

2006s-17

**The Build-up of
Cooperative Behavior among
Non-cooperative Agents**

Hassan Benchekroun, Ngo Van Long

Série Scientifique
Scientific Series

Montréal
Octobre 2006

© 2006 *Hassan Benchekroun, Ngo Van Long*. Tous droits réservés. *All rights reserved*. Reproduction partielle permise avec citation du document source, incluant la notice ©.
Short sections may be quoted without explicit permission, if full credit, including © notice, is given to the source.



CIRANO

Le CIRANO est un organisme sans but lucratif constitué en vertu de la Loi des compagnies du Québec. Le financement de son infrastructure et de ses activités de recherche provient des cotisations de ses organisations-membres, d'une subvention d'infrastructure du Ministère du Développement économique et régional et de la Recherche, de même que des subventions et mandats obtenus par ses équipes de recherche.

CIRANO is a private non-profit organization incorporated under the Québec Companies Act. Its infrastructure and research activities are funded through fees paid by member organizations, an infrastructure grant from the Ministère du Développement économique et régional et de la Recherche, and grants and research mandates obtained by its research teams.

Les partenaires / The Partners

Partenaire majeur

Ministère du Développement économique,
de l'Innovation et de l'Exportation

Entreprises partenaires

Alcan inc.
Banque du Canada
Banque Laurentienne du Canada
Banque Nationale du Canada
Banque Royale du Canada
Bell Canada
BMO Groupe financier
Bombardier
Bourse de Montréal
Caisse de dépôt et placement du Québec
Fédération des caisses Desjardins du Québec
Gaz Métro
Hydro-Québec
Pratt & Whitney Canada
Raymond Chabot Grant Thornton

Autres partenaires gouvernementaux

Industrie Canada
Ministère des Finances du Québec
Ville de Montréal

Partenaires universitaires

École Polytechnique de Montréal
HEC Montréal
McGill University
Université Concordia
Université de Montréal
Université de Sherbrooke
Université du Québec
Université du Québec à Montréal
Université Laval

Le CIRANO collabore avec de nombreux centres et chaires de recherche universitaires dont on peut consulter la liste sur son site web.

Les cahiers de la série scientifique (CS) visent à rendre accessibles des résultats de recherche effectuée au CIRANO afin de susciter échanges et commentaires. Ces cahiers sont écrits dans le style des publications scientifiques. Les idées et les opinions émises sont sous l'unique responsabilité des auteurs et ne représentent pas nécessairement les positions du CIRANO ou de ses partenaires.

This paper presents research carried out at CIRANO and aims at encouraging discussion and comment. The observations and viewpoints expressed are the sole responsibility of the authors. They do not necessarily represent positions of CIRANO or its partners.

The Build-up of Cooperative Behavior among Non-cooperative Agents

Hassan Benchenkroun^{*}, Ngo Van Long[†]

Résumé / Abstract

Nous développons un modèle théorique dans lequel les individus sont motivés par la satisfaction égoïste que leur procure l'accumulation de biens, mais où le contrat social incite chaque individu à travailler pour les autres afin d'accroître le bien-être collectif. Nous introduisons une variable d'état représentant le stock de capital social, ou « stock de coopération ». Nous démontrons que cette variable peut influencer les actions des agents non-coopératifs. Les agents accumulent le stock de coopération de la société et réussissent à régler de manière progressive le problème du passager clandestin pour un jeu de contributions privées dans un bien public. Nous supposons qu'il n'existe pas de stratégies de pénalité, de sentiment de culpabilité chez les individus et que chaque agent est rationnel.

Mots clés : biens public, confiance, règle de conduite, stock de coopération

We develop a theoretical model in which each individual is, in some ultimate sense, motivated by purely egoistic satisfaction derived from the goods accruing to him, but there is an implicit social contract such that each performs duties for the others in a way that enhances the satisfaction of all. We introduce a state variable that acts as a proxy for social capital of trustworthiness and that we call the stock of cooperation. We show that noncooperative agents might condition their action on this state variable. Agents build-up the society's stock of cooperation and gradually overcome the free riding problem in a game of private contribution to a public good. We assume that there are neither penalties in the sense of trigger strategies, nor guilt and that each individual is rational.

Keywords: behavior rule, public goods, stock of cooperation, trust

Codes JEL : C73, H41, D60

^{*} Department of Economics and CIREQ, McGill University, 855 Sherbrooke St West, Montreal, H3A 2T7, Canada. Email: hassan.benchekroun@mcgill.ca

[†] CIRANO and CIREQ, Department of Economics, McGill University, 855 Sherbrooke Street West, Montreal, H3A 2T7, Canada. Email: ngo.long@mcgill.ca

1 Introduction

According to Arrow (1972, 1974) the role of trust in economic behavior cannot be exaggerated. He pointed out that “virtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time. It can be plausibly argued that much of the economic backwardness in the world can be explained by lack of mutual confidence” (Arrow, 1972, p 357). Subsequent to Arrow’s pioneering work, a number of authors have developed theoretical, empirical, and experimental models to address the importance of trust and reciprocity in economic life.

Experimental studies of trust indicate that certain behavior is based on trust even when all agents face one-shot games. Berg, Dickhaut and McCabe (1995, p. 123) reported the following experiment. “Subjects in room A decide how much of their \$10 show-up fee to send to an anonymous counterpart in room B. Subjects were informed that each dollar sent would triple by the time it reaches room B. Subjects in room B then decide how much of the tripled money to keep and how much to send back to their respective counterparts. The unique Nash equilibrium prediction for this game is to send zero money. This prediction is rejected in our first (no history) treatment where 30 of 32 room A subjects sent money.” They also reported that room B subjects do reciprocate¹.

Zak and Knack (2001) build a general equilibrium growth model where agents are heterogeneous and face a moral hazard problem. In each period, to access the capital market consumers must use the service of investment brokers who have more information about the return on investment than their clients. Consumers choose the extent to which they trust their broker by choosing the amount of spending on verification of their broker’s trustworthiness. The authors show that trust is linked to social, economic and institutional environments and that there exists a “low-trust poverty trap”. In societies where trust is low, savings cannot sustain positive output growth. Their results are supported by an empirical investigation of a cross-section of countries. The model provides an explanation of

¹A double-blind procedure was used to guarantee complete anonymity with respect to other subjects and the experimenter. (p.127.)

the empirical finding of Knack and Keefer (1997) that higher trust is conducive to growth for a sample of 29 market economies.

Trust and reciprocity are as important facts of economic life as the price system. Voluntary donations to a blood bank, for example, could be explained in terms of trust. Richard Titmuss (1971, p. 239) argued that “in not asking for or expecting any payment of money, these donors signify their belief in the willingness of other men to act altruistically in the future, and to combine together to make a gift freely available should they have a need for it. By expressing confidence in the behavior of future unknown strangers, they were denying the Hobbesian thesis that men are devoid of any distinctive moral sense.”

In trying to explain the motives of blood donors and other gift givers, Arrow (1972, p. 348) advanced three alternative hypotheses:

(H1) The welfare of each individual depends positively on his own satisfaction, and on the satisfactions obtained by others.

(H2) The welfare of each individual depends not only on the utilities of himself and others, but also on his contributions to the utilities of others.

(H3) Each individual is, in some ultimate sense, motivated by purely egoistic satisfaction derived from the goods accruing to him, but there is an implicit social contract such that each performs duties for the other in a way calculated to enhance the satisfaction of all.

It has often been pointed out that a methodological problem with hypotheses (1) and (2) is that it is “too easy” to explain a behavior by assuming that such behavior contributes directly to utility. For this reason, it is the third hypothesis that interests us most, and this paper formulates a theoretical model that elaborates on the idea that self-interested individuals have an incentive to “do good”, provided that they see this incentive in other equally self-interested individuals with whom they interact.

There are two important issues concerning hypothesis H3. First, what is the origin of such tacit social contracts? Second, how are such contracts carried out? In this paper, we are not addressing the first issue, though it is worthwhile to note in passing that Arrow (1972, p. 349) did mention the evolutionary mechanism, which, as he pointed out, had been discussed in Kropotkin (1902) and Wynne-Edwards (1962). According to the evolutionary

view, trust emerges because it maximizes genetic fitness. See also more recent work by Güth et al.(1993), and Cosmides and Tooby (1992). Altruism, for example, might be propagated by genes (Dawkins, 1976) or by “meme”, a controversial concept introduced by Dawkins (1986).

Kurzban and Houser (2004) conduct an experimental investigation of cooperative types in humans in a public goods game. Their findings support the existence of three types of subjects: cooperative, non-cooperative and reciprocal types who condition their contribution to the public good on beliefs about others’ contribution. However, the authors point out that they "remain agnostic with respect to the very important issue of the correct ultimate explanation for the existence of cooperative types, and how cooperative strategies are stabilized." The authors suggest punishment and the possibility of non participation as two promising directions to answer how cooperation in groups might have evolved. In this paper we offer an alternative answer to the question of the stability or sustainability of cooperative strategies.²

Our purpose is to present a theoretical model showing how behavior based on trust can be self-sustaining: if each agent believes that other agents will "do good", it is in this agent’s best interest to "do good" and adopt a "cooperative" behavior. Note that this is in contrast with approaches that explain cooperation by example-setting behavior, where an agent "does good" to induce the other agents to "do good". In the model presented here, the causality is reversed: given a "cooperative" behavior of the other agents, an agent chooses to adopt a "cooperative" behavior. This is consistent with the experimental findings in Kurzban and Houser (2004) where the majority type is the reciprocal type. In their experiment the agent is given a chance to revise his contribution to the public good after getting the information.

It is important to point out that we do not rely on repeated games with trigger strategies and punishment by reverting to the one-shot non-cooperative equilibrium, nor on reputation

²Kosfeld et al. (2005) give an alternative biological explanation of trust among humans: they established a link between the level of trust and the level of oxytocin, a neuropeptide that was shown to play a key role in positive social interactions among non-human mammals. They study the effect of intranasal administration of oxytocin on individual’s decisions in a trust game. The experiments show that oxytocin increases the investors’ trusting behavior, although it does not appear to have an impact on the trustees’ behavior.

building in a sequence of games (Kreps, 1990). Instead we model the conditioning of behavior on a “social history” of trust or cooperation. In general, a social history need not be the history of plays of current players. It is possible that a history of behavior of past players (who have exited the game) may influence the behavior of new players³. It has been observed that, in experiments, the provision of social history of past players does influence the behavior of current players (Berg et al., 1995, section 4).

Instead of a complicated social history that records the behavior of each agent at all points of time, we postulate a very simple social history that we call the stock of cooperation. It is simply a rough indicator of past behavior. We suppose that the stock of cooperation builds up when agents "behave cooperatively" and that it is subject to “decay” (people easily forget things that happened a long time ago). We assume that there are no penalties in the sense of trigger strategies, nor guilt.

Technically, in our model a social history of trustworthiness or the stock of cooperation is a state variable that is subject to decay. We show that noncooperative agents might condition their action on this state variable, and gradually (and partially) overcome the free riding problem in a game of private contribution to a public good. An implication of our model is that two societies with identical endowments, preferences and technology may end up with two different levels of public good (and welfare), if in one society agents condition their action on the intangible stock of cooperation, and build up the stock, while agents in the other society disregard the possible relevance of such a stock. Both sets of agents are rational, and coordination in the first society is only tacit.

We consider a repeated game of contribution to a public good. As expected, the constant contribution level (that of a static Nash equilibrium) constitutes a Nash equilibrium level of contributions of the infinitely repeated game. A novel feature of our model is, without resorting to trigger strategies, the existence of a continuum of contribution levels that can be sustained as the steady state of a Markov Perfect Nash Equilibrium (MPNE). These equilibria exhibit two important features. First, we show that the highest sustainable level

³A related concept is social norms, which we do not model here. Violations of norms result in sanctions, which can be external (ostracism) or internal (guilt).

of contribution is strictly between the constant contribution level of a static Nash equilibrium and the socially optimal contribution level. Second, the contribution level of the repeated static game changes over time. Agents gradually adjust their contributions, and eventually achieve a (steady state) sustainable level. These two features contrast with the properties of the equilibria obtained in the repeated game literature, where agents can use strategies that allow for punishment. In that literature it is shown that (i) different contribution levels, including the socially desirable contribution level, can be sustained when agents use trigger strategies and (ii) the equilibrium outcome is such that level of contribution to the public good is constant. The dynamics of trust in an a priori stationary environment or the gradual building of cooperation through time that we show, is an important contribution of this paper. Other models that generates a dynamic of cooperation typically require the existence of at least two types of agents, a cooperative agent and a selfish agent, where the proportion of agents' type follows an evolutionary process described by a replicator dynamics. For example Carpenter (2004) presents a model where agents can either choose a cooperative level of contribution to a public goods game or a level that corresponds to a free rider's contribution. The proportion of free riders is assumed to follow a replicator dynamics and in the long-run all the population becomes free riders. This prediction is partially supported by the results of some laboratory experiments on public goods games (see e.g. Friedman (1996), Carpenter and Matthews (2005), Ledyard (1995), Miller and Andreoni (1991)). Carpenter (2004) shows that conformity increases the growth rate of free riding. In our model agents are homogenous and all selfish. The equilibria exhibited in our paper also have the following feature observed in experiments: starting from high levels of cooperation, the level of cooperation will decrease over time to a steady state level, as agents interact. On the other hand starting from low levels of cooperation the level of cooperation will increase to its steady state level.

Before proceeding, we would like to draw attention to a related literature on dynamic contribution to public goods. This literature assumes that there is a growing physical stock of public good that enters the utility function. Fershtman and Nitzan (1991) modelled public good as a stock that grows with additional contributions. They assumed agents condition

their additional contributions on the current level of the stock of the public good. They showed that the free riding problem is worse, compared with the case where agents are able to commit to a time path of contributions. Wirl (1996) showed that if non-linear strategies are admitted, then the outcome can be better than that predicted by Fershtman and Nitzan. Itaya and Shimomura (2001) obtained results similar to that of Wirl, but in a more general setting. Marx and Matthews (2000) introduced imperfect information and focussed on Bayesian equilibria. In all these models, the public good is a stock that grows over time. In contrast, we model the public good as a flow, so that the only stock is the intangible stock of cooperation (which does not exist in the models mentioned above). In differential game theory it is well established that even linear quadratic games may admit a continuum of non-linear equilibrium strategies. This was shown in a dynamic oligopoly competition with sticky prices (Tsutsui and Mino (1990)) and in a transboundary pollution game (Dockner and Long (1993)). In these models however the state variable entered directly in the utility of each agent. In our model, this intangible stock is of “no intrinsic value” in that it does not enter the utility function nor the production function.

Two questions could be raised: why would agents condition their behavior on an intangible stock of social history of cooperation, and not on any other variable, such as the weather? And how do people coordinate such conditioning? While these questions are beyond the scope of this paper, we would like to note that perhaps, in determining what is a relevant variable, education and culture play a role. As argued in Bisin and Verdier (2000, 2001), the behavioral patterns of children are acquired through an adaptation and imitation process which depends on their parents’ socialization actions, and on the cultural and social environment in which children live. Carpenter et al. (2004) conduct field experiments in urban slums in Viet Nam and Thailand and measure the impact of demographic factors on trust and cooperation. In both countries the contribution level to a voluntary contribution game is high. However, they find that the correlations of trust to many demographic and associational factors differ significantly across the two countries, evidence of the important role of culture.

2 The Model

There are two final goods: a private good, and a public good, the supply of which is denoted by G . (Here G is a flow, not a stock.) Both goods are produced using labor (measured in unit of effort). Examples of private contributions to a public good include dyke maintenance, charity, monitoring and denouncing corrupt behavior of government officials, etc. There are n identical agents. If agent i contributes g_i units of labor as input to the public good, he is left with $x_i = x(g_i)$ units of private good. The function $x(g_i)$ can be thought of as the upper boundary of the individual's production possibilities set. Assume $x'(g_i) < 0$ for all $g_i \geq 0$. The maximum amount of private goods that an individual can produce depends negatively on his contribution to public good. We do not impose further restrictions on $x(g)$ in order to simplify the exposition.

Each person has the utility function

$$u_i = U(G, x_i)$$

where G is the quantity of the public good. Assume $U_G > 0$ and $U_x > 0$.

Let G_{-i} denote the total contribution of all agents other than agent i . Then

$$G = G_{-i} + g_i$$

In this model, there is no physical stock of productive input. (This assumption is made so that we can focus on the role of "social trust" or the stock of cooperation which is modelled as a stock, with its own law of motion.)

Individuals live for ever. Agent i ($i = 1, 2, \dots, n$) maximizes his life-time utility, as described below.

PROBLEM P_i^1 : Choose the time path $g_i(t)$ to maximize

$$\int_0^{\infty} U(g_i(t) + G_{-i}(t), x(g_i(t))) e^{-rt} dt$$

2.1 Static Nash equilibrium

Suppose agent i believes that $G_{-i}(t) = \bar{G}_{-i}$ (a positive constant). Then the solution of PROBLEM P_i^1 is to set $g_i(t) = \bar{g}_i$, where \bar{g}_i maximizes

$$U(g_i + G_{-i}, x(g_i))$$

We assume an interior optimum for agent i 's problem. Then \bar{g}_i satisfies the first order condition:

$$U_G(\bar{g}_i + \bar{G}_{-i}, x(\bar{g}_i)) + U_x(\bar{g}_i + \bar{G}_{-i}, x(\bar{g}_i))x'(\bar{g}_i) = 0 \quad (1)$$

We assume that the following second order condition is satisfied

$$J = U_{GG} + 2U_{Gx}x' + U_{xx}(x')^2 + U_x x'' < 0 \quad (2)$$

All the functions ($U_{GG}, U_{Gx}, U_{xx}, x'$ and x'') in condition (2) are evaluated at the interior equilibrium.

The first order condition defines the reaction function

$$\bar{g}_i = R(\bar{G}_{-i})$$

The slope of the reaction function is

$$\frac{d\bar{g}_i}{d\bar{G}_{-i}} = - \left(\frac{1}{J} \right) [U_{GG} + x'(\bar{g}_i)U_{xG}]$$

This slope is negative if the term inside the square bracket is negative. (In that case, \bar{g}_i is said to be a strategic substitute for \bar{G}_{-i}). It is useful, though not essential, to assume the strategic substitute property:

Assumption A: (Strategic substitutes)

We assume that U and x are such that

$$U_{GG} + x'U_{xG} < 0 \text{ for all } g \geq 0$$

A symmetric static Nash equilibrium is a vector $\mathbf{g}^N = (g^N, \dots, g^N)$ such that

$$g^N = R(ng^N)$$

This equilibrium is characterized by the equality between each individual's marginal rate of substitution and his marginal rate of transformation:

$$MRS = \frac{U_G(ng^N, x(g^N))}{U_x(ng^N, x(g^N))} = -x'(g^N) = MRT \quad (3)$$

Clearly, there exists a unique symmetric Nash equilibrium if the function $-x'(g)$ is upward sloping, with $x'(0) = 0$ and the function $U_G(ng^N, x(g^N))/U_x(ng^N, x(g^N))$ is downward sloping (a sufficient set of conditions for this is $U_{xG} \geq 0$, $U_{GG} \leq 0$ and $U_{xx} \leq 0$.)

Under standard assumptions, it is well known that the symmetric static Nash equilibrium implies the underprovision of the public good (as compared to the social optimum)⁴. The social optimum is obtained by maximizing the utility of the representative individual with respect to g :

$$\max_g U(ng, x(g))$$

Let g^{so} denote the social optimal solution. The first order condition for the social optimum is

$$nU_G(ng^{so}, x(g^{so})) + U_x(ng^{so}, x(g^{so}))x'(g^{so}) = 0$$

This condition yields the well-known Samuelsonian rule, namely the MRT must be equated

⁴See, for example, Cornes and Sandler (1986).

to the sum of MRS across all agents:

$$\sum MRS_i = \frac{nU_G(ng^{so}, x(g^{so}))}{U_x(ng^{so}, x(g^{so}))} = -x'(g^{so}) = MRT \quad (4)$$

We show in Appendix 1 that, under Assumption A⁵, the social optimum g^{so} exceeds g^N .

2.2 Introducing the stock of cooperation

Let us now introduce a state variable $S(t)$ which is a summary of the social history and that we call the stock of cooperation. We assume that the rate of change of the variable $S(t)$ is given by

$$\dot{S}(t) = G(t) - ng^N - \delta S(t) \quad (5)$$

where $\delta > 0$ is the rate of decay of the stock of cooperation. It follows that if $S(0) = 0$ and $G(t) = ng^N$ always, then $S(t)$ will remain zero: no social capital is built-up. On the other hand, as soon as $G(t)$ exceeds ng^N , $S(t)$ will become positive. Starting with $S(0) = 0$, if agents never contribute less than the static Nash equilibrium level, $S(t)$ will never become negative. When all agents contribute to the public good at the socially desirable rate g^{so} the resulting steady-state stock of cooperation is $S_\infty^{so} = n\frac{g^{so}-g^N}{\delta} > 0$. We note that if the game starts with a large level of cooperation stock S_0 , then even if $G(t) - ng^N > 0$ the stock of cooperation may decrease. This is consistent with observations from in experiments that the provision of social history of past players does influence the behavior of current players (Berg et al., 1995, section 4) and that when the level of cooperation at the beginning of experiments can be ‘high’, it diminishes as the experiment proceeds (Ledyard (1995) and the citations therein, see also Fehr and Gächter (2000)).

It is important to note that the variable $S(t)$ is not an argument of the utility function U (recall that U depends only on G and x). Nor is $S(t)$ a stock from which a good (public or private) can be extracted. So, we say $S(t)$ has “no intrinsic value”. We assume that the

⁵Assumption A is satisfied, for example, when $U_{GG} < 0$, $U_{Gx} \geq 0$ and $x' < 0$.

law of motion (5) is known to all individuals. They can observe the current $G(t)$.

Even though it has no intrinsic value, we argue below that $S(t)$ can indirectly influence welfare of individuals, if each individual i believes that all other individuals $j \in N$ follow a “rule of behavior” that conditions their public good contribution $g_j(t)$ on the level of the stock of cooperation $S(t)$:

$$g_j(t) = \phi_j(S(t)) \text{ for } j \neq i \quad (6)$$

Consider the problem facing individual i if all agents $j \neq i$ adopt behavior rules of the type specified in equation (6). For simplicity, assume

$$\phi_j(S) = \tilde{\phi}(S) \text{ for all } j \neq i$$

Given this rule of behavior, it follows that

$$G_{-i}(t) = m\tilde{\phi}(S(t)) \text{ where } m = n - 1$$

Thus individual i faces a standard optimal control problem involving a single state variable, S , and a control variable, g_i . The individual i determines $g_i(t)$ that solves

PROBLEM P_i^2 :

$$\max \int_0^\infty U(g_i(t) + m\tilde{\phi}(S(t)), x(g_i(t)))e^{-rt} dt$$

subject to

$$\dot{S} = g_i + m\tilde{\phi}(S) - ng^N - \delta S \text{ with } S(0) = 0. \quad (7)$$

Since this is an autonomous problem (i.e. time appears explicitly only in the term e^{-rt}) and since the time horizon is infinite, the solution of this problem can be represented in the feedback form with $g_i(t) = \phi_i(S(t))$. The set of strategies we consider is the set of Markovian strategies. A Markovian Nash equilibrium is a profile of Markovian strategies, one for each player, that are best replies to each others.

We seek to characterize Markovian Nash equilibria for this differential game which

involves n players. This can be done either by using either the Hamilton-Jacobi-Bellman equation, or by solving the optimal control problem of the representative agent, and appealing to symmetry. We adopt the second approach, and show in Appendix 2 that each solution $\phi(S)$ to the following differential equation

$$\phi'(S) = \frac{(r + \delta)F(\phi(S))}{mU_G + mF(\phi(S)) + F'(\phi(S))(n\phi(S) - ng^N - \delta S)} \quad (8)$$

where

$$F(\phi(S)) = -U_G(n\phi(S), x(\phi(S))) - U_x(n\phi(S), x(\phi(S)))x'(\phi(S))$$

represents a potential Markovian Nash equilibrium strategy. Any solution of (8) for which the transversality condition of the optimal control problem P_i^2 (equation (23) in Appendix 2) is met qualifies as a symmetric Markovian Nash equilibrium. Suppose $S(t)$ remains finite, and U is bounded, then if $\phi(S)$ is such that

$$\lim_{t \rightarrow \infty} S(t) = \bar{S}$$

for some finite $\bar{S} > 0$, one can easily verify that the transversality condition is satisfied.

Lemma 1: *The static Nash equilibrium (found in the preceding subsection) is a Markovian Nash equilibrium of the differential game. At that equilibrium,*

$$\phi(S) = g^N \quad \text{for all } S \geq 0$$

Proof: With $\phi(S) = g^N$, we have $\phi'(S) = 0$. The right-hand side of (8) is also zero, because at g^N

$$-U_G(ng^N, x(g^N)) - U_x(ng^N, x(g^N))x'(g^N) = 0 \quad \blacksquare$$

Remark: We would like to point out that while we allow players to condition their action on the level S , choosing a strategy that does not depend S is admissible. In particular, it is feasible to choose the equilibrium constant level of contribution g^N of the non-cooperative static game. The idea of allowing agents to condition on S reflects the fact that in repeated

transactions the level of past cooperation may induce higher trust among agents and a more cooperative behaviour of others. For instance, it is customary in business that acceptable payment methods for long time customers are different and more accommodating than for a new customer (e.g. accepting personal checks or debt acknowledgement instead of requiring certified checks or cash). Similarly, the practice of credit ratings used by loan companies is an example where the main information that matters is the past payment behaviour of a customer, over a certain period of time. The credit history is then summarised by a state variable, i.e. the credit rating. In public good games, there are numerous laboratory experiments that support the prevalence of (partial) cooperation even in one shot games, where the use of punishment is not feasible. The main assumption that we impose in our paper is that the perceived indicator of cooperation and its dynamics as described by equation (5) is the same for all agents. This assumption is made for tractability. We believe that equilibria where agents do condition their actions on their perceived level of cooperation would still exist if each agent had a specific indicator of cooperation with its own dynamics.

Our specification of the dynamics, i.e. equation (5), of the perceived indicator of cooperation is by no means the only one that would produce the desirable outcome. Here we have assumed that the baseline expectation of an agent is g^N and any contribution above g^N builds up the indicator of cooperation. Alternatively, we could have used a different baseline such as g^{so} and have the indicator decrease when the level of contribution is below g^{so} . One can also consider a more sophisticated natural decay of the stock of cooperation. The essential point is that any dynamic specification, such as equation (5), should exhibit two simple features: there exists a perceived level of trust or cooperative behavior by agents if they observe a level of public goods different from the noncooperative level, this perception of cooperation is enhanced the higher the contributions to the public good. Specifications such as equation (5) are supported by results from laboratory experiments of trust games and games of public goods which show that in repeated interactions, as the contributions fall below the cooperative level trust among agents decreases.

Measuring the initial stock of cooperation S and other parameters of the differential equation that describes its dynamics (5), in particular the rate of decay δ , is a critical

issue in empirical applications of this model. Although we do not address this question we would like to mention the existence of surveys that have been used to measure trust between individuals. For example Zak and Knack (2001) use a measure based on data from the World Values Survey (WVS) conducted in three periods (1981, 1990-1 and 1995-6) in several dozen countries (Inglehart et al. 2000). The measure of trust used by Zak and Knack (2001) is the percentage of respondents in a given country that stated that ‘most people can be trusted’ against those who stated the alternative, ‘you can’t be too careful in dealing with people’. We believe that a similar measure of the stock of cooperation can be used, and the data from several years can be used to measure the rate of decay of the stock of cooperation or trust. For certain public goods of a specialized nature, potential contributors are given information about the aggregate history of cooperations. For example, the website wikipedia.org (the 25th most popular website in the world) publishes statistics about daily financial donations (of mostly anonymous donors) to wikipedia, and statistics about recent additions/changes to articles (all of which are by anonymous contributors) to the spread of encyclopedic knowledge. Similarly, alumni of many universities are provided with information about the outcomes of fund raising efforts of previous years. These provide some proxy measures of the stock of cooperation for a given community.

We now investigate the existence of non-constant Markovian Nash equilibria.

2.3 Some properties of non-constant Markovian Nash equilibria

Write (8) as

$$\frac{dg}{dS} = \frac{(r + \delta)F(g)}{mU_G(ng, x(g)) + mF(g) + F'(g)(ng - ng^N - \delta S)} \quad (9)$$

We can trace the integral curves of this first order differential equation in a diagram with S along the horizontal axis and g along the vertical axis. In the diagram, we draw the line $g = g^N + (\delta S/n)$. Along this line, $\dot{S} = 0$. The intersection of this line with an integral curve of the first order differential equation (9) yields a potential steady state. We are particularly interested in finding the integral curves $g = \phi(S)$ that cut this line from above (to ensure

stability of the steady state), i.e. we want $n\phi'(S) < \delta$. At the point where the curve $g = \phi(S)$ cuts the curve $g = g^N + (\delta S/n)$, we have $\dot{S} = 0$, hence

$$\phi'(S) = \frac{(r + \delta) \{-U_G - U_x x'\}}{mU_G(ng, x(g)) + mF(g)} = \frac{(r + \delta) \{-U_G - U_x x'\}}{-mU_x x'}$$

or

$$\phi'(S) = \frac{r + \delta}{m} \left\{ 1 + \frac{U_G}{U_x x'} \right\} \quad (10)$$

Note that if we evaluate the right hand side of (10) at $g = g^N$, we obtain zero, which means that $\phi(S)$ is a horizontal line. This is the static Nash equilibrium. If we evaluate the right hand side of (10) at the social optimum g^{so} , then it is equal to

$$\frac{r + \delta}{m} \left\{ 1 - \frac{1}{n} \right\}$$

and thus

$$n\phi'(S) = (r + \delta)(n - 1)/m = r + \delta > \delta$$

which implies instability of the steady state; i.e. the stock of cooperation is not bounded. We thus obtain the following impossibility result:

Lemma 2: *The socially optimal level of contribution to the public, g^{so} , cannot be supported as a stable steady state level of contribution of a Markovian Nash equilibrium.*

We would like to point out that this result contrasts with the result obtained for standard repeated games with the use of punishment strategies of the trigger type. When trigger punishment strategies are allowed it is well known in the repeated games literature that the socially desirable level of contributions to the public good can be sustained as the outcome of a Nash equilibrium. In our case although the level of contributions to the public good can be larger than the contributions under a "myopic" static Nash equilibrium g^N , the maximum steady-state level of contributions to the public good is some number g^* that is strictly smaller than g^{so} . This is because we restrict the permissible strategies to a set that does not

involve the switching to a triggered punishment mode in case any deviation is observed⁶. Therefore g^{so} cannot be sustained as a stable steady of our differential game.

Although the socially optimum level of contribution to the public good cannot be supported in the long-run, we can state the following

Proposition 1: *There exists $g^* \in (g^N, g^{so})$ such that any level of the contributions to the public good $g \in [g^N, g^*]$ can be sustained as the stable steady state level of contributions to the public good of a Markovian Nash equilibrium.*

Proof: This follows from Lemma 1 and Lemma 2 and the fact the solutions to the first order differential equation (9) is a continuum. Through each point on the segment IS , in the (S, g) space, joining the two points $(0, g^N)$ and (S^{so}, g^{so}) where $S^{so} = \frac{g^{so} - g^N}{\delta/n}$, passes one and only one integral curve of the differential equation (9). Moreover, since $\frac{U_G}{U_{xx'}}$ is a continuous function of g , we can state that the slope of the integral curves at their intersection with the segment IS , given by $\phi'(S) = \frac{r+\delta}{m} \left\{ 1 + \frac{U_G}{U_{xx'}} \right\}$, varies continuously from $\phi'(0) < \frac{\delta}{n}$ to $\phi'(S^{so}) > \frac{\delta}{n}$. Therefore there exists a unique point (g^*, S^*) on the segment IS such that the integral curve to the differential equation that cuts the segment IS at (g^*, S^*) has slope exactly equal to $\frac{\delta}{n}$ at (g^*, S^*) . Moreover each element of the continuum of integral curves to the differential equation (9) that cut the segment IS at $g \in [g^N, g^*]$ has a slope that is strictly smaller than $\frac{\delta}{n}$ (i.e. the stability condition (10) is satisfied) at its intersection with the segment IS . Therefore each of these integral curves characterize a MPNE strategy and any level of the contributions to the public good $g \in [g^N, g^*]$ can be sustained as the stable steady state level of contributions to the public good of a Markovian Nash equilibrium ■

The existence of a continuum of non-linear equilibria in differential games was shown in a dynamic oligopoly competition with sticky prices (Tsutsui and Mino, 1990) and in a transboundary pollution game (Dockner and Long, 1993). In contrast with the model at hand, in these models the state variable entered directly in the utility of each agent. We would like to point out that, as in Tsutsui and Mino (1990) and Dockner and Long (1993), each equilibrium strategy that can support a steady state level of contributions $g^* \in (g^N, g^{so})$

⁶For a discussion of strategies that allow for the use of threats in differential games see Dockner et al. (2000, Chapter 6).

is only locally defined. Note that while Tsutsui and Mino (1990) and Dockner and Long (1993) use a linear quadratic model, our Proposition 1 is established without assuming that the utility function is quadratic. Also, Tsutsui and Mino (1990) and Dockner and Long (1993), use a dynamic programming approach to characterize the equilibria, here we use the maximum principle and exploit the fact that problem is autonomous to express the controls and the co-state variables in a feedback form. Many authors have used the maximum principle approach to obtain Markovian equilibria of differential games (see, for examples, Tornell and Lane, 1999, Tornell and Velasco, 1992).

Another important feature of the equilibria we exhibit is that the contribution to the public good is not constant along the equilibrium path. Agents gradually and smoothly adjust their contribution as the stock of cooperation converges to its steady state level. This also contrasts with the outcome of a repeated game where agents are allowed to use strategies that allow for punishments. That setup does not generate a dynamics of the actions played along an equilibrium, unless one introduces imperfect information and learning about some characteristics of the game as it unfolds (as in Marx and Matthews, 2000). In our model agents are informed about the characteristics of the game and in particular the strategies played by the other agents, yet the level of contribution, and the stock of cooperation, can increase gradually and only through the actions taken over time.

3 Complete analytical solution

We now specify functional forms so that a complete analytical characterization of the equilibria is possible. In particular we characterize a continuum of Markovian Nash equilibria with non-linear strategies $\phi(S)$ that lead to a steady state where everyone contributes more than their static Nash equilibrium contribution level to the public good. We also show that if the discount rate r is small enough then it is possible to achieve approximately the social optimum.

Let the utility function be

$$U(G, x) = EG + \frac{wG^2}{2} + x \quad (11)$$

where $E > 0$ is a constant. (At this stage we do not specify whether w is negative or positive.) Let the transformation curve be

$$x(g) = Y - \frac{cg^2}{2} \text{ where } Y > 0 \text{ and } c > 0.$$

Assume $c - n^2w > 0$. The static Nash equilibrium satisfies $U_G + U_x x' = 0$ which gives

$$g^N = \frac{E}{c - nw}$$

The social optimum satisfies $nU_G + U_x x' = 0$ which gives

$$g^{so} = \frac{nE}{c - n^2w} = \frac{E}{(c/n) - nw} > g^N$$

Thus the social optimum contribution exceeds the static Nash contribution.

3.1 A graphical solution

Equation (9) becomes

$$\frac{dg}{dS} = \frac{[(c - wn)g - E](r + \delta)}{mcg + (c - wn)(ng - ng^N - \delta S)} = \frac{T(g)}{D(g, S)} \quad (12)$$

Let us plot the integral curves of this first order differential equation on the plane (S, g) .

Since the numerator is

$$T(g) = [(c - wn)g - E](r + \delta)$$

we know that $T(g) > 0$ if $g > g^N$, and $T(g) < 0$ if $g < g^N$. The denominator is

$$D(g, S) = mcg + (c - wn)(ng - ng^N - \delta S)$$

Thus $D(g, S) = mcg > 0$ for $g > 0$ along the (possible steady state) line $g = g^N + (\delta/n)S$ denoted SS , and $D(g, S) = 0$ along the line RR defined by

$$g = \frac{\delta S}{\frac{mc}{(c-wn)} + n} + \frac{ng^N}{\frac{mc}{(c-wn)} + n}$$

which is below the line $g = g^N + (\delta/n)S$ for all $S \geq 0$. (See Figure 1).

Thus the horizontal line $g = g^N$ and the line RR divide the positive orthant into four regions, which we label clockwise as regions I , II , III , IV , where regions I and II are above the line RR . The steady state line $g = g^N + (\delta/n)S$ lies inside region II .

Integral curves are positively sloped in regions II and IV and negatively sloped in regions I and III . There exists a unique integral curve that is tangent to the possible steady state line SS and g^* denotes the vertical coordinate of the tangency point. There is a continuum of integral curves in region II that cut the steady state line $g = g^N + (\delta/n)S$ from above. (See Figure 1). Each of these integral curves represents a Markovian Nash equilibrium strategy that leads to a stable steady state level of contributions to the public good that belongs to the interval $[g^N, g^*)$.

There are no linear strategies that lead to a steady state, apart from the constant strategy $\phi(S) = g^N$.

3.2 An analytical characterization

We solve the differential equation (12). Let us consider the inverse of the function $\phi(S)$ and interpret (12) as a differential equation characterizing $S(g)$. We obtain the following "inverted" first order differential equation:

$$\frac{dS}{dg} = \frac{(mc + (c - wn)n)g - \delta(c - wn)S - (c - wn)ng^N}{[(c - wn)g - E](r + \delta)} \quad (13)$$

Let $a = \frac{\delta}{(r+\delta)}$, $b = \frac{(mc+(c-wn)n)}{(r+\delta)(c-wn)}$, $B = \frac{mc}{(r+\delta)(c-wn)}$ and $z = g - g^N$. The equation (13) can be written in the form (see Appendix 3):

$$\frac{dS}{dz} + \frac{aS}{z} = b + \frac{Bg^N}{z} \quad (14)$$

The family of solutions is:

$$S(z) = \frac{bz}{1+a} + \frac{Bg^N}{a} + z^{-a}K \quad (15)$$

For each K , (15) defines an implicit relationship between z (or g) and S . (Note: for $K = 0$ we have a linear solution). Given K , expressing S as a function of g for $g > g^N$, we have

$$S_K(g) = \frac{b(g - g^N)}{1+a} + \frac{Bg^N}{a(1+a)} + K(g - g^N)^{-a} \quad (16)$$

For $K < 0$, and $g > g^N$, $S_K(g)$ is a strictly concave and increasing function, with $\lim_{g \rightarrow g^N} S_K(g) = -\infty$ and $\lim_{g \rightarrow \infty} S_K(g) = \infty$. For $K > 0$, and $g > g^N$, $S_K(g)$ is a strictly convex function, with $\lim_{g \rightarrow g^N} S_K(g) = +\infty$ and $\lim_{g \rightarrow \infty} S_K(g) = +\infty$. Furthermore, $\lim_{g \rightarrow g^N} S'_K(g) = -\infty$ and $\lim_{g \rightarrow \infty} S'_K(g) = b/(1+a)$. We must determine whether the curve $S_K(g)$ intersects the curve $S = n(g - g^N)/\delta$ for $g > g^N$. If an intersection exists, it is a steady state.

Steady states are denoted (with a hat) by $\hat{g} = g^N + (\delta/n)\hat{S}$. They are implicitly determined by

$$\hat{S} = \frac{b}{1+a} \left((\delta/n)\hat{S} \right) + \frac{Bg^N}{a} + \left((\delta/n)\hat{S} \right)^{-a} K$$

Let

$$\eta \equiv \left(\frac{b}{1+a} \right) \left(\frac{\delta}{n} \right) - 1 < 0$$

then

$$\eta \hat{S} + \frac{Bg^N}{a} = -K \left((\delta/n)\hat{S} \right)^{-a} \quad (17)$$

The left-hand side of (17) is linear and decreasing in \hat{S} and is positive for $\hat{S} \in (0, Bg^N/(-\eta a))$ and negative for $\hat{S} > Bg^N/(-\eta a) > 0$. There are three cases: $K > 0$, $K = 0$ and $K < 0$. It can be shown that for $K > 0$ and $K = 0$ the stability condition is not satisfied.

For any $K < 0$, the right-hand side of (17) is a positive, convex, and decreasing function of

\widehat{S} for all $\widehat{S} > 0$. It follows that if the absolute value of K is not too large, the convex curve must intersect the downward sloping straight-line that represents the left-hand side of (17) exactly twice, at values which we denote by S_K^L and S_K^H where

$$Bg^N / (-\eta a) > S_K^H > S_K^L > 0.$$

It is clear that S_K^L is locally stable and S_K^H is unstable. This is because the function $S_K(g)$ is concave, and thus the curve representing it in the space (g, S) (with g measured along the horizontal axis) cuts the line $S = n(g - g^N)/\delta$ from below at the point S_K^L and from above at the point S_K^H .

There is a critical value of K , denoted by K^* such that for all K smaller (larger in absolute value) than K^* the convex and decreasing curve that represents the right-hand side of (17) does not intersect the downward sloping straight-line that represents the left-hand side of (17). If $K = K^*$, we have $S_{K^*}^H = S_{K^*}^L$, that is, the steady state is unique. We will denote this unique steady state by \widehat{S}_{K^*} . Let $g^* = g^N + (\delta/n)\widehat{S}_{K^*}$ be the associated steady state level of contribution to the public good. We show in Appendix 4 that the critical value of K^* and \widehat{S}_{K^*} and g^* are given by

$$\widehat{S}_{K^*} = \frac{Bg^N}{(-\eta)(1+a)} > \left(\frac{ng^N}{\delta} \right) \quad (18)$$

and

$$g^* - g^N = \frac{\delta Bg^N}{n(-\eta)(1+a)} > 0$$

Proposition 1b: *Let $g^* = g^N + \frac{\delta Bg^N}{n(-\eta)(1+a)}$. We have $g^* < g^{so}$ and any level of the contributions to the public good $g \in [g^N, g^*]$ can be sustained as the stable steady state level of contributions to the public good of a Markovian Nash equilibrium.*

Substituting g^* and g^N gives

$$\frac{g^N}{g^*} = \frac{n(r+2\delta)(c-wn) - \delta[m + (c-wn)n]}{n(r+2\delta)(c-wn) - \delta[m + (c-wn)n] + \delta mc}$$

we show that as $r \rightarrow 0$, $g^* \rightarrow g^{so}$

$$\lim_{r \rightarrow 0} \frac{g^N}{g^*} = 1 - \frac{mc}{nc - wn^2} = \frac{c - wn^2}{nc - wn^2}$$

On the other hand

$$\frac{g^N}{g^{so}} = \frac{c - wn^2}{nc - wn^2}$$

Thus

$$\lim_{r \rightarrow 0} g^* = g^{so}$$

Thus, we have proved:

Proposition 2:

If the rate of discount r is close enough to zero, then it is possible to approximate the social optimum.

To summarize, there is maximum level of contribution to the public good g^* (and a corresponding stock of cooperation \widehat{S}_{K^*}) that can be supported as the steady state level of contribution to the public good a Markovian Nash equilibrium. Although g^* is smaller than the socially optimal level of contributions to the public good g^{so} (and $\widehat{S}_{K^*} < S^{so}$) it is larger than g^N , the contribution to the public good under the static Nash equilibrium (and $\widehat{S}_{K^*} > S^N = 0$).

The steady state level of contribution to the public good g^* is supported by a Markovian Nash equilibrium where the equilibrium contribution strategy is an increasing function of the stock of cooperation. When agent i takes the contribution of agent $j \neq i$ as given, he can still influence the amount contributed at each moment by agent j by influencing the stock of cooperation. This feedback effect increases the marginal benefit of the contribution to the public good and the resulting equilibrium level of contributions exceeds the contribution level under the static Nash equilibrium where the feedback effect is absent.

We have shown the existence of a stable steady state level of cooperation S in the case for all negative K such that $K^* < K < 0$. While the value of K^* (corresponding to the ‘most cooperative’ equilibrium) depends on parameters of the model, the selection

of a particular level of K may not depend on the model parameters. The existence of a continuum of equilibria can be naturally followed up by a discussion of the selection process of an equilibrium among the continuum of equilibria. The selection of a specific equilibrium amounts to the selection of a specific value of K . This process is not described in our paper. It can be for example determined in a pre-play phase where agents have cheap talks and agree on a level of K . An alternative device to induce agents to condition their actions on the level of cooperation is the existence and participation of a third party such as government agency or regulator that conditions its actions on the level of the stock of cooperation and announces a specific value of K . Note that for a given negative K in $(K^*, 0)$ the stock of cooperation converges to a positive steady level even when the initial stock of cooperation is zero. This is because at time zero the equilibrium initial level of contribution is above the non-cooperative level which initiates the build up process of cooperation.

4 Concluding remarks

We have shown that if each individual in a society expects that other individuals use a behavior rule that conditions their public good contribution on the level of trust or the stock of cooperation in the society, then he will have an incentive to build up the stock of cooperation. This can result in behavior rules that lead to a superior performance in terms of welfare. However it is only when the rate of discount tends to zero that the steady state of this game can approximate what a social planner would want to achieve.

Our model serves to concretize Arrow's third hypothesis, that each individual is, in some ultimate sense, motivated by purely egoistic satisfaction derived from the goods accruing to him, but there is an implicit social contract such that each performs duties for the other in a way calculated to enhance the satisfaction of all.

It is important to note that in our model the variable that represents social history is not an argument of the utility function nor is it a stock from which a good (public or private) can be extracted. In fact, this variable has "no intrinsic value". A remarkable feature of our model is that a variable of no intrinsic value can influence behavior and improve welfare even

when individuals do not resort to trigger strategies. Formally, this can happen for example when agents hold cheap talks prior to the start of the game and agree on the self-enforcing equilibrium strategy (uniquely defined by a given value of K) that will be followed. The main conditions that need to be met are that K is in $(K^*, 0)$ and that initial stock of cooperation is not too high since the non-linear equilibria are not defined globally.

Appendix 1

Comparing the Samuelsonian rule with the Nash equilibrium condition (3), we can see that, if $(U_{GG} + U_{Gx}x')(h - 1) + J < 0$ for all $h \in [1, n]$, the Nash equilibrium level of private provision of the public good is too low. To prove this, define the function

$$\Omega(g, h) = hU_G(ng, x(g)) + U_x(ng, x(g))x'(g)$$

Then, as is clear from (3) and (4),

$$\Omega(g^N, 1) = 0$$

$$\Omega(g^{so}, n) = 0$$

Now along the curve $\Omega(g, h) = 0$, it holds that

$$\frac{dg}{dh} = -\frac{\frac{\partial \Omega}{\partial h}}{\frac{\partial \Omega}{\partial g}}$$

Since

$$\frac{\partial \Omega}{\partial h} = U_G > 0$$

it follows that $g^{so} > g^N$ if, for all $h \in [1, n]$

$$\frac{\partial \Omega}{\partial g} < 0$$

Now

$$\begin{aligned} \frac{\partial \Omega}{\partial g} &= hU_{GG} + (h + 1)U_{Gx}x' + U_{xx}(x')^2 + U_x x'' \\ &= (h - 1)(U_{GG} + U_{Gx}x') + J \end{aligned}$$

where $J < 0$ by the second order condition, and $U_{GG} + U_{Gx}x' \leq 0$ under Assumption A. This completes the proof. (Note that Assumption A is not a necessary condition for g^{so} to exceed g^N .)

Appendix 2: Derivation of (8)

PROBLEM P_i^2 :

$$\max \int_0^{\infty} U(g_i(t) + m\tilde{\phi}(S(t)), x(g_i(t)))e^{-rt} dt$$

subject to (7).

We use the maximum principle to characterize the solution to this problem. The Hamiltonian function is

$$H = U(g_i + m\tilde{\phi}(S), x(g_i)) + \lambda \left[m\tilde{\phi}(S) + g_i - ng^N - \delta S \right] \quad (19)$$

where λ is the costate variable associated with the stock of social environment.

The first order condition from the maximization of H is:

$$U_G(g_i + m\tilde{\phi}(S), x(g_i)) + U_x(g_i + m\tilde{\phi}(S), x(g_i))x'(g_i) + \lambda = 0 \quad (20)$$

We assume that the following second order condition is satisfied

$$U_{GG} + 2U_{Gx}x'(g_i) + U_{xx}(x')^2 + U_x x''(g_i) < 0 \quad (21)$$

The adjoint equation is

$$\dot{\lambda}(t) = r\lambda(t) - \frac{\partial H}{\partial S} = (r + \delta - m\tilde{\phi}'(S))\lambda(t) - m\tilde{\phi}'(S)U_G \quad (22)$$

and the transversality condition is

$$\lim_{t \rightarrow \infty} e^{-rt}\lambda(t)S(t) = 0 \quad (23)$$

Since this is an autonomous problem (i.e. time appears explicitly only in the term e^{-rt}) and since the time horizon is infinite, the solution of this problem can be represented in the feedback form with $\lambda(t) = \lambda^*(S(t))$ and $g_i(t) = \phi_i(S(t))$. The equations (20) and (22) may be written as

$$\lambda^*(S) = -U_G(\phi_i(S) + m\tilde{\phi}(S), x(\phi_i(S))) - U_x(\phi_i(S) + m\tilde{\phi}(S), x(\phi_i(S)))x'(\phi_i(S)) \quad (24)$$

and

$$\frac{d\lambda^*(S)}{dS} \frac{dS}{dt} = (r + \delta - m\tilde{\phi}'(S))\lambda^*(S) - m\tilde{\phi}'(S)U_G \quad (25)$$

We substitute the term $\frac{dS}{dt}$ from (7). The term $\frac{d\lambda^*(S)}{dS}$ is obtained by differentiating equation (24) with respect to S . We then substitute the result into the left-hand side of (25) and we get the following first order differential equation for $\phi_i(S)$:

$$\begin{aligned} & \left((-U_{GG} - U_{Gx}x')(\phi'_i + m\tilde{\phi}') - U_{Gx}x'\phi'_i - U_{xx}(x')^2\phi'_i \right) \left(\phi_i + m\tilde{\phi} - ng^N - \delta S \right) = \\ & (r + \delta - m\tilde{\phi}') \left(-U_G(\phi_i + m\tilde{\phi}, x(\phi_i)) - U_x(\phi_i + m\tilde{\phi}, x(\phi_i))x'(\phi_i) \right) - m\tilde{\phi}'U_G \end{aligned} \quad (26)$$

Agent i takes $\tilde{\phi}(S)$ (and $\tilde{\phi}'(S)$) as given; his problem consists of solving the first order differential equation (26) to obtain $\phi_i(S)$.

We seek to characterize Markovian Nash equilibria for this differential game which involves n players. Restricting our attention to a symmetric equilibrium, we want to find a function $\phi(S)$ such that when $\phi_i(S) = \tilde{\phi}(S) = \phi(S)$ the equation (26) is satisfied.

Let us define

$$F(\phi(S)) = -U_G(n\phi(S), x(\phi(S))) - U_x(n\phi(S), x(\phi(S)))x'(\phi(S))$$

Then, for a symmetric equilibrium, equation (26) reduces to

$$\begin{aligned} & F'(\phi(S))\phi'(S) \left[\phi_i(S) + m\tilde{\phi}(S) - ng^N - \delta S \right] = \\ & (r + \delta)F(\phi(S)) - mF(\phi(S))\phi'(S) - mU_G\phi'(S) \end{aligned}$$

Hence we obtain the first order differential equation:

$$\phi'(S) = \frac{(r + \delta)F(\phi(S))}{mU_G + mF(\phi(S)) + F'(\phi(S))(n\phi(S) - ng^N - \delta S)} \quad (27)$$

Appendix 3:

The numerator of (13) can be separated to obtain

$$\frac{dS}{dg} = \frac{(mc + (c - wn)n)(g - g^N) + mcg^N}{[(c - wn)g - E](r + \delta)} - \frac{\delta(c - wn)S}{[(c - wn)g - E](r + \delta)}$$

using

$$g^N = \frac{E}{c - nw}$$

gives

$$\frac{dS}{dg} + \frac{\delta(c - wn)S}{(g - g^N)(c - wn)(r + \delta)} = \frac{(g - g^N)(mc + (c - wn)n)}{(g - g^N)(c - wn)(r + \delta)} + \frac{mcg^N}{(g - g^N)(c - wn)(r + \delta)}$$

Substituting $a = \frac{\delta}{(r + \delta)}$, $b = \frac{(mc + (c - wn)n)}{(r + \delta)(c - wn)}$, $B = \frac{mc}{(r + \delta)(c - wn)}$ and $z = g - g^N$ yields (14).

Appendix 4: derivation of K^* , \widehat{S}_{K^*} and g^*

To find K^* , we note that the curve $S_{K^*}(g) = \frac{b(g - g^N)}{1 + a} + \frac{Bg^N}{a} + K^*(g - g^N)^{-a}$ must be tangent to the line $S = n(g - g^N)/\delta$ at the unique steady state value \widehat{S}_{K^*} . Thus

$$S'_{K^*}(g^*) = \frac{b}{1 + a} - aK^* \left[\frac{\delta \widehat{S}_{K^*}}{n} \right]^{-a-1} = \frac{n}{\delta} \quad (28)$$

Hence

$$-\frac{a\delta}{n} K^* \left[\frac{\delta \widehat{S}_{K^*}}{n} \right]^{-a-1} = 1 - \frac{b}{1 + a} \left(\frac{\delta}{n} \right) = -\eta \quad (29)$$

On the other hand, by definition of a steady state,

$$\eta \widehat{S}_{K^*} + \frac{Bg^N}{a} = -K^* \left(\frac{\delta \widehat{S}_{K^*}}{n} \right)^{-a} \quad (30)$$

Using (29) and (30) to eliminate K^*

$$\frac{a\delta}{n} \left[\eta \widehat{S}_{K^*} + \frac{Bg^N}{a} \right] = -\eta \left(\frac{\delta \widehat{S}_{K^*}}{n} \right)$$

Solving for \widehat{S}_{K^*}

$$\widehat{S}_{K^*} = \frac{Bg^N}{(-\eta)(1+a)} > \left(\frac{ng^N}{\delta} \right) \quad (31)$$

Having determined \widehat{S}_{K^*} , we can use (31) and (30) to solve for K^* .

It follows that

$$g^* - g^N = \frac{\delta Bg^N}{n(-\eta)(1+a)} > 0$$

References

- [1] Arrow, K. J., 1972. Gifts and Exchanges. *Philosophy and Public Affairs*, Volume 1, 343-362.
- [2] Arrow, K. J., 1974. *The Limits of Organization*. Norton, New York.
- [3] Berg, J., Dickhaut, J., and McCabe, K., 1995. Trust, Reciprocity, and Social History. *Games and Economic Behavior* 10, 122-142.
- [4] Bisin, A., and Verdier, T., 2000. A Model of Cultural Transmission, Voting, and Political Ideology. *European Journal of Political Economy* 16, 5-29.
- [5] Bisin, A., and Verdier, T., 2001. The Economics of Cultural Transmission and the Dynamics of Preferences. *Journal of Economic Theory* 97, 298-319.
- [6] Carpenter, J. P., 2004. When in Rome: Conformity and the Provision of Public Goods. *Journal of Socio-Economics*, 33(4): 395-408.
- [7] Carpenter, J. P. and Matthews, P., 2005. No Switch Backs: rethinking aspiration-based dynamics in the ultimatum game. *Theory and Decision*, 58(4): 351-385.
- [8] Carpenter, J. P., Daniere A. G., and Takahashi, L.M. 2004. Cooperation, Trust, and Social Capital in Southeast Asian Urban Slums. *Journal of Economic Behavior & Organization* Vol. 55 : 533-551.
- [9] Cornes, R., and Sandler, T., 1986. *The Theory of Externalities, Public Goods, and Club Goods*. Cambridge University Press.

- [10] Cosmides, L., and Tooby, J., 1992. Cognitive Adaptation for Social Exchange. In J. Barkow, L. Cosmides, and J. Tooby (eds.), *The Adapted Mind*, Oxford University Press, New York.
- [11] Dawkins, R., 1976. *The Selfish Gene*. Oxford University Press, Oxford, UK.
- [12] Dawkins, R., 1986. *The Blind Watchmaker*. Longman Science and Technology, Harlow, UK.
- [13] Dockner, E. and Long, N. V., 1993. International Pollution Control: Cooperative versus Non-Cooperative Strategies. *Journal of Environmental Economics and Management*, 24, 13-28.
- [14] Dockner, E., Jorgensen, S., Long, N.V., and Sorger, G., 2000. *Differential Games in Economics and Management Science*. Cambridge University Press, Cambridge, UK.
- [15] Fehr, E., and Gächter, S., 2000. Cooperation and Punishment in Public Goods Experiments. *American Economic Review* 90: 980–994.
- [16] Fershtman, C. and Nitzan, S., 1991. Dynamic Voluntary Provision of Public Goods. *European Economic Review*, 35, 1057-67.
- [17] Friedman, D., 1996. Equilibrium in Evolutionary Games: Some Experimental Results. *Economic Journal*, 106, 1-25.
- [18] Güth, W., Ockenfels, P., and Wendel, M., 1993. Efficiency by Trust in Fairness? Multiperiod Ultimatum Bargaining Experiments with an Increasing Cake. *International Journal of Game Theory* 22, 51-73.
- [19] Güth, W., Ockenfels, P., and Wendel, M., 1997. Cooperation based on trust: An Experimental Investigation. *Journal of Economic Psychology* 8, 15-43.
- [20] Inglehart, R. et al. (2000). *World values surveys and european values surveys, 1981-1984, 1990-1993, and 1995-1997*. ICPSR study No. 2790, Inter-University Consortium and Political and Social Research, Institute for Social Research.

- [21] Itaya, J. and Shimomura, K., 2001. A Dynamic Conjectural Variations Model in the Private Provision of Public Goods: A Differential Game Approach. *Journal of Public Economics*, 81,153-172.
- [22] Knack, S., and Keefer, P., 1997. Does Social Capital Have an Economic Payoff? A Cross-Country Investigation. *Quarterly Journal of Economics* 112 (4), 1251-88.
- [23] Kosfeld, M., Heinrichs, M., Zak, P.J., Fischbacher, U. and Fehr, E., 2005. Oxytocin increases trust in humans. *Nature* 435, 673-676.
- [24] Kreps, D.M., 1990. Corporate Culture and Economic Theory. In J. Alt and K. Shepsle, eds., *Perspective on Positive Political Economy*, Cambridge University Press, Cambridge, UK.
- [25] Kropotkin, P.A., 1902. *Mutual Aid: A Factor of Evolution*. London.
- [26] Kurzban, R. and Houser, D., 2004. Experiments investigating cooperative types in humans: a complement to evolutionary theory and simulations. *Proceedings of the National Academy of Sciences*, vol. 102 (5), 1803-1807.
- [27] Ledyard, J., 1995. Public goods: A survey of experimental research. In J. Kagel and A. Roth, editors, *Handbook of Experimental Game Theory*. Princeton University Press, Princeton.
- [28] Marx, L.M. and Matthews, S.A., 2000. Dynamic Voluntary Contribution to a Public Project. *Review of Economic Studies*, 67, 327-58
- [29] Miller, J. and Andreoni, J., 1991. Can Evolutionary Dynamics Explain Free Riding in Experiments? *Economics Letters*, 36, 9-15.
- [30] Titmuss, R., 1971. *The Gift Relationship: From Human Blood to Social Policy*. London and New York.
- [31] Tornell, A. and Lane, P., 1999. The voracity effect. *American Economic Review* 89, 22-46.

- [32] Tornell, A., and Velasco, A., 1992. The tragedy of the commons and economic growth: why does capital flow from poor to rich countries? *Journal of Political Economy* 100, 1208-1231.
- [33] Tsutsui, S. and Mino, K., 1990. Non-linear Strategies in Dynamic Duopolistic Competition with Sticky Prices. *Journal of Economic Theory*, 52, 136-61
- [34] Wirl, F., 1996. Dynamic Voluntary Provision of Public Goods: Extension for Non-Linear Strategies. *European Journal of Political Economy*, 12, 555-60
- [35] Wynne-Edwards, V.C., 1962. *Animal Dispersion in Relation to Social Behavior*. New York.
- [36] Zak, P.J. and Knack, S., 2001. Trust and Growth. *Economics Journal*, 111, p 295-321.

Figure 1

