to the two adjacent states, with the convention that both state 0 and state M are adjacent to themselves, one can conclude that any equality matching automaton is irreducible.

To show that each equality matching automaton I_M is a subgame perfect equilibrium, we will start by considering the case in which players' payoff in the repeated game equals

$$U_i^{\delta}(\boldsymbol{\omega}, \mathbf{a}) = (1-\delta)\sum_{k=1}^{\infty}\delta^k u_i(\omega_k, a_k), \qquad (15)$$

for all i, where $\delta \in (0,1)$, $\boldsymbol{\omega} = \{\omega_k\}_{k=1}^{\infty} \subseteq \Omega$, and $\mathbf{a} = \{a_k\}_{k=1}^{\infty} \subseteq A$. Existing results guarantee that if I_M is subgame perfect for all discount factors close to 1, then I_M is subgame perfect in our game.

The following lemma estimates the benefit for a given player i of having a larger balance (which also implies that the other player will have a smaller one). It shows that, with probability 1, either player i receives exactly one more favor or provides exactly one less favor.

Lemma 1 *Let* $m \in \{0, \ldots, M-1\}$*. Then there is* $\alpha_m \in (0,1)$ *such that*

$$\lim_{\delta \to 1} \frac{1}{1-\delta}(U_i^\delta(I_M, m+1) - U_i^\delta(I_M, m)) = \alpha_m u + (1-\alpha_m)d,$$

for $i = 1, 2$.

Proof. Because player 1's case is symmetric to player 2's, we deal only with the first.

Step 1: Some definitions.

Denote $m(0) = (m+1, M-m-1)$ and $m'(0) = (m, M-m)$. Also, let $\boldsymbol{\Omega} := \Omega \times \Omega \times ...$ be the countable infinite Cartesian product of Ω, and let $(\boldsymbol{\Omega}, \mathcal{G}, \mu)$ denote the usual corresponding probability space. A generic element of $\boldsymbol{\Omega}$ is denoted by $\boldsymbol{\omega} = \{\omega_t\}_{t=1}^{\infty}$, where $\omega_t \in \Omega$, for all $t \in \mathbb{N}$. Given $\boldsymbol{\omega}$, let $m(k)(\boldsymbol{\omega}) = (m_1(k)(\boldsymbol{\omega}), m_2(k)(\boldsymbol{\omega}))$ denote the balance players have at the end of stage k if they started with $m(0)$ and let $m'(k)(\boldsymbol{\omega})$ denote the balance players have at the end of stage k if they started with $m'(0)$.

With this notation, we can write

$$\begin{aligned}
&\frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m)) = \\
&\frac{1}{1-\delta}\left(\int_{\boldsymbol{\Omega}} U_1^\delta(I_M, m+1)(\boldsymbol{\omega})d\mu - \int_{\boldsymbol{\Omega}} U_1^\delta(I_M, m)(\boldsymbol{\omega})d\mu\right) = \\
&\frac{1}{1-\delta}\left(\int_{\boldsymbol{\Omega}} \left[U_1^\delta(I_M, m+1,)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega})\right] d\mu\right),
\end{aligned} \tag{16}$$

where the last equality follows because both of the functions $\boldsymbol{\omega} \mapsto U_1^\delta(I_M, m)(\boldsymbol{\omega})$ and $\boldsymbol{\omega} \mapsto U_1^\delta(I_M, m+1)(\boldsymbol{\omega})$ are integrable.

Step 2: There exists $\{A_t, B_t\}_{t=1}^{\infty}$ such that $\frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m,)) = \sum_{t=1}^{\infty} d\delta^{t-1}\mu(B_t) + \sum_{t=1}^{\infty} u\delta^{t-1}\mu(A_t)$.

Let $A_1 := \{\boldsymbol{\omega} \in \boldsymbol{\Omega} : m_1(1)(\boldsymbol{\omega}) = 0 \text{ and } \omega_1 = 2\}$; in A_1 player 1 is able to receive a favor under m but not under m'. Note also that $m(1) = m'(1)$. So, for $\boldsymbol{\omega} \in A_1$, the difference in payoffs is u. That is, for $\boldsymbol{\omega} \in A_1$,

$$U_1^\delta(I_M, m+1)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega}) = (1-\delta)u. \tag{17}$$

Let $B_1 := \{\boldsymbol{\omega} \in \boldsymbol{\Omega} : m_2'(1)(\boldsymbol{\omega}) = 0 \text{ and } \omega_1 = 1\}$; in B_1 player i has to provide a favor under m' but not under m. Note also that $m(1) = m'(1)$. Thus, for $\boldsymbol{\omega} \in B_1$,

$$U_1^\delta(I_M, m+1)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega}) = (1-\delta)d. \tag{18}$$

We proceed by induction: let $t \geq 2$. Let

$$A_t := \{\boldsymbol{\omega} \in \boldsymbol{\Omega} \backslash ((\overset{t-1}{\underset{k=1}{\cup}} A_k) \cup (\overset{t-1}{\underset{k=1}{\cup}} B_k)) : m_1(t)(\boldsymbol{\omega}) = 0 \text{ and } \omega_t = 2)\} \tag{19}$$

and

$$B_t = \{\boldsymbol{\omega} \in \boldsymbol{\Omega} \backslash ((\overset{t-1}{\underset{k=1}{\cup}} A_k) \cup (\overset{t-1}{\underset{k=1}{\cup}} B_k)) : m_2'(t)(\boldsymbol{\omega}) = 0 \text{ and } \omega_t = 1\}. \tag{20}$$

Similarly as before, we have that for $\boldsymbol{\omega} \in A_t$,

$$U_1^\delta(I_M, m+1)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega}) = (1-\delta)\delta^{t-1}u, \tag{21}$$

and for $\boldsymbol{\omega} \in B_t$,

$$U_1^\delta(I_M, m+1)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega}) = (1-\delta)\delta^{t-1}d. \tag{22}$$

Finally let $C = \boldsymbol{\Omega} \backslash \left[\left(\overset{\infty}{\underset{t=1}{\cup}} A_t \right) \cup \left(\overset{\infty}{\underset{t=1}{\cup}} B_t \right) \right]$.

For each $t \in \mathbb{N}$, A_t is measurable since it can be written as $D_1 \times \cdots \times D_t \times \Omega \times \Omega \times \cdots$ for some $D_1, \ldots, D_t \subseteq \Omega$. Similarly, B_t is measurable for each $t \in \mathbb{N}$ and so is C. Note also that for all $j, k \in \mathbb{N}$, we have that $A_j \cap A_k = \emptyset$, $B_j \cap B_k = \emptyset$ and $A_j \cap B_k = \emptyset$.

Claim 1 $\mu(C) = 0$.

Proof. Let $S_n(\boldsymbol{\omega})$ be the number of times that $\omega_k = 1$ in the first n periods. If $\boldsymbol{\omega} \in C$, then it follows that $m_1(k)(\boldsymbol{\omega}) > 0$ whenever $\omega_k = 2$, for all k. Therefore, $n - S_n$ (which equals the number of times that $\omega_t = 2$) is also the "amount spent" by player 1 in the first n periods. Since the "amount received" by player 1 in the first n periods is at most S_n, then for each $n \in \mathbb{N}$, $m_1(0) + S_n \geq n - S_n$, that is $S_n/n \geq 1/2 - m_1(0)/2n$. Hence $C \subseteq \bigcap_{n=1}^{\infty} \{\boldsymbol{\omega} \in \boldsymbol{\Omega} : \frac{S_n(\boldsymbol{\omega})}{n} \geq \frac{1}{2} - \frac{m_1(0)}{2n}\}$, which has measure zero by lemma 2 applied to the sequence of random variables $\{X_n\}_{n=1}^{\infty}$, where for all n, $X_n(\boldsymbol{\omega}) = \chi_{\{\boldsymbol{\omega}:\, \omega_n = 1\}}$. ■

Hence, we obtain

$$\begin{aligned}
&\frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m)) = \\
&\frac{1}{1-\delta}(\sum_{t=1}^{\infty} \int_{A_t} [U_1^\delta(I_M, m+1)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega})] d\mu + \\
&+ \sum_{t=1}^{\infty} \int_{B_t} [U_1^\delta(I_M, m+1)(\boldsymbol{\omega}) - U_1^\delta(I_M, m)(\boldsymbol{\omega})] d\mu) = \\
&\sum_{t=1}^{\infty} u\delta^{t-1}\mu(A_t) + \sum_{t=1}^{\infty} d\delta^{t-1}\mu(B_t).
\end{aligned} \tag{23}$$

Step 3: There exists $\alpha_m \in (0, 1)$ such that

$$\lim_{\delta \to 1} \frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m)) = \alpha_m u + (1 - \alpha_m)d. \tag{24}$$

By Abel's theorem (see DePree and Swartz (1988, Theorem 11.17, p. 135)),

$$\lim_{\delta \to 1} \frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m)) = d\mu(\bigcup_{t=1}^{\infty} B_t) + u\mu(\bigcup_{t=1}^{\infty} A_t). \tag{25}$$

So define $\alpha_m := \mu(\bigcup_{t=1}^{\infty} A_t)$. Finally note that $\mu(\bigcup_{t=1}^{\infty} A_t) \geq (\frac{1}{2})^{m(0)} > 0$, $\mu(\bigcup_{t=1}^{\infty} B_t) \geq (\frac{1}{2})^{M-m(0)}$, $(\bigcup_{t=1}^{\infty} A_t) \cap (\bigcup_{t=1}^{\infty} B_t) = \emptyset$ and that $\boldsymbol{\Omega} = (\bigcup_{t=1}^{\infty} A_t) \cup (\bigcup_{t=1}^{\infty} B_t) \cup C$ implying that $\mu(\bigcup_{t=1}^{\infty} A_t) + \mu(\bigcup_{t=1}^{\infty} B_t) = 1$. ■

The following lemma was used above:

Lemma 2 *Let* $\{X_n\}_{n=1}^{\infty}$ *be a sequence of independent and identically distributed random variables with mean equal to* $\rho \leq \frac{1}{2}$ *and finite variance* $\sigma^2 > 0$ *and let* $c \in \mathbb{R}$*. Then* $\mu(\bigcap_{n=1}^{\infty}\{\boldsymbol{\omega} : \frac{S_n(\boldsymbol{\omega})}{n} \geq \frac{1}{2} - \frac{c}{n}\}) = 0$.

Proof. The result will follow from the Law of Iterated Logarithms: *Let* $\{Y_k\}_{k=1}^{\infty}$ *be independent and identically distributed random variables with* $E[Y_1] = 0$ *and* $\sigma^2(Y_1) = 1$*. Then* $\liminf_{n\to\infty} \frac{S_n}{\sqrt{2n\log\log n}} = -1$ *almost surely (see Billingsley (1995, Theorem 9.5, p. 154)).*

Define $Y_n(\boldsymbol{\omega}) = \frac{X_n(\boldsymbol{\omega})-\rho}{\sigma}$ and $S_n^y(\boldsymbol{\omega}) = \sum_{k=1}^{n} Y_k(\boldsymbol{\omega})$. Then $\frac{S_n^y(\boldsymbol{\omega})}{n} = \frac{1}{\sigma}\left(\frac{S_n(\boldsymbol{\omega})}{n} - \rho\right)$. By the Law of Iterated Logarithms, there is $Z \subset \boldsymbol{\Omega}$ with $\mu(Z) = 0$ such that $\liminf_{n\to\infty} \frac{S_n^y(\boldsymbol{\omega})}{\sqrt{2n\log\log n}} = -1$, for all $\boldsymbol{\omega} \in \boldsymbol{\Omega} \setminus Z$. Let $\boldsymbol{\omega} \in \boldsymbol{\Omega} \setminus Z$. Then, since $\inf_{n\geq k} \frac{S_n^y(\boldsymbol{\omega})}{\sqrt{2n\log\log n}}$ increases to $\liminf_{n\to\infty} \frac{S_n^y(\boldsymbol{\omega})}{\sqrt{2n\log\log n}}$, it follows that $\inf_{n\geq k} \frac{S_n^y(\boldsymbol{\omega})}{\sqrt{2n\log\log n}} \leq -1$, for all $k \in \mathbb{N}$. That is, $\frac{S_n^y(\boldsymbol{\omega})}{\sqrt{2n\log\log n}} \leq -1$ infinitely often. Thus, for n large enough,

$$\frac{S_n(\boldsymbol{\omega})}{n} \leq \rho - \frac{\sigma\sqrt{2\log\log n}}{\sqrt{n}} < \frac{1}{2} - \frac{c}{n} \tag{26}$$

infinitely often (the last inequality follows because, for n large enough, we have that $c < \sigma\sqrt{2n\log\log n} \to \infty$). It follows then that $\boldsymbol{\omega} \notin \bigcap_{n=1}^{\infty}\{\boldsymbol{\omega} : \frac{S_n(\boldsymbol{\omega})}{n} \geq \frac{1}{2} - \frac{c}{n}\}$; hence $\bigcap_{n=1}^{\infty}\{\boldsymbol{\omega} : \frac{S_n(\boldsymbol{\omega})}{n} \geq \frac{1}{2} - \frac{c}{n}\} \subseteq Z$ and the result follows. ■

It is useful to use the discounted version of our game to show that I_M is subgame perfect, because in discounted games we can use the one-shot deviation principle (see Abreu (1988).) For the particular case of an equality matching automaton, we need to show that it is not profitable for a player to refuse to provide a favor when the other has a positive balance, and to follow the equality matching strategy afterwards.

If a player deviates by not providing a favor when the other has a positive balance, his utility increases today by d, i.e., he gains by not having to provide the favor. However, he starts the next period with one less unit on his balance, and the other player starts with one more unit. Thus, the following Lemma follows from Lemma 1.

Lemma 3 *There exists $\delta^* \in (0,1)$ such that for all $\delta \in (\delta^*, 1)$, I_M is subgame perfect.*

Proof. By Proposition 3.11 of Mertens and Parthasarathy (1987)) (or Proposition 1 of Abreu (1988)) it is enough to show that no player can profitably deviate from I_M by deviating just in the first stage. Again, because player 1's case is symmetric to player 2's, we deal only with the first.

It is clear that player 1 does not want to deviate from $B_1^M(M) = NP$, since by choosing P when his balance equals M he would reduce his utility today by d, and receive the same future utility. So we are left with showing that he does not want to deviate from $B_1^M(m)$, for all $m = 0, \ldots, M-1$.

Let $m \in \{0, \ldots, M-1\}$. If player 1 deviates from $B_1^M(m)$, and therefore chooses NP, his utility will be equal to

$$\bar{U} := (1-\delta)\delta U_1^\delta(I_M, m), \tag{27}$$

while if he does not deviate, his utility will be equal to

$$U_1^\delta(I_M, m) = (1-\delta)(-d + \delta U_1^\delta(I_M, m+1)). \tag{28}$$

Thus,

$$U_1^\delta(I_M, m) - \bar{U} = (1-\delta)\left[-d + \delta \frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m))\right]. \tag{29}$$

By lemma 1,

$$\begin{aligned} &- d + \delta \frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m)) \underset{\delta \to 1}{\rightarrow} \\ &\underset{\delta \to 1}{\rightarrow} -d + \alpha_m u + (1-\alpha_m) d > 0. \end{aligned} \tag{30}$$

Therefore, if we let δ^* be such that for all $\delta > \delta^*$

$$-d + \delta \frac{1}{1-\delta}(U_1^\delta(I_M, m+1) - U_1^\delta(I_M, m)) > 0 \tag{31}$$

for all $m \in \{0, \ldots, M-1\}$, then

$$U_1^\delta(I_M, m) - \bar{U} > 0 \tag{32}$$

for all $m \in \{0, \ldots, M-1\}$. ■

The second step of the proof is to show that $U_i(I_M) \geq U_i(I)$ for any $i = 1, 2$, and any symmetric, irreducible, and subgame perfect automaton I with $|S| \leq |S_M|$. The following Lemma uses standard ergodic theorems for Markov chains to compute the payoff of any symmetric, irreducible, and subgame perfect automaton.

Lemma 4 *Let $I \in \mathcal{A}_N$. Then,*

$$U_i(I) = \frac{1}{2\,|S|} \sum_{s \in S} \sum_{\omega \in \Omega} u_i(\omega, B(s)). \tag{33}$$

Proof. Let $I \in \mathcal{A}_N$, and let $\tilde{S} = \Omega \times S$. Then I also induces a Markov chain $\tilde{\Pi}$ on $\tilde{S}$ satisfying

$$\tilde{\pi}_{i,j} = \pi_{i_2,j_2}, \tag{34}$$

for all $i = (i_1, i_2)$, and $j = (j_1, j_2)$ in $\tilde{S}$.

Since Π is symmetric and irreducible, then so will be $\tilde{\Pi}$. Denoting $\tilde{s}_1 = (1, \bar{s}_1)$, and $\tilde{s}_2 = (2, \bar{s}_2)$, we have that

$$U_i^n(I) = \frac{1}{2n} \sum_{k=1}^{n} \sum_{(\omega,s)} \tilde{\pi}_{\tilde{s}_1,(\omega,s)}^{(k)} u_i(\omega, B(s)) + \frac{1}{2n} \sum_{k=1}^{n} \sum_{(\omega,s)} \tilde{\pi}_{\tilde{s}_2,(\omega,s)}^{(k)} u_i(\omega, B(s)). \tag{35}$$

By Theorems A.1, and A.4 of Derman (1970), we have that

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=1}^{n} \tilde{\pi}_{\tilde{s}_i,(\omega,s)}^{(k)} = \frac{1}{2\,|S|}, \tag{36}$$

for $i = 1, 2$, since the uniform distribution is the unique stationary distribution of $\tilde{\Pi}$. Thus,

$$U_i(I) = \lim_{n\to\infty} U_i^n(I) = \frac{1}{2|S|}\sum_{s\in S}\sum_{\omega\in\Omega} u_i(\omega, B(s)). \tag{37}$$

■

The following lemma states that in any subgame perfect equilibrium there has to be a "punishment" state, which, in our model, corresponds to a player refusing to provide a favor.

Lemma 5 *Let $I \in \mathcal{A}_N$. Then, for all $i = 1, 2$, there exists $s \in S$ such that $B_i(s) = NP$.*

Proof. Suppose that for some $i \in \{1, 2\}$, we have $B_i(s) = P$, for all $s \in S$. Let $\tilde{I}_{-i}$ be such that $B_{-i}(s) = NP$, for all $s \in S$, which implies that $U_{-i}(I_i, \tilde{I}_{-i}) = u/2$. Suppose, in order to reach a contradiction, that $I_{-i} \neq \tilde{I}_{-i}$. Since $I \in \mathcal{A}_N$, then by lemma 4 we obtain

$$U_{-i}(I) \leq \frac{u}{2} - \frac{d}{2|S|} < U_{-i}(I_i, \tilde{I}_{-i}), \tag{38}$$

a contradiction since I is a subgame perfect equilibrium. Thus, $I_{-i} = \tilde{I}_{-i}$, and so $U_i(I) = -d/2$.

However, letting $\tilde{I}_i$ be such that $B_i(s) = NP$, for all $s \in S$, we obtain $U_i(\tilde{I}_i, I_{-i}) = U_i(\tilde{I}) = 0 > U_i(I)$. This shows that I_i is not a best response to I_{-i}, which is a contradiction. ■

With the above lemmas, we can prove Proposition 1.

Proof of Proposition 1. Let $N \in \mathbb{N}$ be given, and let $M = N - 1$. We first establish that I_M belongs to $\mathcal{A}_N$. It is clear that I_M is symmetric, and since, for all $m \in S_M$,

$$\pi_{mM}^{(M)} \geq \underbrace{\pi_{mm+1}\cdots\pi_{M-1M}}_{M-m \text{ terms}}\underbrace{\pi_{MM}\cdots\pi_{MM}}_{m \text{ terms}} > 0. \tag{39}$$

we obtain that for all $m, m' \in S_M$,

$$\pi_{mm'}^{(2M)} \geq \pi_{mM}^{(M)} \pi_{Mm'}^{(M)} > 0; \tag{40}$$

that is, the Markov Chain induced by I_M is irreducible.

Finally, to show that I_M is a subgame perfect equilibrium, we proceed as follows: Let $i = 1$, and $s \in S_M$. Given I_2^M, player 1 faces a Markovian decision problem, where the state space is $\tilde{S} = \Omega \times S$, the initial state is either $(1, s)$ or $(2, s)$, each with 1/2 probability, and the transition probabilities $q_{l,j}(a)$ are as follows: let $l = (\omega, m)$; if either $\omega = 2$ or $\omega = 1$, and $a = P$, then for all $m \in S_M$,

$$q_{l,j}(a) = \begin{cases} \frac{1}{2} & \text{if} \quad j = (1, T_M(\omega, m, B_\omega(m))) \\ \frac{1}{2} & \text{if} \quad j = (2, T_M(\omega, m, B_\omega(m))) \\ 0 & \text{otherwise.} \end{cases} \tag{41}$$

If $\omega = 1$, and $a = NP$ then for all $m \in S_M$,

$$q_{l,j}(a) = \begin{cases} \frac{1}{2} & \text{if} \quad j = (1, m) \\ \frac{1}{2} & \text{if} \quad j = (2, m) \\ 0 & \text{otherwise.} \end{cases} \tag{42}$$

By Corollary 3.1 in Derman (1970), there exists a finite automaton I_1^* that is a best reply to I_2^M at state s. By Theorem A.1 in Derman (1970) and Abel's Theorem (see DePree and Swartz (1988, Theorem 11.17, p. 135)) we have that $U_1(I_M, s) = \lim_{\delta \to 1} U_1^\delta(I_M, s)$, and also that $U_1((I_1^*, I_2^M), s) = \lim_{\delta \to 1} U_1^\delta((I_1^*, I_2^M), s)$. By Lemma 3, it follows that $U_1(I_M, s) \geq U_1((I_1^*, I_2^M), s)$. Hence, I_1^M is a best reply to I_2^M at state s. Using an analogous argument for $i = 2$, we conclude that I_M is a subgame perfect equilibrium.

It is left to show that $U_1(I_M) + U_2(I_M) \geq U_1(I) + U_2(I)$ for all $I \in \mathcal{A}_N$. For each $I \in \mathcal{A}_N$, recall that by Lemma 4 we have

$$U_i(I) = \frac{1}{2\,|S|} \sum_{s \in S} \sum_{\omega \in \Omega} u_i(\omega, B(s)). \tag{43}$$

Let $S_P = \{s \in S : B_2(s) = P\}$ and $S_{NP} = \{s \in S : B_2(s) = NP\}$. By symmetry, $|S_P| = |\{s \in S : B_1(s) = P\}|$ and $|S_{NP}| = |\{s \in S : B_1(s) = NP\}|$. Then, we obtain

$$\begin{aligned} U_1(I) =& \frac{1}{2\,|S|}\left(\sum_{s\in S} u_1(1, B(s)) + \sum_{s\in S} u_1(2, B(s))\right) \\ =& \frac{1}{2\,|S|}\left(-d\,|S_P| + u\,|S_P|\right) = \frac{|S_P|}{|S|}\frac{u-d}{2}. \end{aligned} \tag{44}$$

Because I is a subgame perfect equilibrium, $|S_{NP}| \geq 1$, and so $|S_P| = |S| - |S_{NP}| \leq |S| - 1$. Hence, it follows that,

$$U_1(I) \leq \frac{u-d}{2}\left(1 - \frac{1}{|S|}\right) \leq \frac{u-d}{2}\left(1 - \frac{1}{M+1}\right) = U_1(I_M). \tag{45}$$

Since, by symmetry, $U_2(I_M) = U_1(I_M) \geq U_1(I) = U_2(I)$, the result follows. ■

A.2 Proof of Proposition 2

Since $Y_n(\omega) = 2$ for all $n \in \mathbb{N}$ and $\omega \in \Omega$, we have that $\text{cov}(c_t^1, y_{t-k}^1 | Y_{t-k}) = \text{cov}(c_t^1, y_{t-k}^1)$ for all $0 \leq k < t$.

For convenience, let μ^n denote the uniform measure on Ω^n, i.e., $\mu^n(\omega^n) = 2^{-n}$ for all $\omega^n \in \Omega^n$.

By definition,

$$\text{cov}(c_t^1, y_{t-k}^1) = \sum_{\omega^t \in \Omega^t} \frac{1}{2^t}(c_t^1(\omega^t) - \bar{c}_t^1)(y_{t-k}^1(\omega^t) - \bar{y}_{t-k}^1) \tag{46}$$

where $\bar{y}_{t-k}^1 = 1$. If $q_t(s)$ denotes the probability that in period t the state is $s \in \{0, \ldots, M\}$, then, $\bar{c}_t^1 = 1 + (q_t(M) - q_t(0))/2$. Since $\lim_{t\to\infty} q_t(s) = 1/(M+1)$ for all s, then $\bar{c}_t^1 \to 1$. Hence, we may compute $\text{cov}(c_t^1, y_{t-k}^1)$ using $\bar{c}^1 = 1$ instead of $\bar{c}_t^1$, by considering t sufficiently large.

Let $\sigma_1(s)$ denote the probability that the state s_t is s and $\omega_{t-k} = 1$; similarly, let $\sigma_2(s)$ denote the probability that the state s_t is s when $\omega_{t-k} = 2$. We have that $\sigma_i(s) = \mu^t(\{\omega^t \in \Omega^t : \omega_{t-k} = i \text{ and } s_t(\omega^{t-1}) = s\})$ for $i = 1, 2$.

If $\omega_t = 1$, $s_t = M$ and $\omega_{t-k} = 1$ then $y^1_{t-k} = 2$, $c^1_t = 2$ and so

$$(c^1_t(\omega^t) - \bar{c}^1)(y^1_{t-k}(\omega^t) - \bar{y}^1_{t-k}) = 1.$$

Similarly, if $\omega_t = 1$, $s_t = M$ and $\omega_{t-k} = 2$ then $y^1_{t-k} = 0$, $c^1_t = 2$ and so

$$(c^1_t(\omega^t) - \bar{c}^1)(y^1_{t-k}(\omega^t) - \bar{y}^1_{t-k}) = -1.$$

Given $\omega_t = 1$, in all remaining cases we have

$$(c^1_t(\omega^t) - \bar{c}^1)(y^1_{t-k}(\omega^t) - \bar{y}^1_{t-k}) = 0,$$

since $c^1_t(\omega^t) = 1 = \bar{c}^1$.

For the case $\omega_t = 2$ we obtain

$$(c^1_t(\omega^t) - \bar{c}^1)(y^1_{t-k}(\omega^t) - \bar{y}^1_{t-k}) = \begin{cases} -1 & \text{if } s_t = 0 \text{ and } \omega_{t-k} = 1, \\ 1 & \text{if } s_t = 0 \text{ and } \omega_{t-k} = 2, \\ 0 & \text{otherwise.} \end{cases} \tag{47}$$

Then,

$$\text{cov}(c^1_t, y^1_{t-k}) = \frac{\sigma_1(M) - \sigma_2(M)}{2} + \frac{\sigma_2(0) - \sigma_1(0)}{2}. \tag{48}$$

Thus, it is enough to show that $\sigma_1(M) \geq \sigma_2(M)$ and $\sigma_2(0) > \sigma_1(0)$.

For any $s \in S_M$ and $\omega^{k-2} = (\omega_{t-k+1}, \ldots, \omega_{t-1})$ let $\{s^i_j(s, \omega^{k-2})\}^t_{j=t-k+1}$ denote the sequence of states resulting from having state s in period $t-k$ and $\omega_{t-k} = i$ for $i = 1, 2$. Using the definition of T_M, one easily sees that $s^1_j(s, \omega^{k-2}) \geq s^2_j(s, \omega^{k-2})$ for any j, s and ω^{k-2}. So, given $s \in S_M$, if ω^{k-2} is such that $s^1_t(s, \omega^{k-2}) = 0$, then $s^2_t(s, \omega^{k-2}) = 0$. Similarly, if ω^{k-2} is such that $s^2_t(s, \omega^{k-2}) = M$, then $s^1_t(s, \omega^{k-2}) = M$. This implies that $\sigma_1(M) \geq \sigma_2(M)$ and $\sigma_2(0) \geq \sigma_1(0)$. Hence, it is enough to show that there exists $s \in S_M$, possible to reach at period $t - k$ starting from $\bar{s}_M$, for which the following holds: there exists ω^{k-2} such that $s^1_t(s, \omega^{k-2}) > 0$ and $s^2_t(s, \omega^{k-2}) = 0$ since this implies $\sigma_2(0) > \sigma_1(0)$.

Let $t \geq \alpha$ and $0 \leq k \leq t-\alpha$, i.e., $t-k \geq \alpha$. By choosing $\alpha > 0$ sufficiently large, any state $s \in S_M$ can be reached at period $t-k$ starting from $\bar{s}_M$: simply take $\omega = 2$ in the beginning in order to get to $s = 0$, then continue with $\omega = 2$ to keep $s = 0$ until period $t-k-s-1$, and take $\omega = 1$ from period $t-k-s$ until period $t-k-1$. If k is odd, let $s_{t-k} = 0$ and $\omega^{k-2} = (1, 2, 1, 2, \ldots)$. This will produce $s_t^1(s, \omega^{k-2}) = 1$ and $s_t^2(s, \omega^{k-2}) = 0$. If k is even, let $s_{t-k} = 1$ and $\omega^{k-2} = (2, 1, 2, 1, \ldots)$. Again, this will produce $s_t^1(s, \omega^{k-2}) = 1$ and $s_t^2(s, \omega^{k-2}) = 0$. This completes the proof that $\mathrm{cov}(c_t^1, y_{t-k}^1 | Y_{t-k}) > 0$.

Finally, we show that $\mathrm{cov}(\theta_t, \theta_{t-1}) < 0$ if t is sufficiently large.

Since $t_{\omega,n}^i \in \{0, 1\}$ for all $n \in \mathbb{N}$ and $\omega \in \Omega$, then

$$\begin{aligned} \bar{\theta}_t &= \mu^t(\{\omega^t : s_t(\omega^t) > 0 \text{ and } \omega_t = 2\}) - \mu^t(\{\omega^t : s_t(\omega^t) < M \text{ and } \omega_t = 1\}) \\ &= \frac{1 - q_t(0) - 1 + q_t(M)}{2} = \frac{q_t(M) - q_t(0)}{2}. \end{aligned} \tag{49}$$

Letting $\bar{\theta} = 0$, it follows that $\bar{\theta}_t$ converges to $\bar{\theta}$ and so we can use $\bar{\theta}$ to compute $\mathrm{cov}(\theta_t, \theta_{t-1})$ by considering t sufficiently large.

Note that $(\theta_t(\omega^t) - \bar{\theta})(\theta_{t-1}(\omega^t) - \bar{\theta}) = \theta_t(\omega^t)\theta_{t-1}(\omega^t)$ and that

$$\theta_t(\omega^t)\theta_{t-1}(\omega^t) = \begin{cases} 1 & \text{if } s_t(\omega^t) > 0, s_{t-1}(\omega^t) > 0, \omega_t = 2 \text{ and } \omega_{t-1} = 2, \\ 1 & \text{if } s_t(\omega^t) < M, s_{t-1}(\omega^t) < M, \omega_t = 1 \text{ and } \omega_{t-1} = 1, \\ -1 & \text{if } s_t(\omega^t) > 0, s_{t-1}(\omega^t) < M, \omega_t = 2 \text{ and } \omega_{t-1} = 1, \\ -1 & \text{if } s_t(\omega^t) < M, s_{t-1}(\omega^t) > 0, \omega_t = 1 \text{ and } \omega_{t-1} = 2. \end{cases} \tag{50}$$

Since, if $\omega_{t-1} = 2$, then both $s_t > 0$ and $s_{t-1} > 0$ if and only if $s_{t-1} > 1$, it follows that

$$\begin{aligned} \mu^t(s_t > 0, s_{t-1} > 0, \omega_t = 2, \omega_{t-1} = 2) &= \frac{\mu^t(s_t > 0, s_{t-1} > 0, \omega_{t-1} = 2)}{2} \\ &= \frac{\mu^t(s_{t-1} > 1, \omega_{t-1} = 2)}{2} = \frac{\sum_{s=2}^M q_t(s)}{4}. \end{aligned} \tag{51}$$

Similarly, if $\omega_{t-1} = 1$, then both $s_t < M$ and $s_{t-1} < M$ if and only if $s_{t-1} <$

$M - 1$. Thus, it follows that

$$\mu^t(s_t < M, s_{t-1} < M, \omega_t = 1, \omega_{t-1} = 1) = \frac{\mu^t(s_t < M, s_{t-1} < M, \omega_{t-1} = 1)}{2} = \frac{\mu^t(s_{t-1} < M - 1, \omega_{t-1} = 1)}{2} = \frac{\sum_{s=0}^{M-2} q_t(s)}{4}. \tag{52}$$

If $\omega_{t-1} = 1$, then both $s_t > 0$ and $s_{t-1} < M$ if and only if $s_{t-1} < M$. Thus, it follows that

$$\mu^t(s_t > 0, s_{t-1} < M, \omega_t = 2, \omega_{t-1} = 1) = \frac{\mu^t(s_t > 0, s_{t-1} < M, \omega_{t-1} = 1)}{2} = \frac{\mu^t(s_{t-1} < M, \omega_{t-1} = 1)}{2} = \frac{\sum_{s=0}^{M-1} q_t(s)}{4}. \tag{53}$$

Finally, if $\omega_{t-1} = 2$, then both $s_t < M$ and $s_{t-1} > 0$ if and only if $s_{t-1} > 0$. Thus, it follows that

$$\mu^t(s_t < M, s_{t-1} > 0, \omega_t = 1, \omega_{t-1} = 2) = \frac{\mu^t(s_t < M, s_{t-1} > 0, \omega_{t-1} = 2)}{2} = \frac{\mu^t(s_{t-1} > 0, \omega_{t-1} = 2)}{2} = \frac{\sum_{s=1}^{M} q_t(s)}{4}. \tag{54}$$

It then follows that $\text{cov}(\theta_t, \theta_{t-1})$ converges to

$$-\lim_{t\to\infty} \frac{q_t(1) + q_t(M-1)}{4} = -\frac{1}{2(M+1)}. \tag{55}$$

Hence, if t is sufficiently large, we conclude that $\text{cov}(\theta_t, \theta_{t-1}) < 0$.

References

Abdulkadiroğlu, A., and K. Bagwell (2005): "Trust, Reciprocity and Favors in Cooperative Relationships," mimeo, Columbia University.

Abreu, D. (1988): "On the Theory of Infinitely Repeated Games with Discounting," *Econometrica*, 56, 383–396.

Aumann, R. (1981): "Survey of Repeated Games," in *Essays in Game Theory and Mathematical Economics in Honor of Oskar Morgenstern*. Bibliographisches Institut, Mannheim.

Billingsley, P. (1995): *Probability and Measure*. Wiley, New York.

DePree, J., and C. Swartz (1988): *Introduction to Real Analysis*. Wiley, New York.

Derman, C. (1970): *Finite State Markovian Decision Processes*. Academic Press, New York.

Ellison, G. (1994): "Cooperation in the Prisoner's Dilemma with Anonymous Random Matching," *Review of Economic Studies*, 61, 567–588.

Fafchamps, M., and S. Lund (2001): "Risk-Sharing Networks in Rural Philippines," mimeo, University of Oxford.

——— (2003): "Risk-Sharing Networks in Rural Philippines," *Journal of Development Economics*, 71, 261–287.

Fiske, A. (1992): "The Four Elementary Forms of Sociality: Framework for a Unified Theory of Social Relations," *Psychological Review*, 99, 689–723.

Hauser, C., and H. Hopenhayn (2004): "Trading Favors: Optimal Exchange and Forgiveness," mimeo, University of Rochester and UCLA.

Henrich, J., R. Boyd, S. Bowles, C. Camerer, E. Fehr, H. Gintis, and R. McElreath (2001): "In Search of Homo Economicus: Behavioral Experiments in 15 Small-Scale Societies," *American Economic Review*, 91, 73–78.

Kalai, E. (1990): "Bounded Rationality and Strategic Complexity in Repeated Games," in *Game Theory and Applications*, ed. by A. N. Tatsuro Ichiishi, and Y. Tauman. Academic Press, New York.

Kalai, E., and W. Stanford (1988): "Finite Rationality and Interpersonal Complexity in Repeated Games," *Econometrica*, 56, 397–410.

Kandori, M. (1992): "Social Norms and Community Enforcements," *Review of Economic Studies*, 59, 63–80.

Kocherlakota, N. (1996): "Implications of Efficient Risk Sharing without Commitment," *Review of Economic Studies*, 63, 595–609.

Koeppl, T. (2006): "Differentiability of the Efficient Frontier when Commitment to Risk Sharing is Limited," *Topics in Macroeconomics*, 6, Article 10.

La Ferrara, E. (2003): "Kin Groups and Reciprocity: A Model of Credit Transactions in Ghana," *American Economic Review*, 93, 1730–1751.

Ligon, E., J. Thomas, and T. Worrall (2002): "Informal Insurance Arrangements with Limited Commitment: Theory and Evidence from Village Economies," *Review of Economic Studies*, 69, 209–244.

Mertens, J.-F., and T. Parthasarathy (1987): "Equilibria for Discounted Stochastic Games," CORE Discussion Paper 8750, Université Catholique de Louvain.

Möbius, M. (2001): "Trading Favors," mimeo, Harvard University.

Okuno-Fujiwara, M., and A. Postlewaite (1995): "Social Norms and Random Matching Games," *Games and Economic Behavior*, 9, 79–109.

Rincón-Zapatero, J., and M. Santos (2006): "Differentiability of the Value Function without Interiority Assumptions," mimeo, Universidad Carlos III de Madrid and Miami University.

Rubinstein, A. (1986): "Finite Automata Play the Repeated Prisoner's Dilemma," *Journal of Economic Theory*, 39, 83–96.

Schotter, A. (1981): *The Economic Theory of Social Institutions.* Cambridge University Press, Cambridge.

Selten, R. (1975): "Reexamination of the Perfectness Concept for Equilibrium Points in Extensive Games," *International Journal of Game Theory*, 4, 25–55.

Townsend, R. (1994): "Risk and Insurance in Village India," *Econometrica*, 62, 539–591.