Self-regulation and trust under endogenous transaction costs*

PIER LUIGI SACCO

Department of Economics

University of Florence

Via Montebello, 7

I-50123 Firenze

Italy

Phone: 055-293229

Fax: 055-211533

E-mail: sacco@facec.cce.unifi.it

* I am deeply grateful to Marco Sandri for his help with the computer simulations. The usual disclaimer applies.

1. Introduction

The enforcement of cooperation between firms for the achievement of a common endeavor is an issue of increasing relevance in the post-Fordist era of vertical disintegration and subcontracting. As the large scale production of standardized items has become less and less viable in increasingly segmented markets, firms have gradually discovered the virtues of productive flexibility and of fine-tuned differentiation of their product lines. The firms' concern for cooperation is then clear: under vertical disintegration, the establishment of actual coordination and cooperation between partners is crucial for productive efficiency. In addition, the need of frequent product and process innovation makes it very costly and risky for a single firm to develop resource- and time-consuming R&D investment programs on its own, thereby looking for joint ventures with firms with similar or complementary R&D interests. The ability of building and preserving cooperative relationships is therefore going to become (and possibly is already) a key component of a firm's competitive potential. But then the issue becomes enforceability, in that cooperative advantages are typically medium/long-run ones whereas any of the involved firms ordinarily faces a large number of short-run defectionist temptations that might be hard to resist, especially in view of the uncertainty about the eventual outcome of the cooperative deal. In this context, firms have clearly a strong incentive to learn to evaluate the reliability of a potential cooperative partner but have to carry the costs that derive from the gathering and processing of the relevant information.

In this context, transaction costs play an undeniable role. As long as the monitoring of potential partners' reliability is easy and cheap, it is reasonable to expect that most firms will buy it and that eventually reliability will become a social standard of behavior as would-be defectors are systematically detected and kept out of business. Conversely, if monitoring is difficult and costly, most firms will do without it and consequently the viability of cooperation will be at a serious risk. The above line of reasoning, however, sounds disturbingly close to a plain truism insofar as the monitoring costs are thought as exogenous. If on the other hand such costs are exclusively traced back to ‘technical’ factors like the ease of access to, and use of, information databases, a rather immediate implication is that the promotion of social cooperation has to rely essentially upon technical progress in informational highways and in information-processing technologies and that, a fortiori, technologically sophisticated societies should be the most cooperative, an intuition that finds little empirical support. Indeed, many researchers have pointed out that the reliability issue has an intrinsic social dimension, which amounts to say that societies with a comparative level of technical development may display remarkably different propensities to the creation and maintainance of cooperative relationships. A pioneering series of studies stressing the importance of socially systemic factors for the enforceability of single cooperative relationships is the well known collective volume edited by Gambetta (1988), which focuses upon the notion of trust. The link between propensity to cooperation and trust is clear: trusting others means essentially relying upon their responsibility in taking a pro-social course of action in situations where they face defectionist temptations. As a matter of fact, one can cooperate with others in the absence of trust: this is exactly what happens when economic agents monitor each other to learn about their respective reliability; the opposite, however, cannot be true: trust may only emerge as a consequence of a relatively long and successful past record of cooperation that leads economic agents to give up monitoring. The diffusion of trust as a social standard of behavior can therefore be taken as a good observable proxy of a successful and lasting enforcement of cooperation at the social level. In a recent book, Fukuyama (1995) draws a distinction between high- and low-trust societies that is mainly based on sociological factors pointing out that even G7 countries like Italy and France can be characterized by a weak propensity to trust and thus, more fundamentally, to cooperation. Another recent study that is worthy of mention in this respect is Putnam’s (1993) study on Italy showing that, even in a country that is being labelled by other authors as low-trust, there are substantial regional differences as to cooperative reliability and pro-social propensity that again can be traced back to huge differences in the respective social contexts.

The implications of the previous considerations for our discussion of transaction costs are twofold: in the first place, the actual relevance of transaction (monitoring) costs for the viability of a cooperative relationship heavily depends on the social context in which such relationship is embedded, and specifically on whether or not reliability and responsibility have the status of a social standard of behavior; moreover, the level of monitoring costs may itself be influenced by the social context in that, given the informational technology in use, in a situation of generalized compliance a habitual defector is much more 'visible' and recognizable than in a situation of low compliance, due to the action of social mechanisms like stigma and ostracism which actually work as social channels of trasmission of the relevant information. The paradoxical implication is that once this social dimension is taken into account, one can have that economic agents are willing to buy monitoring about partners when its costs are relatively (but not exceedingly) high but not when they are low: the reason is that in the first case cooperation is problematic at the social level and low trust results, and therefore it is wise to monitor potential partners even if at some cost; on the other hand, when trust is high defectors are very visible and easily detectable so that the cost of monitoring is low, and nevertheless economic agents tend to economize on it because defection becomes exceptional. If not to a complete reversal, this amounts to a deep modification of the intuition coming from a transaction cost analysis entirely focused on technological factors. In a recent and insightful retrospective essay, Coase (1988) himself stresses how reductive interpretations of transaction costs like the 'technological' one in the present context often miss the point and that social factors may play an important role in determining the actual level of such costs in many situations of interest, for example through their effects on the structure of contractual agreements.

A natural way to accommodate for the role of such social forces within a formal model of strategic interaction between firms is by means of an evolutionary game theoretic model. Evolutionary game theory is one of the most rapidly growing areas of research in contemporary economic theory, and the social selection of cooperative (and more generally pro-social) behaviors is one of the outstanding issues in this new body of literature. A widely chosen line of attack has been that of justifying the emergence of cooperation in social selection processes by assuming that self-interested players may be willing to act according to some pseudo-altruistic set of preferences rather than according to their own self-serving preferences if other self-interested players rationally choose to do the same and if this leads to the enforcement of otherwise unattainable (or to the more likely emergence of previously attainable) 'desirable' outcomes (think e.g. of the cooperative outcome in prisoner's dilemma games) [see Sen (1974, 1985), Mueller (1986), Raub and Voss (1990), Hechter (1990), Kliemt (1990), Guth and Kliemt (1994), Menicucci and Sacco (1996a,b), Sethi (1996); for a more conventional 'business-oriented' view about the viability of pro-social behavior see Maitland (1985)]. If pseudo-altruistic play has to be viable and rewarding, it is needed that such players are able to anticipate correctly the strategy of the opponent (and therefore to recognize each other) and to adjust optimally their behavior to the anticipated behavior of the opponent. When such foresight may be obtained at a cost, choosing to buy it is however not necessarily optimal [see Guth and Kliemt (1994), Sethi and Franke (1995)]. Molander (1993) shows that, in the special context of prisoner's dilemma games, the enforcement of cooperation via conditional cooperation is more difficult in n-wise interaction than it is in pairwise interactions that are typically considered in the previously cited literature. Sugden (1993) takes an alternative approach justifying cooperation on the basis of the superposition of individual and group incentives. Kranton (1996) studies a model where firms can form stable cooperative partnerships but with an 'exit' option that they can choose should they doubt about the future viability of the actual relationship; however, firms have to take into account that the establishment of a new relationship entails a cost. She finds that stable cooperative relationships can be formed in this context if the value of the actual partnership can be increased trhough time via specific investments that would be lost if quitting the relationship. Finally, in Landi and Sacco (1996) the cooperative relationship requires a costly two-stage specific investment and entails an opportunity cost that in some cases may be larger than the relationship's completion value; in this context, if actual cooperation does emerge as a social standard of behavior, it is of the conditional (viz., ‘cautious’, type): trust is too risky because of the big stake that players have to put to enter the relationship and because of the large opportunity costs. Nevertheless, a certain amount of trust may emerge even in this somewhat ‘hostile’ environment under certain conditions.

This very brief and inevitably partial survey suggests that cooperative relationships can emerge as a result of a social selection process under suitable conditions. Other interesting inputs come from the experimental literature in which the role of social factors in determining individual attitudes toward cooperation has been clearly established. Of particular interest is the study of Parks and Vu (1994) showing that, in experimental prisoner’s dilemma interactions, people coming from cultures with a strong collecivistic bias tend to cooperate much more often than people from cultures with an individualistic bias; the cooperative attitude of people with a collectivistic cultural orientation was not undermined even in face of a perfectly defectionist opponent, which seems to imply that the cooperative norm at work is one of pure trust without any monitoring or conditioning on the expected behavior of the opponent. A useful survey of cross-cultural research on social norms of cooperation may be found in Smith and Bond (1993), chapter 7. Other recent experimental results by Hackett (1994) and Berg et al. (1995) point out, respectively, that social norms of equity are an important factor for the explanation of sharing rules in actual partnerships even in the absence of reputational effects and of repeated interaction, and that people show a propensity to participate to reciprocity-oriented cooperative schemes provided that an even weak evidence of viability of such schemes is available.

In this paper we pick up some of the suggestions coming from the previously discussed literature to build a model where transaction costs are endogenously determined by their co-evolution with the social pattern of cooperative vs. defectionist behaviors in a context where a large number of firms interact to carry out, say, R&D joint ventures. Firms have access to a monitoring technology whose cost depends on the actual diffusion of cooperative behavior across the economy. The model explores whether, and if so under what conditions, trust emerges as the eventual social standard of behavior, i.e. under what conditions firms find it optimal to cooperate with partners without having to monitor them. The behavior of firms is assumed to evolve through time according to a social selection process that is indexed by three parameters: one for the order of magnitude of transaction costs, another for the actual completion value of the joint venture, and the last for a cooperation externality that increases with the number of successfully completed joint ventures: in other words, the value of being cooperative is larger the better the past record of cooperation (this may be due to knowledge spillovers, to network externalities, to technological complementarities etc.). The interplay of the three parameters is complex and gives rise to nontrivial patterns of social dynamics. Results are not easily summarized in a few sentences, also because the complexity of the social dynamics mandates resolution by means of simulation analysis. However, some interesting stylized facts emerge.

First, and unsurprisingly, if the completion value of the joint venture is large enough, plain trust always emerges as the social standard of behavior, i.e. cooperation is so rewarding that everybody finds it optimal to comply without the need to monitor the opponent's behavior (on the other hand, when the whole economy is made of cooperators the costs of monitoring would drop to zero anyway). Convergence to cooperation is quicker the larger the completion value; however, for relatively small but still sensible levels of the completion value, cooperation may eventually emerge even after a first transitional phase in which it seems to die out. When the initial distribution of behaviors across the population prescribes an even splitting between cooperators and noncooperators, the role of cautious cooperation (i.e., cooperation with monitoring of the opponent) does not have a major role in the establishment of the cooperative norm, and the emergence of unconditional cooperation (i.e. of trust) is observed. If in the other hand the completion value is positive but relatively low, cooperation dies out eventually. As to transaction costs, their order of magnitude may play a very important role when the completion value is low. When the cost of monitoring is low enough, the existence of cautious cooperators who monitor their opponent may help bringing about trust (i.e. unconditional cooperation) in that their presence lowers the value of defectionist behavior; once defectionists begin to decline, unconditional cooperation becomes rewarding and cautious cooperation dies out. In a sense, then, cautious cooperation seems to have the role of ‘breaking the path’ toward the eventual emergence of trust. The effect of the cooperation externality is symmetrically opposite to that of transaction costs: when the externality is relatively low, cooperators and defectionists can coexist for moderate levels of the completion value and the latter prevail for low enough levels. The comparative strength of transaction costs vs. cooperation externality as critical parameters for the eventual selection of cooperative behaviors seems to be roughly equal for a uniform initial distribution of behaviors, i.e., the effects of a simultaneous decrease of the two parameters by the same amount tend to cancel out.

If we consider an initial distribution of behaviors in which defectionist players are a large majority, a very insteresting dynamic pattern emerges for moderate levels of the completion value. If the order of magnitude of transaction costs is very low, one has a two-stage dynamics such that, at stage one, cautious cooperators spread over whereas unconditional cooperators die out; as a consequence of this, defectionists, who can hardly find someone to exploit, die out as well; but then phase two suddenly takes over: once the proportion of defectionist players has been cut, the proportion of cautious cooperators drops and that of unconditional cooperators ‘explodes’ bringing about a social standard of behavior based on plain trust. If however the completion value is too low, cautious cooperators do not manage to set the stage for unconditional cooperators and defectionist behavior spreads over.

These results are of course but a preliminary investigation of the dynamic patterns generated by the social selection process. Further work will be needed to arrive at a complete characterization for arbitrary initial distributions of behaviors and arbitrary parameter constellations.

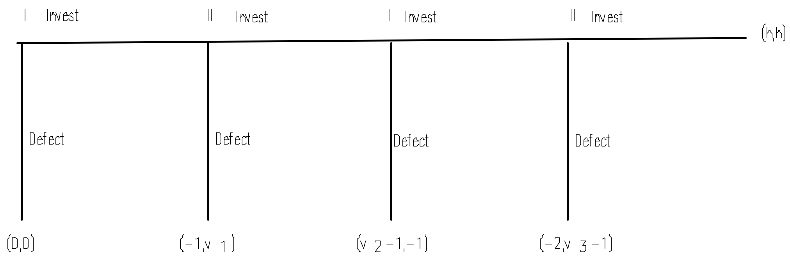
The remainder of the paper is organized as follows. Section 2 is devoted to the exposition of the basic model. Section 3 introduces the social dynamics. Section 4 presents and discusses the results. Section 5 concludes.

2. The basic model

Consider a cooperative relationship (e.g. a R&D joint venture) whose successful completion requires a sequence of incremental investments to be made by two parties; we will refer to these parties as player I and II, respectively. For simplicity, we assume that both players have to make two investments: player I begins, then it is the turn of player II and then again I and finally II. Such investments are however not entirely specific to the project being undertaken. At each stage players may choose to defect selling the asset being built so far to a third party; the existing structure of property rights does not allow to rule out this possibility. Keeping in mind the R&D joint venture as the reference case, defection may be interpreted as one firm selling the knowledge produced in the R&D joint venture to another rival firm; it is then difficult for the partner being cheated to prove that the illegal transaction has taken place and that the third party has not acquired such knowledge through its own R&D activity. Once one player has defected, the value of the project for the other player drops to zero and investment previously made amounts to a net loss. As to the defecting player, the value of selling out knowledge to a third party is of course higher the larger the aggregate amount of investment put in the project: that is to say, the project has an opportunity cost that increases as the project advances[1]. We can therefore represent the strategic interaction between players I and II by means of the following extensive form game with perfect information:

[1] In Kranton (1996) the opportunity cost for the involved firms *decreases* as the project advances because of the specificity of the investment decision. Here cooperation seems more difficult to enforce than in Kranton's framework because the salability of realized investments on the side market implies that the temptation to defect grows as the relationship goes on.

(1)



where h is the *net* completion value of the relationship (i.e., actual completion value minus investment costs) and the v_is are the opportunity costs at the various stages of the project, i.e., the amount that can be obtained the defecting player if selling the asset to a third party when the aggregate level of realized investment is i, $i=1,2,3$, respectively; one has $v_1<v_2<v_3$. Clearly, the payoff to the defecting player is equal to the net opportunity cost, i.e. the amount v_i minus the cost of the player's own past investment, if any.

In this paper we assume that the roles of players I and II are not interchangeable, i.e., the two roles, that we will henceforth call position I and II, respectively, correspond to two distinct sets of specific skills (for example, technological or informational skills). That is to say, players I and II are not distinguished only by their timing of play, but also by some intrinsic characteristics; one can think that the two kinds of firms give complementary, and unsustitutable, contributions, and that the sequence of play reflects the order in which the specific skills of the two partners are to be supplied for the eventual, successful completion of the project[2]. We associate to both players I and II a set of strategic types, that will be denoted as *(I,r)* and *(II,s)*. By (I,r) we clearly mean a position I player whose strategic type is r, $r=1,...,R$, and accordingly *(II,s)* is a type s position II player, $s=1,...,S$. All players of a given strategic type are characterized by the same behavior. In this paper we will consider four strategic types for each position, i.e., $R=S=4$. Moreover, we assume that the behavioral types for each position are suitable specifications of a common behavioral repertoire. We can therefore index

[2] Antoci, Sacco and Scarpa (1996) study an analogous partnership model in which the roles of the two kinds of firms are interchangeable.

without ambiguity the set of behavioral types as 1, 2, 3, 4 without having to bother about the position; clearly, the actual interpretation of the behavioral type will still depend on the position. We consider the following types:

- Explicit Opportunists (type 1);
- Cooperative Screeners (type 2);
- Pure Cooperators (type 3);
- Subtle Opportunists (type 4).

By Explicit Opportunist we mean a player type that defects as soon as he is called upon to play. If playing at position I, this type of player quits the cooperative relationship at the very beginning. If playing at position II and if the position I player did actually invest, he sells out the investment to a third party, and more specifically to a position I Explicit Opportunist. In conclusion, position I Explicit Opportunists are not interested in cooperative relationships and prefer to buy assets from other defectionist players. Position II Explicit Opportunists are among those who sell out; more precisely, they specialize in 'raw' assets, i.e. sell out assets whose level of incorporated investment is low and never make any investment themselves.

Cooperative Screeners are those cooperative players that we have called 'cautious cooperators' in the discussion in section 1. They are willing to cooperate but do not trust their partner; therefore, they monitor the partner's actual strategic type (carrying the corresponding cost) and play a best reply to the partner's strategy. Specifically, if the partner is a cooperator, they cooperate; if he is a defectionist, they quit (i.e., defect themselves). For this strategic type, positions I and II vary only in the amount of risk that has to be taken: position I Cooperative Screeners must protect themselves from the possibility that the partner defects immediately and from the even more frightening possibility that he defects at the last stage of the game, when the cooperative position I player has already provided 2 units of investment. On the other hand, a position II Cooperative Screener must face only one possibility of defection: that of the position I player at stage 3 of the game, which would imply a net loss of 1 unit of investment; indeed, if the position I player did not invest at stage I, the position II Cooperative Screener would never be called upon to play.

Pure Cooperators are cooperative players who rely on the cooperative attitude of their partner and therefore do not monitor him; in other words, their behavior is based on trust. By doing so, Pure Cooperators economize upon the cost of monitoring but run a substantial risk of being cheated. As for Cooperative Screeners, the level of risk is different for positions I and II. Position I Pure Cooperators may be cheated twice losing up to 2 investment units, position II Pure Cooperators only once. In a well defined sense, the kind of trust being implied by the behavior of position I Pure Cooperators is therefore a deeper one.

Finally, Subtle Opportunists are players whose aim is to cheat their partner but, unlike Explicit Opportunists, they specialize in selling out assets with a high level of incorporated investment. (Clearly, potential buyers are again position I Explicit Opportunists). As a consequence, Subtle Opportunists behave as follows. First of all, they monitor the behavioral type of their partner; if she is a defectionist player (i.e. an Explicit or Subtle Opportunist) or a Cooperative Screener, they defect immediately to prevent the partner's defection. If conversely the partner is a Pure Cooperator, they make their investment so as to defect at the next stage they are called upon to play, in order to build a high-valued salable asset. In other words, Subtle Opportunists disguise as cooperators in order to exploit the trust of Pure Cooperators as much as possible. Position II players obviously have more rewarding opportunities with respect to position I colleagues, in that they can sell out assets that incorporate as many as 3 units of investment, whereas the latter can only sell out assets carrying 2 units of investment at most.

In the remainder of the paper we will indicate the proportion of players of type *(I,r)* as x_r and the proportion of players of type *(II,s)* as z_s, respectively, with *r,s=1,2,3,4*. Of course, $\sum_{r=1}^{4} x_r = \sum_{s=1}^{4} z_s = 1$.

At this point we introduce the structure of transaction (monitoring) costs and of cooperation externalities discussed in section 1; as explained, the level of monitoring costs is assumed to be endogenous and in particular to depend on the distribution of behavioral types. Specifically, the more the 'cheaters' hanging around, the more costly to monitor the partner's type, in that defections are more 'visible' and sanctionable (e.g. through social stigma and bad reputation) when defective behavior is exceptional than when it is the norm. We differentiate between Explicit and Subtle Opportunists as to their impact on monitoring costs and assume that the incidence of Subtle Opportunists on costs is twice as much than that of Explicit Opportunists; the rationale of the assumption should be clear: the unreliability of the former type of players is more difficult to detect than that of the latter, in that only Subtle Opportunists disguise as cooperators whereas Explicit Opportunists don't. As a consequence we write monitoring costs *M(x,z)* as follows (where x, z are the four dimensional vectors (x_r), (z_s)):

$$(2) \qquad M(x,z) = \mathrm{h}(2 - x_2 - x_3 - z_2 - z_3 + x_4 + z_4)$$

where $\mathrm{h} > 0$.Notice in particular from (2) that if all players are cooperators, monitoring costs drop to zero. The different incidence on costs of Opportunists players can be easily checked noting that $2-x_2--x_3--z_2--z_3=x_1+x_4+z_1 +z_4$..

As to the cooperation externality *E(x,z)*, it is assumed that the value of cooperation increases with the number of successfully completed projects. In a large population where matching of partners is random one can therefore write

$$(3) \qquad E(x,z) = \mathrm{e}(x_2 + x_3)(z_2 + z_3)$$

where $\mathrm{e} > 0$. The existence of this externality may therefore bring about a 'bandwagon effect' that amplifies the rewards from cooperation as the proportion of cooperators increases, thereby inciting more players to cooperate, and so on.

Finally, we must specify the structure of the side markets for the assets sold by defecting players. On the demand side we find only one type of player: position I Explicit Opportunists. On the supply side we find, respectively, position II Explicit Opportunists selling an asset that incorporates 1 unit of investment, position I Subtle Opportunists selling an asset that incorporates 2 units of investment, and position II Subtle Opportunists selling an asset that incorporates 3 units of investment. We will assume that when markets clear all three assets give to position I Explicit Opportunists a constant profit that we fix equal to 1 (clearly, the return on the three parallel markets must be the same at equilibrium). When supply is large with respect to demand (i.e. when the number of defectionist players selling out assets is large with respect to the number of players buying them) the profit for position I Explicit Opportunists increases, and conversely when supply is small with respect to demand. Of course, defectionist players selling more valuable assets make larger profits; once again, however, the actual profit earned by defectionists depend on the relative size of supply vs. demand. We assume that the equilibrium profits for a player selling an asset that incorporates *i* units of investment (henceforth we will denote it as *i*-asset for short) are equal to *i*. We are now in the position to compute the values of the opportunity costs v_i, $i=1,2,3$. Let us begin by v_1. In this case (1-asset), the level of demand is equal to the proportion of position I Explicit Opportunists, whereas that of supply is equal to x_3z_1: in order to have one 1-asset supplied it is necessary that a position I Pure Cooperator is matched to a position II Explicit Opportunist. So far for the price being paid for a 1-asset; but if the position II Explicit Opportunist must actually sell such an asset, he has to be matched to a position I Pure Cooperator and has to find a position I Explicit Opportunist buying it; otherwise, he does not make any profit. We assume that the probability of finding a buyer for a defecting player selling out an asset is equal to the frequence of prospective buyers, namely x_1: the market for *i*-assets is then characterized by out-of-equilibrium trade and mismatching between suppliers and demanders. If the defecting player is unlucky and does not find a buyer, the asset is lost and his profit is zero; when x_1 is relatively low this may happen rather frequently.

Denoting by *p* the parameter that measures the impact of the excess demand/supply on the market price, we can then write

(4) $$v_1 = [1 + p(x_1 - x_3 z_1)]x_1 x_3$$

Following the same line of reasoning it is easy to show that v_2 and v_3 are equal, respectively, to

(5) $$v_2 = [2 + p(x_1 - x_4 z_3)]x_1 z_3$$

(6) $$v_3 = [3 + p(x_1 - x_3 z_4)]x_1 x_3$$

3. Social dynamics

We are now in the position to compute the payoff earned by the various player types conditional on the current distribution of player types across the economy. We will denote by p*(I,r)* the payoff earned by player type *(I,r)* and by p*(II,s)* the payoff earned by player type *(II,s)*.

As to position I Explicit Opportunists, by always quitting the cooperative relationship at first stage they always get a zero payoff apart from cases in which they trade with defectionist players selling out assets. Remembering that their equilibrium profit on all *i*-assets is equal to 1, by the same logic followed in the derivation of equations (4-6), it is easy to check that

(7)

$$\mathrm{p}(I,1) = [1 + p(x_3 z_1 - x_1)]z_1 x_3 + [1 + p(x_4 z_3 - x_1)]z_3 x_4 + [1 + p(x_3 z_4 - x_1)]z_4 x_3$$

Position I Cooperative Screeners must pay the cost of monitoring *M(x,z)* as measured by (2) above, but also enjoy the cooperation externality *E(x,z)* as determined by (3) above. In addition, they immediately quit the relationship when matched to defectionist players, whereas when they are matched to cooperative players (position II Cooperative Screeners or Pure Cooperators) they stay in and eventually get the completion value *h*; therefore, we have that

(8) $$\mathrm{p}(I,2) = h(z_2 + z_3) - M(x,z) + E(x,z)$$

The case of position I Pure Cooperators is more complex. They always make their investment and therefore may be cheated; on the other hand they do not have to pay monitoring costs. Like Cooperative Screeners, they always get the completion value *h* when matched to a cooperative position II player. On the other hand, if they are matched to a position II Explicit Opportunist they lose 1 investment unit, whereas if they are matched to a position II Subtle Opportunist they lose 2 investment units. Summing up, we have

(9) $$\mathrm{p}(I,3) = -z_1 + h(z_2 + z_3) - 2z_4 + E(x,z)$$

Position I Subtle Opportunists are very easily handled. They only invest and make profits when matched to position II Pure Cooperators (provided that they find a buyer); in addition they have to pay monitoring costs. Therefore,

(10) $$\mathrm{p}(I,4) = \left[2 + p(x_1 - x_4 z_3)\right] x_1 z_3 - z_3 - M(x,z)$$

Coming now to position II Explicit Opportunists, matters are once again quite simple: they only earn nonzero payoffs if matched to position I Pure Cooperators (an if finding a buyer). That is to say,

(11) $$\mathrm{p}(II,1) = \left[1 + p(x_1 - x_3 z_1)\right] x_1 x_3$$

As to position II Cooperative Screeners, their payoffs are exactly analogous to those of their position I colleagues:

(12) $\pi(II,2) = h(x_2 + x_3) - M(x,z) + E(x,z)$

and the same can be said for position II Pure Cooperators with respect to their position I counterparts (notice that position II Pure Cooperators cannot be cheated by position I Explicit Opportunists):

(13) $\pi(II,3) = h(x_2 + x_3) - x_4 + E(x,z)$

And finally, position II Subtle Opportunists have to pay monitoring costs and once again make profits (and invest) only if matched to a position I Pure Cooperator (and if finding a buyer), from which it follows

(14) $\pi(II,4) = \left[3 + p(x_1 - x_3 z_4)\right] x_1 x_3 - x_3 - M(x,z)$

At this point the payoff structure of all player types is fully specified. To determine how the proportion of the various types evolves through time we have now to introduce the social dynamics. In this respect we are faced with a multiplicity of possible choices; in the evolutionary game-theoretic literature the main trend is that of modelling social dynamics by means of the so called ‘replicator dynamics’, which has initially been used by biologists in their modelling of Darwinian selection processes [see e.g. Hofbauer and Sigmund (1988)]. The basic feature of the replicator dynamics is that of assuming that the rate of growth of the frequence a certain strategy within a given population is equal to the difference between the actual performance of that strategy and the average performance of all strategies that are present in the population. In other words, the more a certain strategy does ‘better than average’, the more it grows; the more it does ‘worse than average’, the more it dies down. As we have argued elsewhere [see Sacco (1994)], this dynamic specification has some rationale also in the case of cultural rather than biological selection processes where transmission of strategic types takes place via purposeful human choice rather than via biological reproduction. Various authors question the use of replicator dynamics in economic applications as a sensible model of social selection processes [see e.g. van Damme (1994)] but on the other hand other authors have shown that this dynamics can be rigorously derived as an aggregation of a large number of interactions between boundedly rational agents [see e.g. Bjornerstedt and Weibull (1995) and also the comments in Sacco (1995)] so that choice of the replicator selection mechanism is not entirely *ad hoc* but has some factual content in that it implies clear cut restrictions on the information gathering and processing routines adopted by players. Moreover, it turns out that a large class of selection mechanisms may be represented as a suitable (generally state-dependent) time rescaling of the replicator mechanism [see e.g. Samuelson and Zhang (1992)], which therefore obtains a somewhat 'focal' position in the choice space of the model maker interested in social selection processes.

In view of the above considerations, in this paper we will adopt the replicator selection mechanism, which can be written as follows:

(15) $$\dot{x}_r = x_r[\pi(I,r) - \pi(I,m)]$$

(16) $$\dot{z}_s = z_s[\pi(II,s) - \pi(II,m)]$$

Here dots denote first derivatives with respect to time; moreover *r,s=1,2,3,4* and the $\pi(\bullet,m)$ are the average payoffs relative to positions I and II, respectively, which are given by

(17) $$\pi(I,m) = \sum_{r=1}^{4} \pi(I,r)x_r$$

(18) $\pi(II,m)=\sum_{s=1}^{4}\pi(II,s)z_s$

Notice that the restriction that both the *x*s and the *z*s have to sum up to one implies that both equation sets (15) and (16) include a redundant equation. The phase space of the dynamical system (15-16) is therefore given by the Cartesian product of two three-dimensional unit simplexes, one for each position.

Equations (15-16) then describe the social dynamics of behaviors in the economy. Its intuitive rationale is the following: players sample the relative profitability of the various strategies and tend to switch from less to more rewarding ones. The better the performance of a strategy with respect to average, the more outstanding it is from the point of view of players, and therefore the more likely it is that it is adopted [see Bjornerstedt and Weibull (1996) for a precise characterization of the sampling and decision making process]. In this way, some types of behavior are selected at the expense of others and the dynamics eventually lead to the establishment of a social standard, i.e. to a set of behaviors (possibly a singleton) that players find it convenient to adopt in that, *given the current distribution of behaviors across the population*, no feasible alternative proves to be more rewarding, or at least equally rewarding. In general, therefore, reaching a social standard of behavior does not imply that the selection process lands on a stationary distribution: when the selected behaviors are more than one, the social dynamics may keep on wandering within an invariant set (i.e. a subset of the whole phase space) [see e.g. Weibull (1995)]. In several cases, however, single behaviors are actually selected; in particular, it has been proved that strict Nash equilibria are always attractors for the replicator dynamics (as well as for a more general class of selection processes) and that, in addition, in multi-population models as it is the one studied in this paper, *only* strict Nash equilibria may be point attractors for such processes [see e.g. Weibull (1995)]. In conclusion, under the replicator selection mechanism a homogeneous standard of behavior will be selected if and only if it is a strict Nash equilibrium for the underlying game.

The search for social standards of behavior then amounts to the search for point- ot set-attractors of the dynamics described by equations (15-16).

4. Results

As already remarked, the dynamic process described by equations (15-16) takes place in a six-dimasional space; an analytical characterization of the dynamics as depending on the whole set of parameters of the model is therefore very demanding and, in principle, not necessarily feasible. For this reason we will rather present and discuss some simulations, which provide us with a first clue of the (presumably much larger) set of behaviors that can be generated by the model, and whose fuller characterization will require further work.

The model of section 2 contains four parameters that can be of potential interest as to their bearings on the social dynamics: η, namely, the order of magnitude of transaction (monitoring) costs; ε, the level of the cooperative externality; h, the (net) completion value of the project, and p, the parameter that measures the sensitivity of prices to the disequilibria that can emerge in the parallel markets for assets that are sold out by defectionist players. We choose to conduct our analysis in terms of the first three parameter only, normalizing p to 1, thereby focusing on the impact of 'fundamental' variables on the social dynamics. The issue of the impact of technical and institutional facets of the structure of the parallel markets for assets (which, as we shall discuss in the final section, are not characterized exhaustively by p) is somewhat away from the main concern of the present paper and is again left for future work.

Intuitively, one can expect that a highly nonlinear, high-dimensional dynamics as it is the one studied in this paper is likely to produce behaviors that display nontrivial sensitivities to choice of the initial conditions. For this reason, and in lack of a complete analytical characterization of the dynamics, meaningful simulations require that initial conditions are selected according to sensible criteria. In this paper we will consider two sets of initial conditions. The first, corresponding to baricentric coordinates (i.e. all behavioral types are present in equal proportions at the beginning of the story), is a natural and somewhat obliged choice, in that it provides information about the social dynamics in a context where no behavioral type can exploit possible advantages deriving from its under- or over-representation in the current distribution of behaviors: for example, the potential payoff to cooperation is increasing in the proportion of cooperators whereas the payoff to defection is decreasing in the number of defectionists of the same type (in that the corresponding side of the market grows 'longer' thereby cutting the profit margin). The choice of the second set of initial conditions is more subtle and refers to a context where, initially, one of the parallel markets for assets is potentially well developed because both its supply and demand sides are well represented in the population. Specifically, we have focused upon the situation where position I Explicit Opportunists and position II Subtle Opportunists are preponderant in the initial population, with all other types being present in very modest amounts. This is to say that the market for high-value assets (those incorporating 3 units of investment) should be very active. Moreover, in this situation it is very dangerious for Pure Cooperators to rely upon trust, i.e. non-monitored cooperation, in that the potential risk of being exploited is very high. This set of initial conditions is therefore a very severe test for the viability of trust as a social standard of behavior: if trust can eventually emerge as a rewarding option in a context in which there is initially a very high probability that such trust is ill-posed, one can conclude that there is a strong evolutionary argument for the social viability of trust.

Formally, we denote initial conditions by $x_r(0)$ and $z_s(0)$, respectively, $r,s=1,2,3,4$. The first set of initial conditions will then be given by $x_1(0)=x_2(0)=x_3(0)=x_4(0)=z_1(0)=z_2(0)=z_3(0)=z_4(0)=0.25$ and will be henceforth referred to as the baricentric scenario. The second set of initial conditions will be given by $x_1(0)=z_4(0)=0.76$, $x_2(0)=x_3(0)=x_4(0)=z_1(0)=z_2(0)=z_3(0)=0.08$ and will be henceforth referred to as the exploitation scenario. In this latter situation, then, position I Explicit Opportunists and position II Subtle Opportunists initially represent (slightly more than) three quarters of the respective populations, with the remaining shares equally distributed among all other types.

4.1. Baricentric scenario.

The most sensible strategy for the investigation of the impact of the three parameters *h*, η, ε on the social dynamics is to begin by isolating the effect of single parameters and then to clarify their interactions. First we focus our attention on the role of *h*. Setting ε=η=*1* as a reference, we find the following pattern. As one could expect, when *h* is low enough, i.e. when the continuation value is small, cooperation dies out. Specifically, one can check that there is a value *h* between 0.3 and 0.46973 below which position I Explicit Opportunists gradually spread over[3] until immediate defection becomes the standard of behavior for position I players. On the other hand, the low relative value of the critical level of the continuation value *h* with respect to the transaction cost η implies that monitoring is expensive enough to imply the eventual disappearance of all player types of *both* positions who monitor their opponents. This implies that the eventual equilibrium distribution for position II players is a mix of Explicit Opportunists and Pure Cooperators. Clearly, at the eventual equilibrium such distribution is irrelevant for the final outcome in that, being all position I players Explicit Opportunists, no project ever goes beyond stage 1 and so position II players are never called upon to play. But during the transition toward the equilibrium, the distribution of position II players is indeed relevant because cooperation is still possible (although less and less likely as position I Explicit Opportunists spread over). Perhaps rather surprisingly, it turns out that even when *h* is as low as 0.3, the equilibrium distribution of position II players contains a 40 per cent of Pure Cooperators. Simple reasoning is however needed to check that this result is indeed sensible: as monitoring players die out because of the high relative level of transaction costs, as

---

[3] Here and in the following, expressions like ‘players of type t spread over’ or ‘die out’ do not mean that they actually add to, or leave from, the players pool. As explained in section 3, the social dynamics studied in this paper are a cultural selection process driven by imitation of successful strategies rather than a natural selection process driven by physical replication and extinction of players. Therefore, the above expressions have to be meant as ‘an increasing proportion of players adopt strategy t’ or ‘dismiss strategy t’, respectively.

far as position II is concerned the competition is between Explicit Opportunists and Pure Cooperators. Now, the payoff of position II Explicit Opportunists entirely comes from the exploitation of position I Pure Cooperators; since the latter die out, the former must die out as well, despite a first transitory phase in which their proportion increases because of the still high number of easily exploitable position I Pure Cooperators. At this point one could object that the disappearance of position I Pure Cooperators is also harmful to position II Pure Cooperators whose chances of successful completion of a project descrease as the proportion of reliable partners vanishes. On the other hand, one must note that the level of the cooperation externality ε is again very high relative to those of the other parameters and therefore, despite the modest stock of successfully completed projects, it is enough to make a difference and to allow position II Pure Cooperators to make bigger profits from the cooperation with their position I colleagues than position II Explicit Opportunists do by exploiting them [figure 1(a)]. Clearly, the more the continuation value h drops below 0.3, and the lower the value of the cooperation externality e, the larger the relative proportion of position II Explicit Opportunists at the equilibrium. Notice the subtlety and complexity of dynamic interactions between cooperator and defectionist player types emerging from the pattern just analyzed: the decline of a cooperator player type in one population may bring about the success of the same type in the other population, and conversely the success of a defectionist player type in one population may bring about the decline of the same type in the other, and this despite the strong complementarities existing between couples of homogeneous player types belonging to different populations.

If however h rises even slightly above h (for example, for h=0.46973), the social dynamics changes entirely. Because of the high relative level of the cooperation externality e, this apparently negligible increase in the continuation value is enough to bring about trust as the eventual social standard of behavior. The success of pure cooperators, however, is not straightforward; once again there is a first transitory phase where position I and II Explicit Opportunists spread over whereas position I Pure Cooperators decline (as before, all monitoring players types quickly die out). The trend is however quickly inverted for position II Explicit Opportunists in that the slight increase in *h*, together with the high value of e, causes the steady increase of the already strong position II Pure Cooperators (see the above analysis for *h*=0.3); as a consequence, position II Pure Cooperators always increase more quickly than Explicit Opportunists and, from a certain point on, even increase at the expenses of Explicit Opportunists. This trend reversal in the position II population is however, at least in a first phase, not strong enough to invert the trend in the position I population, where Explicit Opportunists keep on growing at the expenses of Pure Cooperators; at a certain point their proportion almost reaches 70 per cent. However, after a relatively long phase in which the position I population seems to have stabilized at a distribution with a modest presence of Pure Cooperators, the changes in the position II population suddenly become relevant: the quick growth of position II Pure Cooperators creates strong new opportunities for cooperation and as a consequence the relative shares of position I Explicit Opportunists and Pure Cooperators ‘swap’ very quickly. After this sudden change of regime the dynamics enters in the final phase of the transition where Explicit Opportunists slowly but steadily die out and trust gradually becomes the social standard of behavior (figure 1b). Further increases in *h* imply a more and more steady affirmation of trust. For *h*=0.6, for instance, the continuation value is large enough to allow the shares of both position I and II Pure Cooperators to grow from the beginning, thereby cancelling the initial phase in which defectionist player types seemed to prevail for *h*=0.46973 (figure 1c). Finally, it is interesting to point out how even further increases in *h* cause the survival of a positive share of position II Cooperative Screeners at the equilirium. The reason is once again intuitive after some reflection: as the payoff from cooperation grows larger, position I Pure Cooperators wipe out all rival types almost immediately; as a consequence, despite the relatively high level of η, transaction costs drop to nearly zero very quickly, thereby leaving room for those few position II Cooperative Screeners who have not yet disappeared: once monitoring costs nothing, the performance of Pure Cooperators and Cooperative Screeners becomes identical (remember that defectionist types have previously died out). Thus we have once again an apparently paradoxical result: the quick affirmation of position I Pure Cooperators favors the survival of position II Cooperative Screeners at the expenses of position II Pure Cooperators. The effect is however a modest one, at least for modest increases in *h*. For example, when *h*=0.7, the equilibrium proportion of position II Cooperative Screeners is less than 0.05 (figure 1d). But as *h* diverges, the equilibrium proportion of the latter type of players consequently increases: the more and more sweeping affirmation of position I Pure Cooperators quickly cancels any difference between position II Pure Cooperators and Cooperative Screeners by dropping at once transaction costs to almost zero. In the limit, as *h* tends to infinity, the two types become practically equivalent and a fifty-fifty equilibrium share is observed (figure 1e).

With this reference analysis in mind, we can now explore the impact of the other model parameters. As to transaction costs h, it is not surprising to find that a decrease in the order of magnitude of transaction costs causes the survival at the equilibrium of larger shares of position II Cooperative Screeners for given (high enough) levels of *h* such that position I Pure Cooperators prevail, with respect to the case where h was higher. For example (with ε still fixed at 1), when h=0.4, *h*=0.4 implies a positive even if very small equilibrium share of position II Cooperative Screeners, whereas, as shown in figure 1(c), such share was still null for *h*=0.6 when h was set to 1 (figure 2a). Consequently, as *h* tends to infinity one has again that the equilibrium distribution in the position II population tends to a fifty-fifty state, even more quickly than in the case with h=1.

Another important consequence of the fall of transaction costs is the widening of the range of values of *h* that cause the eventual prevalence of cooperation. A first interesting, and somewhat 'intermediate' case is that with h =0.4, *h*=0.3. In this instance, position I Explicit Opportunists still prevail but there is an intriguing effect on the equilibrium distribution of position II players. The fall in the order of magnitude of transaction costs, although still causing the eventual extinction of monitoring player types, drastically changes the transitional dynamics. On the one hand, position I Cooperative Screeners decline steadily because of the low value of the completion value. On the other hand, the share of position I

Subtle Opportunists has a sudden ‘burst’ after an initial decline: the reason is that the lower level of transaction costs makes it profitable for them to exploit the large and increasing share of position II Pure Cooperators that builds up together with that of position I Explicit Opportunists. However, this state of things does not last: as position II Pure Cooperators decrease as a consequence of their heavy exploitation by position I Subtle Opportunists, the size of the ‘suckers’ pool’ shrinks; in addition, the decrease of position II Pure Cooperators causes the expansion of position II Explicit Opportunists, which in turn advantages position I Explicit Opportunists because of the mutually favorable trading opportunities that derive from the exploitation of the (although vanishing) position I Pure Cooperators (figure 2b). One cannot help noticing the complexity of the dynamic interactions taking place in this situation: the low level of transaction costs advantages position I Subtle Opportunists thereby damaging position II Pure Cooperators; the result is that the equilibrium proportion of position II Pure Cooperators then decreases with respect to the case $\mathrm{h}=1$ for $h=0.3$, despite the fact that, on intuitive grounds, one would expect that the decrease of transaction costs would *facilitate* the eventual viability of cooperation. The reason why this intuition is misplaced here is that the completion value is too low to grant the survival of Cooperative Screeners and therefore the reduction of transaction costs basically benefits those player types which use monitoring to better exploit cooperators. The fall of transaction costs from $\mathrm{h}=1$ to $\mathrm{h}=0.4$ for $h=0.3$ does not alter the final equilibrium outcome which is still the extinction of cooperation. However, the complex dynamic effects just described suggest that further decreases of h might bring about more radical changes. This conjecture is easily verified for $\mathrm{h}=0.15$ (and h still fixed at 0.3): here the level of transaction costs is so low that position I Cooperative Screeners survive long enough to undermine seriously the profitability of position II Explicit Opportunists. In addition, and for the same reasons, position II Cooperative Screeners stabilize on a positive equilibrium share (which is above 0.35). The rapid decline of defectionist position II players creates strong opportunities for position I Pure Cooperators which (after a brief initial phase of transitory decline) eventually wipe out all rival types. One can check that, as for the analogous cases analyzed above, the divergence of *h* implies the (fast) convergence toward a fifty-fifty equilibrium distribution for position II players (figure 2c).

Coming now to the impact of the cooperation externality e, consider the case *h*=0.74703 when e drops to 0.4 (h then switches back to 1). Here, given the relatively high level of the completion value, the small size of the cooperation externality is enough to grant the eventual success of position I Pure Cooperators. A small share of position I Explicit Opportunists coexists with Pure Cooperators for a while but then dies out. As to position II, Pure Cooperators prevail (as it happened for larger values of e) (figure 2d). Further falls in e would then push the fading of position I Explicit Opportunists further away in time. Dropping *h* to 0.2 for e at 0.4 causes a further change of regime: position I Explicit Opportunist now prevail, whereas in position II there is a relatively balanced coexistence between Naive Opportunists and Pure Cooperators: the joint reduction of *h* and e thus reduces the profitability of position II Pure Cooperators at the advantage of Naive Opportunists; on the other hand, the high level of monitoring costs causes the extinction of all monitoring types (figure 2e).

Given the above discussion, the effect of a joint reduction of h and e should now be foreseeable: for example, when *h*=1 and h =e=0.3, position I Pure Cooperators prevail because the relative size of the completion value with respect to h and e is large. On the other hand, the low level of h causes the coexistence of position II Cooperative Screeners and Pure Cooperators (figure 2f). Dropping h further to 0.1 brings about an increase of the share of position II Cooperative Screeners; if *h* increases then the familiar pattern of convergence to the fifty-fifty equilibrium distribution emerges. On the other hand, for h =e=0.3 and *h*=0.5, the effect of the reduction of the cooperative externality tends to prevail on that of the reduction of transaction costs: in the position I population there is the coexistence of a small share of Explicit Opportunists and a large share of Pure Cooperators (which rises above a 0.8 equilibrium share after an initial first phase of decline well below 0.3), whereas in the position II population only Pure Cooperators eventually remain (figure 2g).

4.2. Exploitation scenario.

In the baricentric scenario, Pure Cooperation (viz., trust) and Cooperative Screening (viz., 'cautious' cooperation) play substitute rather than complementary roles: there is no synergetically reinforcing interaction between the two types of cooperators to bring about the final prevalence of cooperation. The exploitation scenario provides different insights. We will confine our attention to the parameter constellations which give rise to new instances of dynamic regimes.

Starting from a reference situation in which $\mathrm{h} = \mathrm{e} = 0.7$, and $h=1$, the resulting behavior is analogous to that observed in figures 2(d) and 2(g) above: coexistence of position I Explicit Opportunists and Pure Cooperators and prevalence of position II Pure Cooperators. However, if h drops to figures such as 0.1, a new and unprecedented effect emerges: due to the relatively high level of h and to the relatively low level of h, position I Cooperative Screeners initially prevail whereas Pure Cooperators nearly disappear; the reason is simple: given that in the initial population there is a very high share of dangerous exploiters, cooperators who rely on trust do badly whereas cooperators who monitor do very good. As a consequence, defectionist types die out. It is interesting to remark that the low initial level (and the steady decline) of position I Subtle Opportunist, i.e. the only potential exploiter for position II Pure Cooperators, implies that the latter are not penalized as their position I colleagues increase their share up to 0.8 (they are advantaged with respect to position II Cooperative Screeners in that they economize on monitoring costs). But the really interesting phase emerges as position I Cooperative Screeners (through their interaction with position II cooperators) drive out all defector types. At this point, position I Pure Cooperators no longer run the risk of being heavily exploited and therefore they 'burst out' and in a very short time completely wipe out Cooperative Screeners themselves. In other words, we observe here an endogenous two-phase transition toward the establishment of trust: first cautious cooperation ‘opens the way’ wiping out potential exploiters, and then the less expensive trust suddenly and completely takes over when the risk of exploitation becomes negligible: in other words, the two types of cooperators now work in a complementary rather than substitute way for the establishment of cooperation (figure 3a). As h drops to 0.85 and h is set at 0.2,, the viability of position II Cooperative Screeners is seriously compromised in that the completion value is no longer sufficient to cover transaction costs. As a consequence, position II Pure Cooperators eventually stand almost alone whereas at position I one still observes the two-stage transition to trust through cautious cooperation (figure 3b). As h drops down to 0.7, the first phase of the transition toward trust becomes more complex and troubled: it takes a longer time to position I Cooperative Screeners to wipe out defection, also because the low completion value and the low level of transaction costs initially create remarkable opportunities for position I Subtle Opportunists (who can exploit a raising stock of position II Pure Cooperators) and for position II Explicit Opportunists (who exploit position I Pure Cooperators). However, the near-extinction of position I Pure Cooperators in this first phase ‘kills’ position II Explicit Opportunists whereas poistion I Subtle Opportunists are eliminated by a joint effect: a sudden ‘burst’ of position II Cooperative Screeners that lowers the opportunities of exploitation, and the parallel diffusion of position I Cooperative Screeners and position II Pure Cooperators which gradually increases the profitability of cooperation through the cooperation externality. Whereas phase one of the transition becomes very long, phase two becomes extremely short: a spectacular switch (as to both its quickness and its size) from ‘cautious cooperation’ to trust suddenly occurs (figure 3c). Moving further down to h=0.5 and easing down h at 0.1 the continuation value becomes too small for position I Pure Cooperators to recover eventually: after a long and complex transient phase, cooperation prevails; specifically, ‘cautious’ cooperation prevails in position I and trust prevails in position II (with a small minority of cautious position II cooperators): once defectionists have been wiped out from position I, monitoring in position II can be dispensed with and Pure Cooperators spread over (figure 3d). A parallel increase of h back to 0.2 and of h to 0.6209835 produces an analogous effect but with a even longer transient: now the relatively higher cost of monitoring causes the eventual disappearance of position II Cooperative Screeners; moreover, position I and II Explicit Opportunists are still substantially present after as many as 200 periods (figure 3e). When the drop in h is less drastic, however, the two-stage transition toward cooperation emerges again and with even more spectacular characteristics. Setting h=0.2 and h=0.67 (with e still fixed at 0.7), we find that, as the share of position I Cooperative Screeners rises to almost 1 in the first phase, it is once again suddenly and entirely substituted by the since then virtually extinct Pure Cooperators (figure 3f). When h=0.6209834, the transition becomes extremely quick and it follows an initial period in which defectionist players predominate; a three-stage pattern then emerges: defection, then cautious trust, and finally pure trust; the transition to the latter phase is almost instantaneous (figure 3g). Compare moreover figure 3g with 3e: in the former case, a marginal *drop* in the continuation value h with respect to the latter case causes the eventual emergence of *Pure* Cooperation, in that it favors an earlier affirmation of Position I Cautious Cooperators and *therefore*, once the latter have spread over soon enough, creates a 'last minute' opportunity for Pure Cooperators to take over before they go irreversibly extinct. Interestingly, a marginally lower continuation value is then conducive to pure trust whereas a marginally higher one is not. Notice how in these last figures the dynamic patterns become more and more complex in their transitional behavior, which is characterized by nonmonotonic transitions and by intriguing dynamic interactions among player types. Notice also how the 'parallel market' effect never works in these cases: the proportion of position II Subtle Opportunists always quickly falls from the initial figure of 0.76 to zero. The reason is that, because of the low level of transaction costs, the parallel market is always killed by the action of position I Cooperative Screeners. On the other hand, an increase of the transaction costs would damage Subtle Opportunists as well as Cooperative Screeners and therefore would not help the former to enhance their survival chances.

If however *h* goes down to 0.62095, cooperation suddenly crumbles down: the point is that the completion value is too low and position I Pure Cooperators die out during the former 'phase one' of the process, whereas position I Cooperative Screeners do not manage to sustain transaction costs. As a consequence, position I Explicit Opportunists take over whereas in position II a mix of (preponderant) Explicit Opportunists and (minoritarian) Pure Cooperators emerges according to a logic similar to that of the case depicted in figure 1(a) above (figure 3h).

Finally, the crucial role of transaction costs in the exploitation scenario should not be understated. Starting from the situation of figure 3h above where cooperation is eventually wiped out, even a very small drop of h from 2 to 1.999 becomes enough to restore the two-phase pattern toward the emergence of pure trust (figure 3i). As h drops further to 0.16, the transition to 'phase two' of pure trust becomes quicker and easier (figure 3l). At h=0.12 one finds in addition that position II Cooperative Screeners are now observed at the equilibrium after a transitional phase in which their share 'plunges down' temporarily: interestingly, this dynamic pattern is the opposite of that of their position I colleagues (figure 3m).

5. Conclusions

Although it provides some interesting and somewhat surpising results, the simulative analysis presented in this paper is only but a very first step toward a full understanding of the dynamic behavior of the model of sections 2-3. Although demanding in itself, this fuller simulative analysis is in turn but a first step in the broader framework that emerges by relativizing some of the key hypotheses of the model.

A first point is the relatively minor role played by parallel market for assets in the simulation runs discussed in the previous section. This may depend among other things on the structure of such markets, and in particular on the fact that mismatching can occur: defectionist players cannot be sure to sell their assets in that they must manage to find a buyer and this is often not the case. Assuming that the market is more transparent and that consequently mutually advantageous trades are always carried out might increase substantially the profitability of subtle cheating, thereby leading to more developed and durable parallel markets.

Another open and important point has to do with the legal enforcement system: in the present model, cooperative players that are exploited by defectionists have no way to punish them, neither they can ask for guarantees *a priori* against the possibility of defection of the partner. One could for example imagine the possibility of judicial vindication against defectors; then, depending on the probability with which defectors are punished and on the time delay with which a decision is reached, the payoffs to defectionist behaviors might change; more specifically, the parameters describing the characteristics of the judicial system may add to the ones considered in this paper as factors that have a bearing on the social dynamics. As for the guarantees issue, one could imagine a precommitment arrangement in which partners who agree to carry out the cooperative project pay a bonus that they will lose in the case of defection [this type of mechanism has been explored in Landi and Sacco (1996), although with different motivations and with a different characterization of player types].

Finally, as pointed out by many authors, replicator dynamics is a sensible specification for the modelization of social dynamics in the medium-long run but not in the very long run where the possibility of random disturbances of various kinds should be taken into account; this aspect is particularly relevant for 'centipede' games like the one studied in this paper. It is clear, for example, that in the exploitation scenario the eventual diffusion of Pure Cooperators creates an ideal context for defectionists; therefore, if for whatever reason a small fraction of defectionists 'invades' the equilibrium population it may spread over at the expense of trustful cooperators. But on the other hand this makes room for the diffusion of cautious cooperators who can successfully monitor defectors, and so on: a complex cyclical pattern emerges, as it has been noted by Ho (1991) following a different but somewhat related evolutionary approach.

A fair evaluation of the descriptive and explanatory potential of the model requires that all these possibilities be explored on their own and in a comparative perspective: a busy but promising research agenda for the understanding of the impact of transaction costs on the social dynamics of cooperation and trust.

References

A. Antoci, P.L. Sacco, C. Scarpa (1996), “Endogenous partnership formation in a transitional economy”, mimeo, Department of Economics, University of Florence.

J. Berg, J. Dickhaut, K. McCabe (1995), “Trust, reciprocity, and social history”, *Games and Economic Behavior* 10, 122-142.

J. Bjornerstedt, J.W. Weibull (1996), “Nash equilibrium and evolution by imitation”, in K.J. Arrow et al. (eds.), *The rational foundations of economic behavior*, Macmillan, London.

R. Coase (1988), “The nature of the firm. 1. Origin, 2. Meaning, 3. Influence”, *Journal of Law, Economics and Organization* 4, 3-17.

F. Fukuyama (1995), *Trust*, The Free Press, New York.

D. Gambetta (ed.) (1988), *Trust. Making and breaking cooperative relations*, Basil Blackwell, Oxford.

W. Guth, H. Kliemt (1994), “Competition or co-coperation: On the evolutionary economics of trust, exploitation an moral attitudes”, *Metroeconomica* 45, 155-187.

S.C. Hackett (1994), “Is relational exchange possible in the absence of reputations and repeated contact?”, *Journal of Law, Economics and Organization* 10, 360-389.

M. Hechter (1990), “The emergence of cooperative social institutions”, in M. Hechter, K.-D. Opp and R. Wippler (eds.) *Social institutions. Their emergence, maintenance and effects*, de Gruyter, Berlin.

T.-H Ho (1991), “An evolutionary analysis of the centipede game using classifier systems with genetic-algorithm reproduction”, University of Pennsylvania Decision Sciences Working Paper.

J. Hofbauer, K. Sigmund (1988), *The theory of evolution and dynamical systems*, Cambridge University Press, London.

H. Kliemt (1990), “The costs of organizing social cooperation”, in M. Hechter, K.-D. Opp and R. Wippler (eds.) *Social institutions. Their emergence, maintenance and effects*, de Gruyter, Berlin.

R.E. Kranton (1996), “The formation of cooperative relationships”, *Journal of Law, Economics and Organization* 12, 214-233.

M. Landi, P.L. Sacco (1996), “Norms of cooperation in a partnership game”, mimeo, Department of Economics, University of Florence.

I. Maitland (1985), “The limits of business self-regulation”, *California Management Review* 27, 132-147.

D. Menicucci, P.L. Sacco (1996a), “Evolutionary dynamics with λ-players”, CITG Working Paper no. 19.

D. Menicucci, P.L. Sacco (1996b), “Rawlsian altruism and efficiency”, Studi e discussioni no. 102, Department of Economics, University of Florence.

P. Molander (1992), “The prevalence of free riding”, *Journal of Conflict Resolution* 36, 756-771.

D. Mueller (1986), “Rational egoism versus adaptive egoism as fundamental postulate for a descriptive theory of human behavior”, *Public Choice* 51, 3-23.

C.D. Parks, A.D. Vu (1994), “Social dilemma behavior of individuals from highly individualist and collectivist culture”, *Journal of Conflict Resolution* 38, 708-718.

R.D. Putnam (1993), *Making democracy work*, Princeton University Press, Princeton.

W. Raub, T. Voss (1990), “Individual interests and moral institutions: An endogenous approach to the modification of preferences”, in M. Hechter, K.-D. Opp and R. Wippler (eds.) *Social institutions. Their emergence, maintenance and effects*, de Gruyter, Berlin.

P.L. Sacco (1994), “Selection mechanisms in economics”, Behavioral Science 39, 311-325.

P.L. Sacco (1996), “Comment on Bjornerstedt and Weibull”, in K.J. Arrow et al. (eds.), *The rational foundations of economic behavior*, Macmillan, London.

L. Samuelson, J. Zhang (1992), “Evolutionary stability in asymmetric games”, *Journal of Economic Theory* 57, 363-391.

A.K. Sen (1974), “Choice, orderings and morality” in S. Korner (ed.), *Practical reason*, Blackwell, Oxford.

A.K. Sen (1985), “Goals, commitment and identity”, *Journal of Law, Economics and Organization* 1, 341-355.

R. Sethi (1996), “Evolutionary stability and social norms”, *Journal of Economic Behavior and Organization* 29, 113-140.

R. Sethi, R. Franke (1995), “Behavioral heterogeneity under evolutionary pressure: Macroeconomic implications of costly optimization”, *Economic Journal* 105, 583-600.

P.B. Smith, M.H. Bond (1993), *Social psychology across cultures*, Harvester, Hemel Hampstead.

R. Sugden (1993), “Thinking as a team: Towards an explanation of nonselfish behavior”, *Social Philosophy and Policy* 10, 69-89.

E. van Damme (1994), “Evolutionary game theory”, *European Economic Review* 38, 847-858.

J.W. Weibull (1995), *Evolutionary game theory*, Mit Press, Cambridge, Mass.

CONTENTS

FIGURE CAPTIONS

Figure 1(a). Baricentric scenario. $\varepsilon=\eta=1$, h=0.3.

Figure 1(b). Baricentric scenario. $\varepsilon=\eta=1$, h=0.46973.

Figure 1(c). Baricentric scenario. $\varepsilon=\eta=1$, h=0.6.

Figure 1(d). Baricentric scenario. $\varepsilon=\eta=1$, h=0.7.

Figure 1(e). Baricentric scenario. $\varepsilon=\eta=1$, h=1502.

Figure 2(a). Baricentric scenario. $\varepsilon=1$, $\eta=0.4$, h=0.4.

Figure 2(b). Baricentric scenario. $\varepsilon=1$, $\eta=0.4$, h=0.3.

Figure 2(c). Baricentric scenario. $\varepsilon=1$, $\eta=0.15$, h=0.3.

Figure 2(d). Baricentric scenario. $\varepsilon=0.4$, $\eta=1$, h=0.74703.

Figure 2(e). Baricentric scenario. $\varepsilon=0.4$, $\eta=1$, h=0.2.

Figure 2(f). Baricentric scenario. $\varepsilon=\eta=0.3$, h=1.

Figure 2(g). Baricentric scenario. $\varepsilon=\eta=0.3$, h=0.5.

Figure 3(a). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.1$, h=1.

Figure 3(b). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.2$, h=0.85.

Figure 3(c). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.2$, h=0.7.

Figure 3(d). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.1$, h=0.5.

Figure 3(e). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.2$, h=0.6209835.

Figure 3(f). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.2$, h=0.67.

Figure 3(g). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.2$, h=0.6209834.

Figure 3(h). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.2$, h=0.62095.

Figure 3(i). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.199$, h=0.62095.

Figure 3(l). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.16$, h=0.62095.

Figure 3(m). Exploitation scenario. $\varepsilon=0.7$, $\eta=0.13$, h=0.62095.

fig. 1a

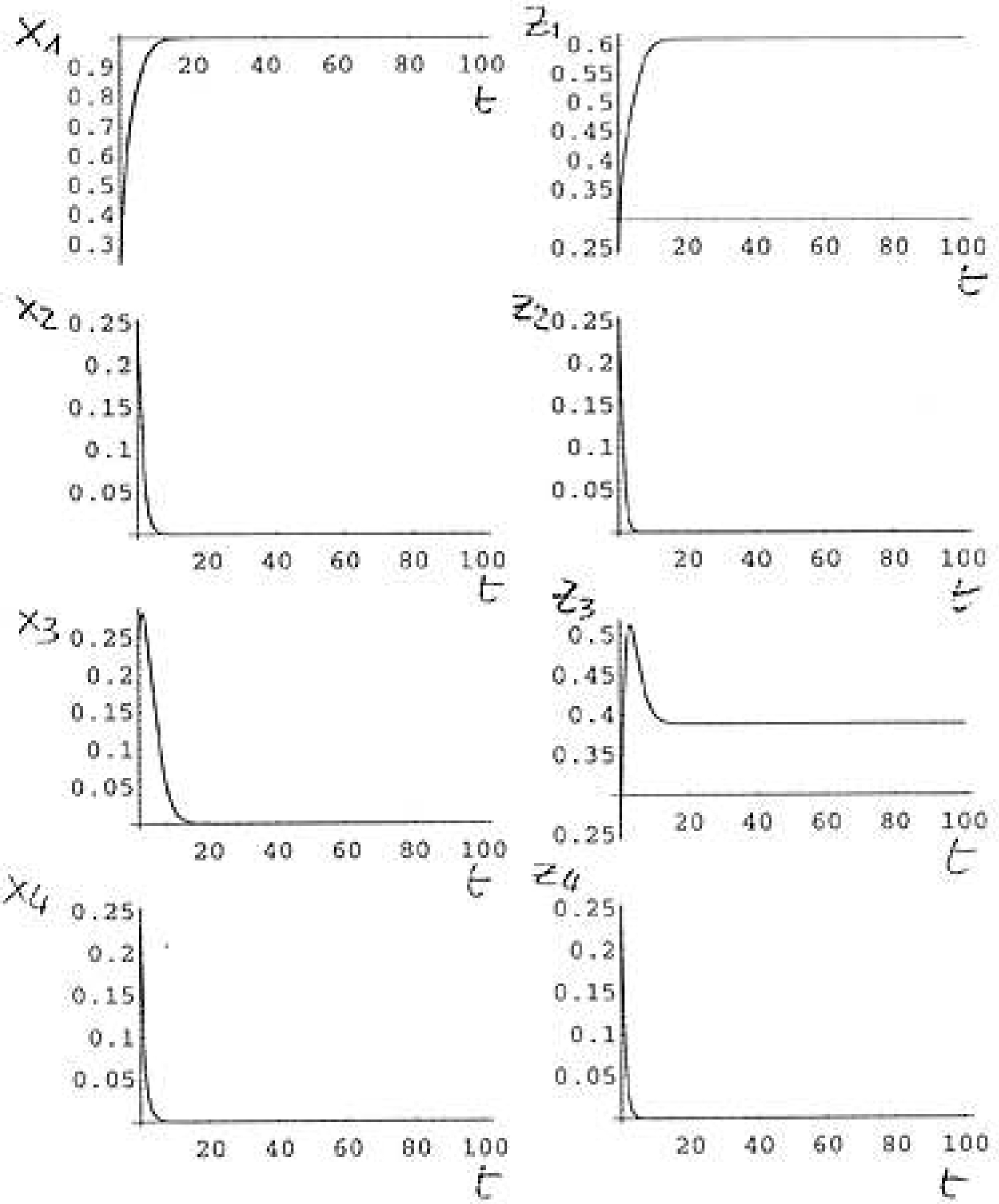

x1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20 40 60 80 100
t
z1
0.6
0.55
0.5
0.45
0.4
0.35
0.25
20 40 60 80 100
t
x2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x3
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z3
0.5
0.45
0.4
0.35
0.25
20 40 60 80 100
t
x4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t

fig. 1b

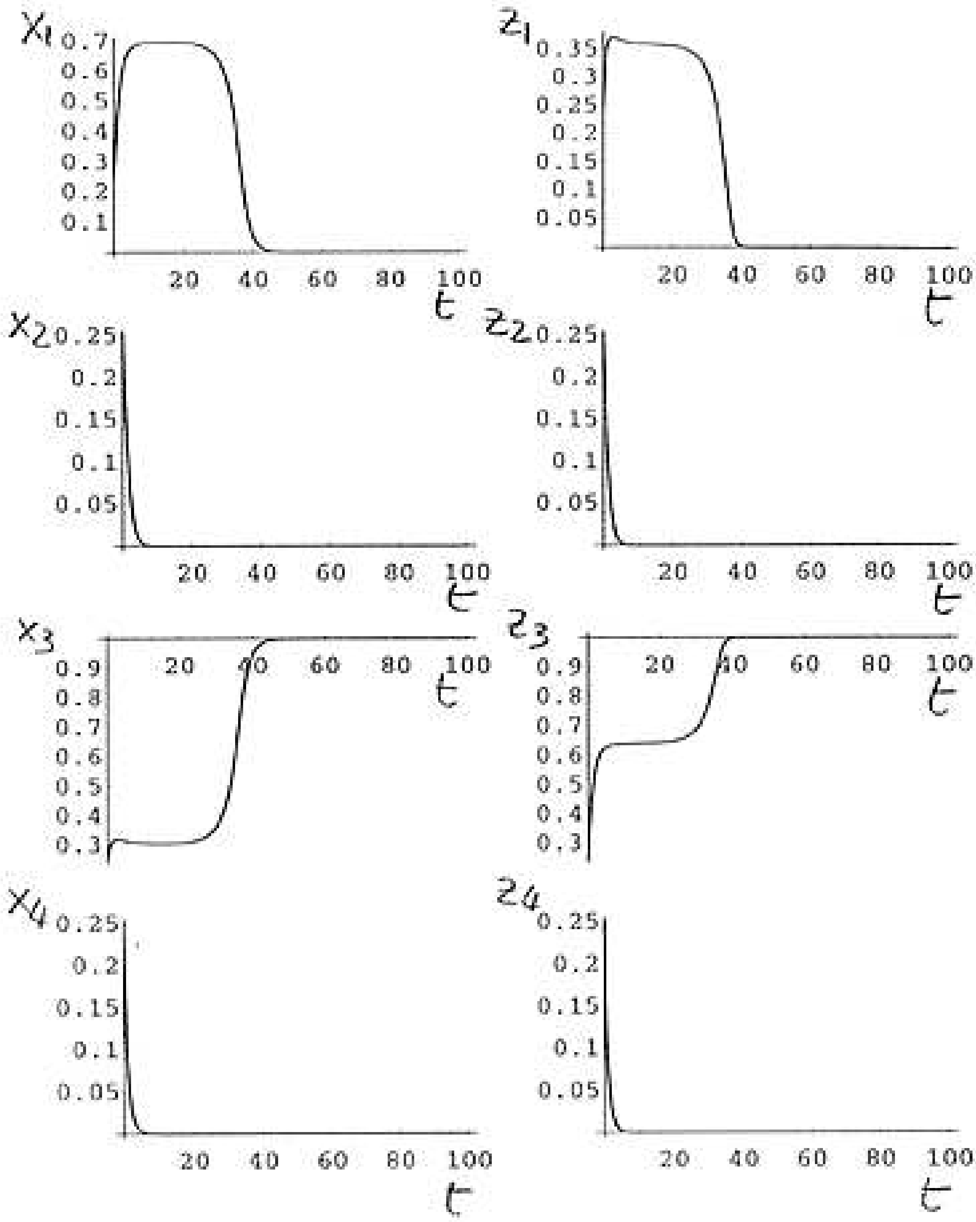

x1
0.7
0.6
0.5
0.4
0.3
0.2
0.1
20
40
60
80
100
t
z1
0.35
0.3
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
x2
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
z2
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
x3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20
40
60
80
100
t
z3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20
40
60
80
100
t
x4
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
z4
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t

fig. 1c

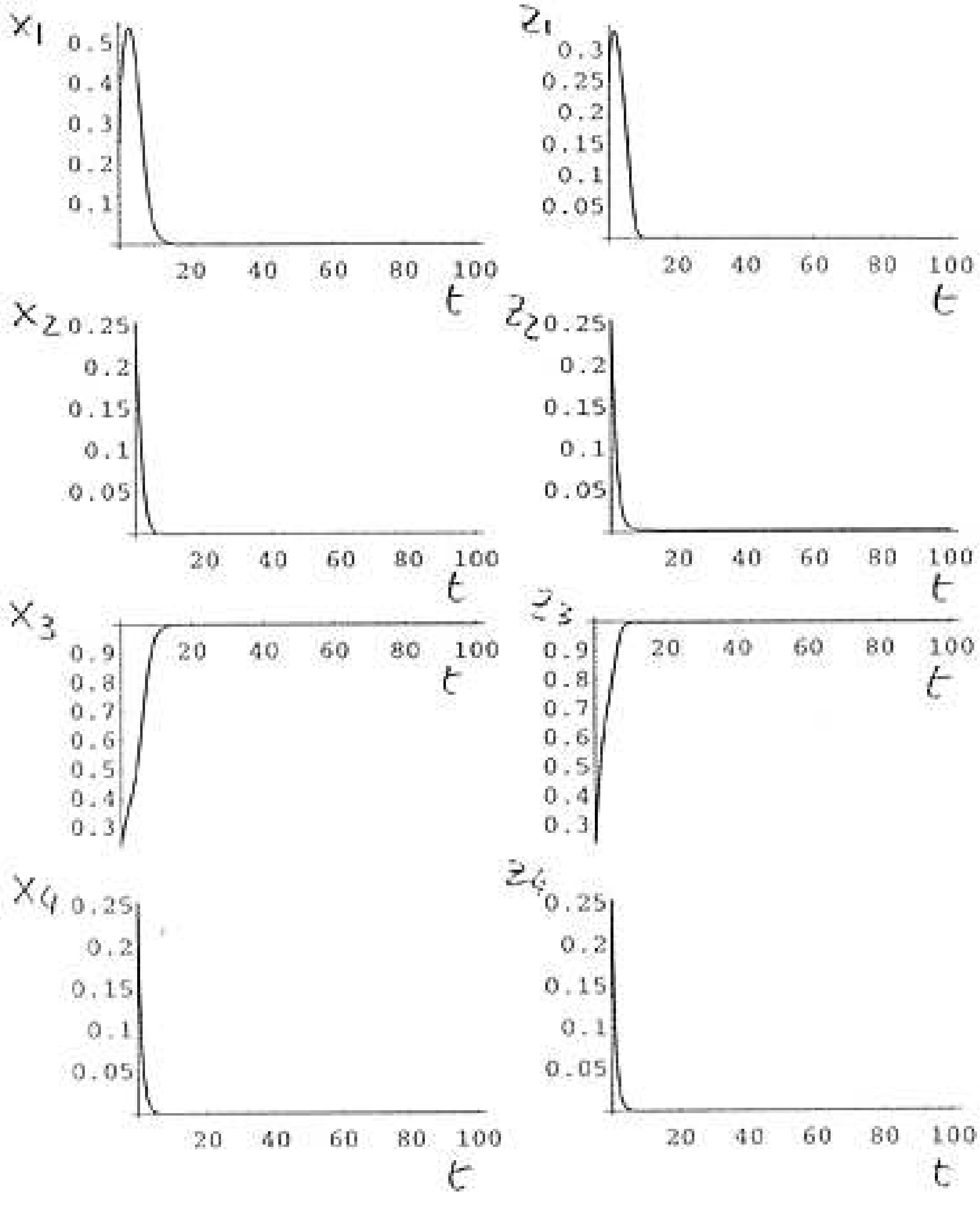

x1
0.5
0.4
0.3
0.2
0.1
20 40 60 80 100
t
z1
0.3
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20 40 60 80 100
t
z3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20 40 60 80 100
t
x4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t

fig. 1d

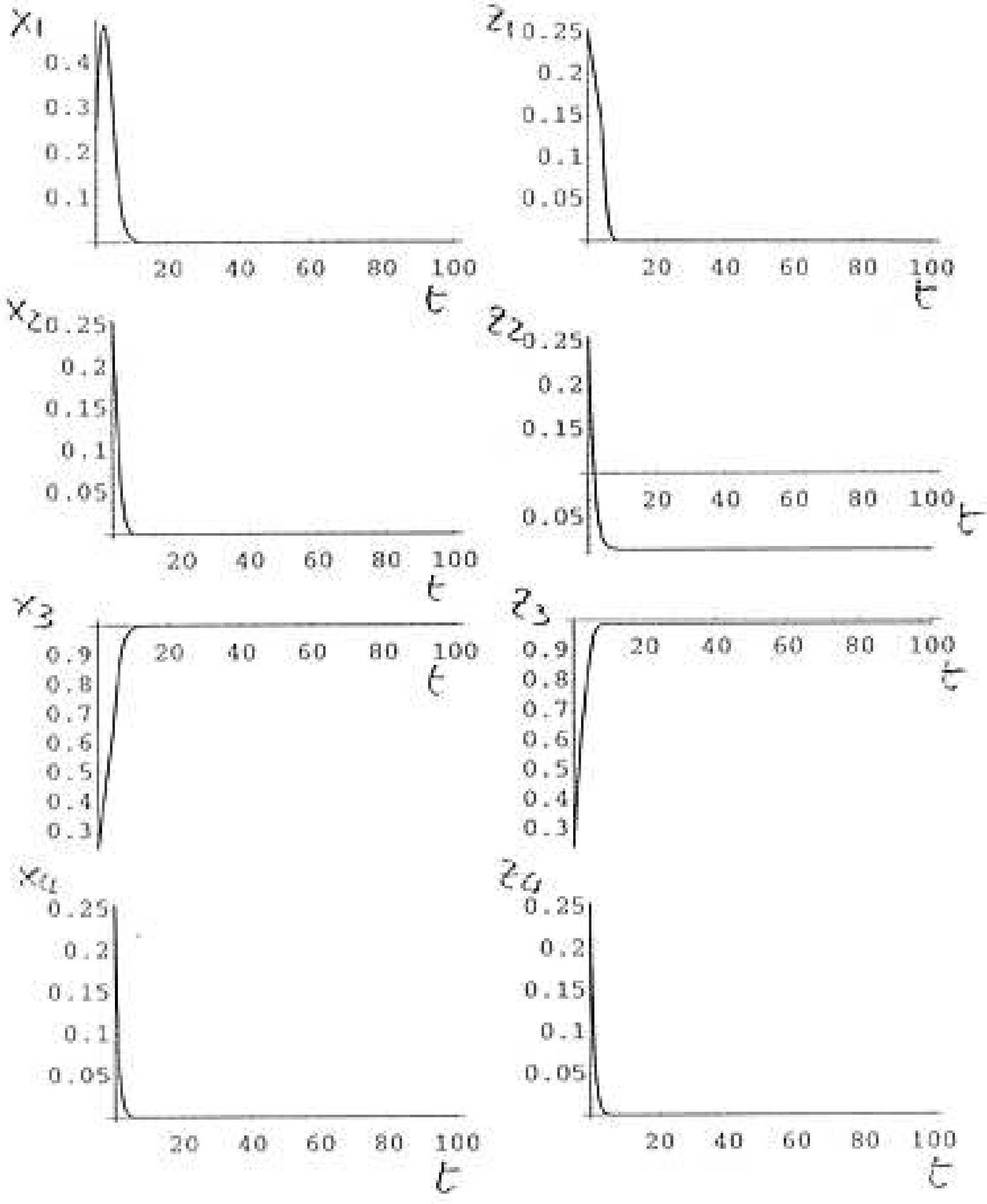
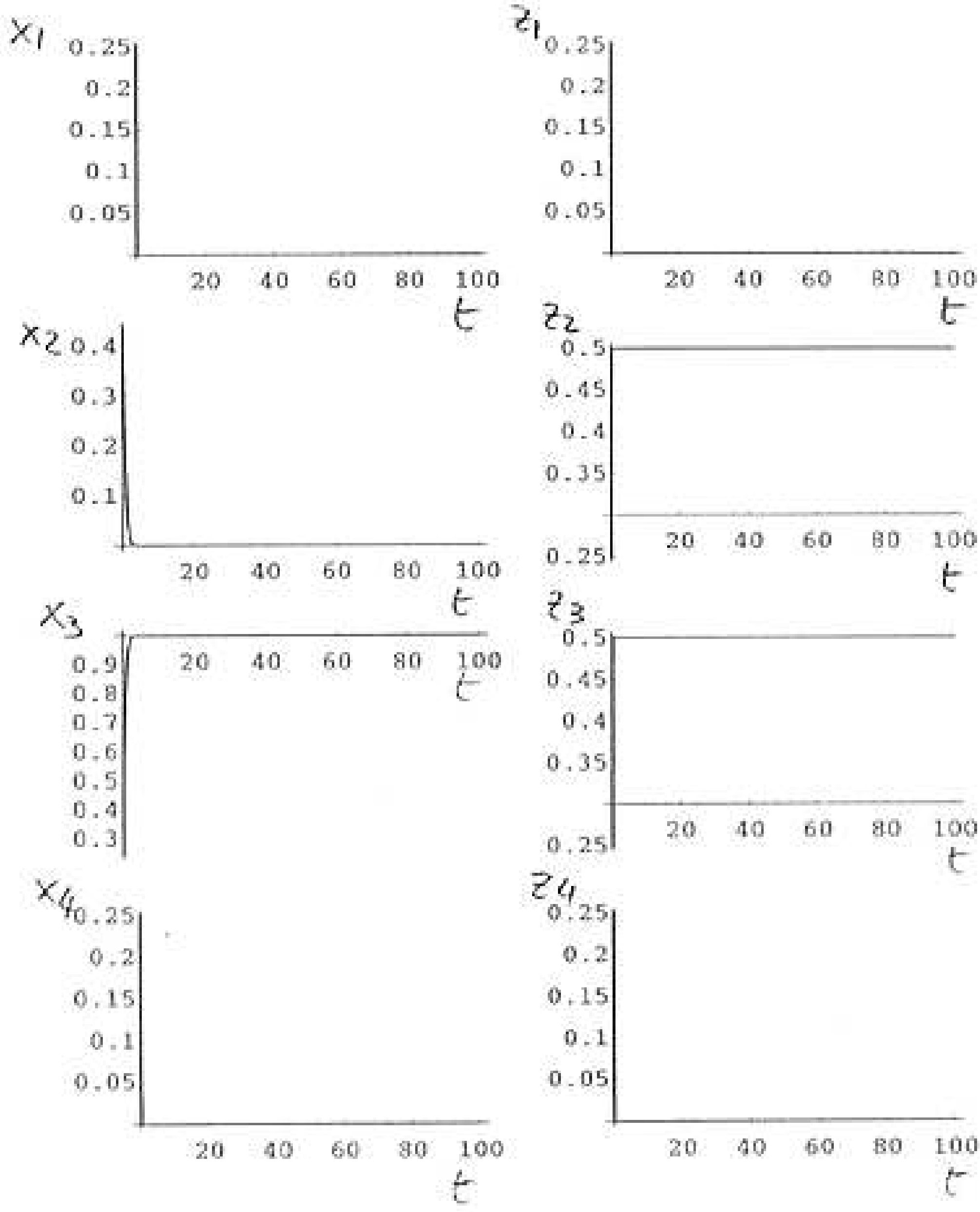
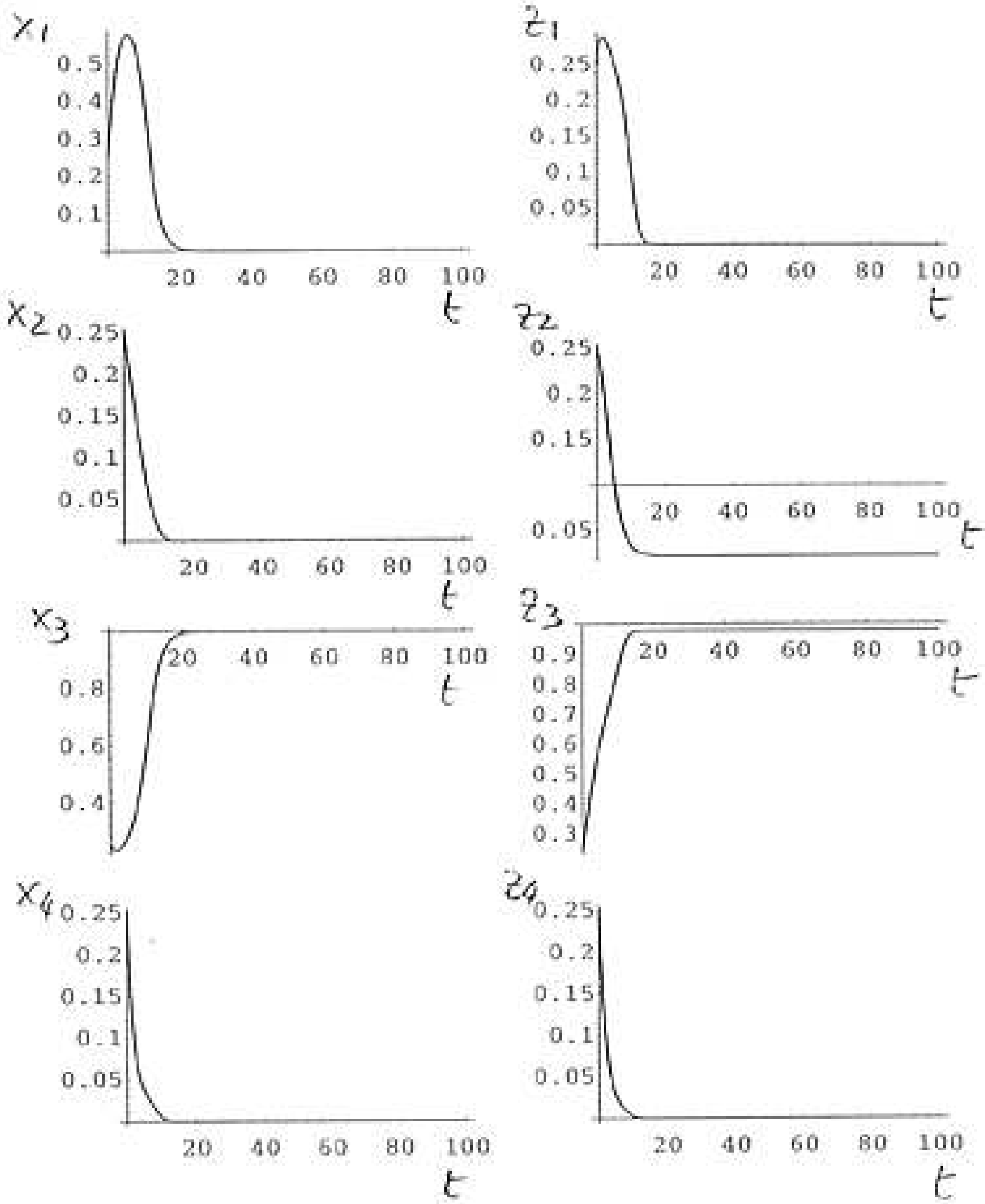


fig. 1e



x1
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z1
x2
0.4
0.3
0.2
0.1
z2
0.5
0.45
0.4
0.35
0.25
x3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
z3
x4
z4

fig. 2a



x1
0.5
0.4
0.3
0.2
0.1
20
40
60
80
100
t
z1
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
x2
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
z2
0.25
0.2
0.15
20
40
60
80
100
t
0.05
x3
20
40
60
80
100
t
0.8
0.6
0.4
z3
0.9
20
40
60
80
100
t
0.8
0.7
0.6
0.5
0.4
0.3
x4
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
z4
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t

fig. 2b

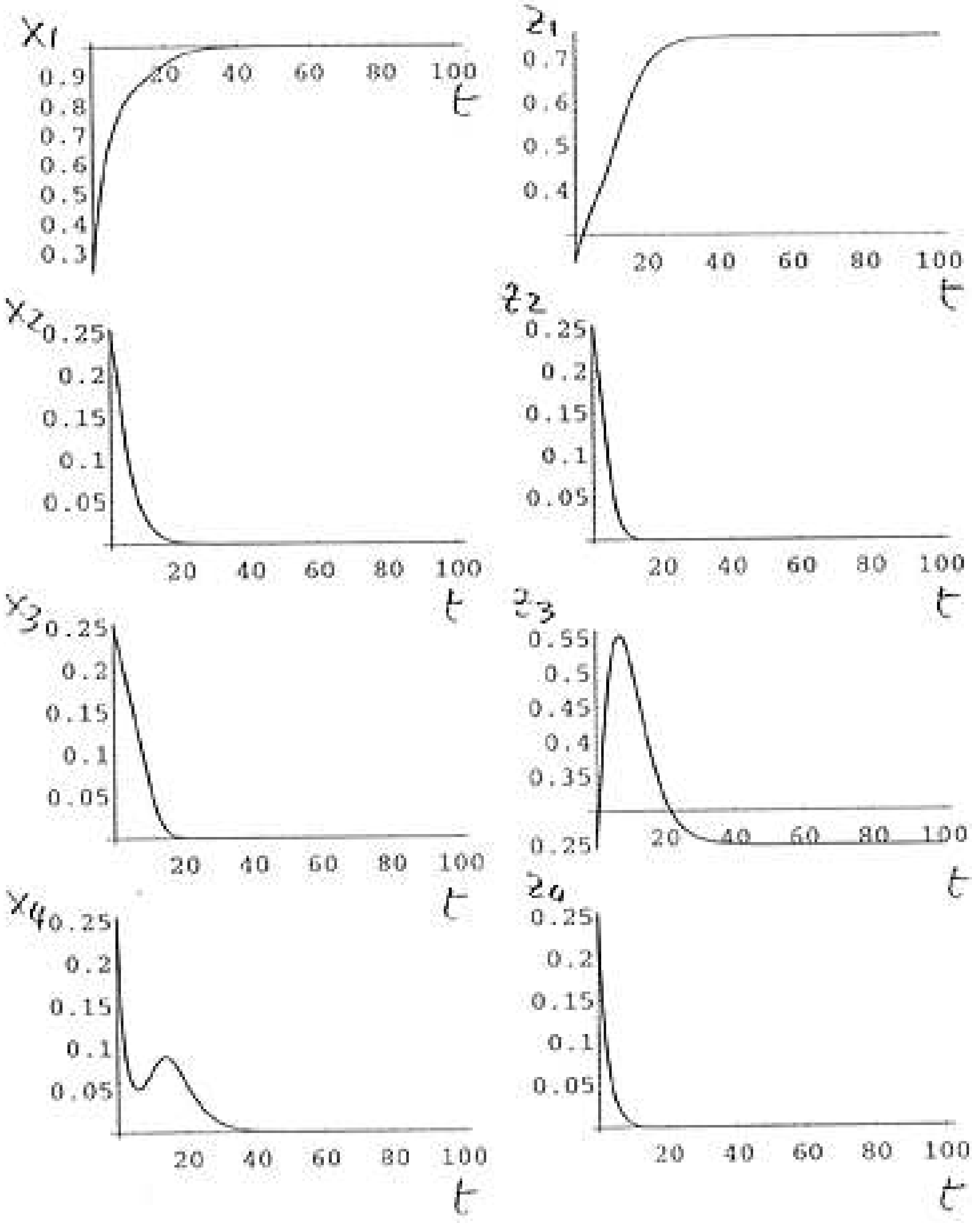
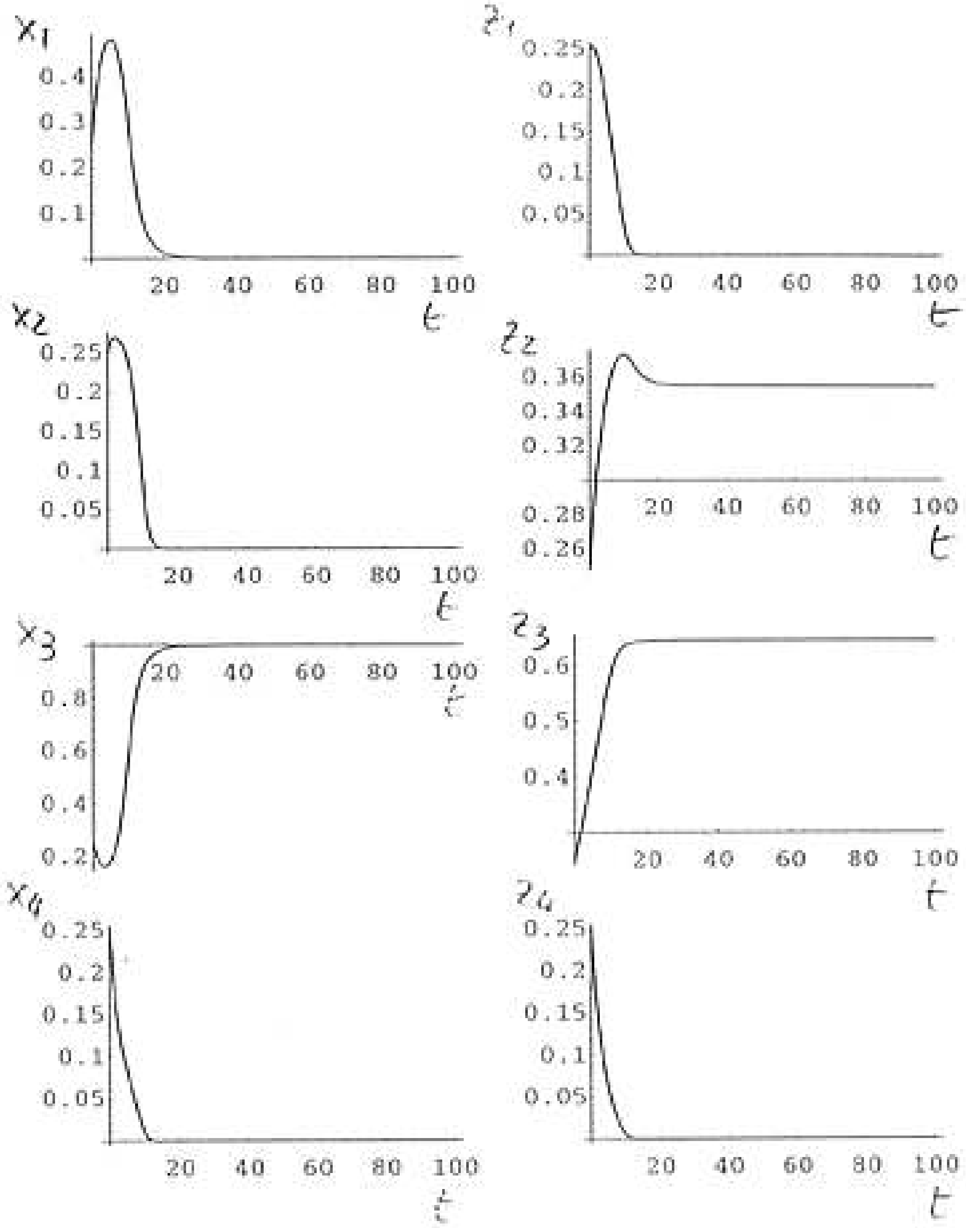


fig. 2c



x1
0.4
0.3
0.2
0.1
20 40 60 80 100
t
z1
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z2
0.36
0.34
0.32
0.28
0.26
20 40 60 80 100
t
x3
0.8
0.6
0.4
0.2
20 40 60 80 100
t
z3
0.6
0.5
0.4
20 40 60 80 100
t
x4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t

fig. 2d

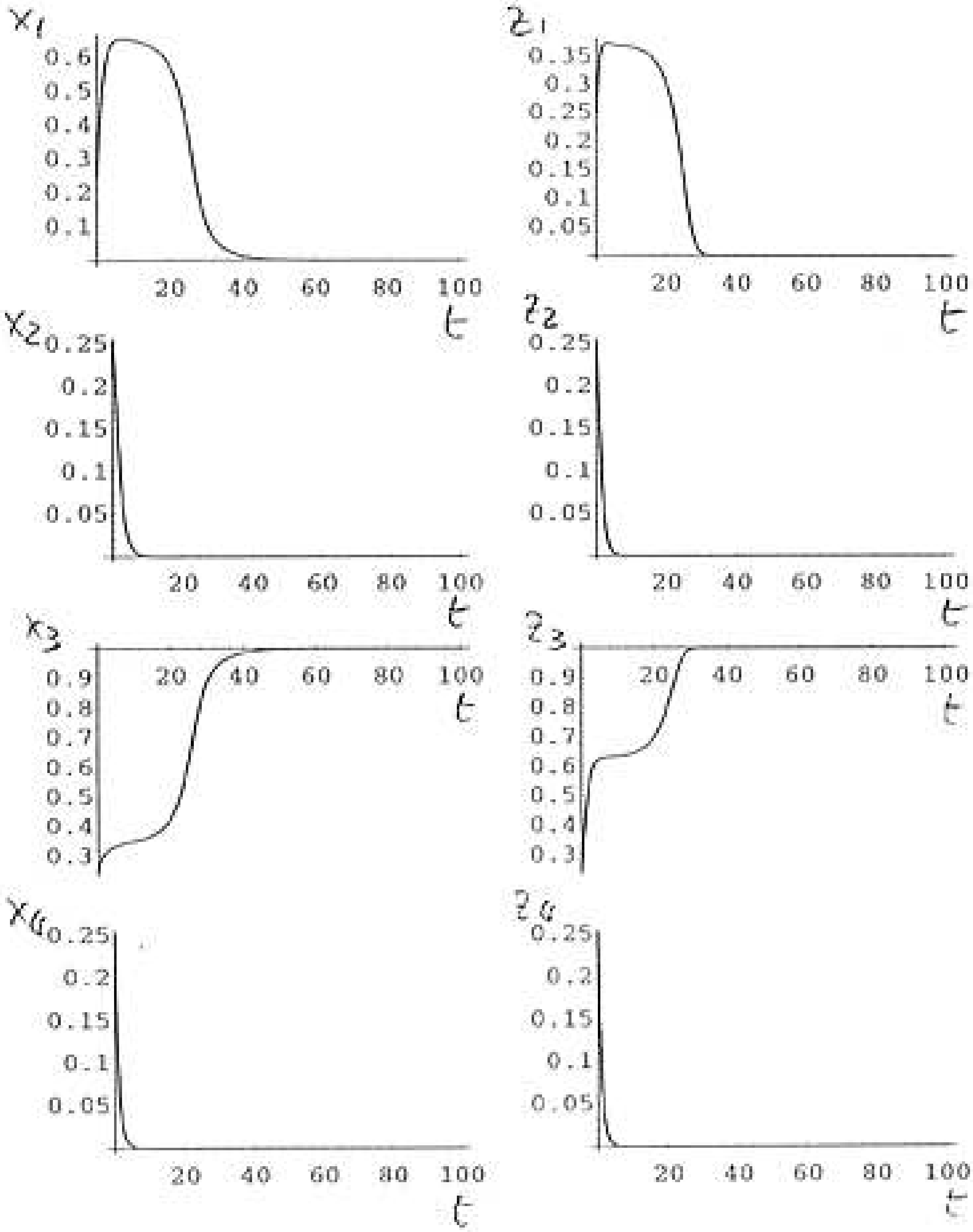

x1
0.6
0.5
0.4
0.3
0.2
0.1
20 40 60 80 100
t
z1
0.35
0.3
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20 40 60 80 100
t
z3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20 40 60 80 100
t
x4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t

fig. 2e

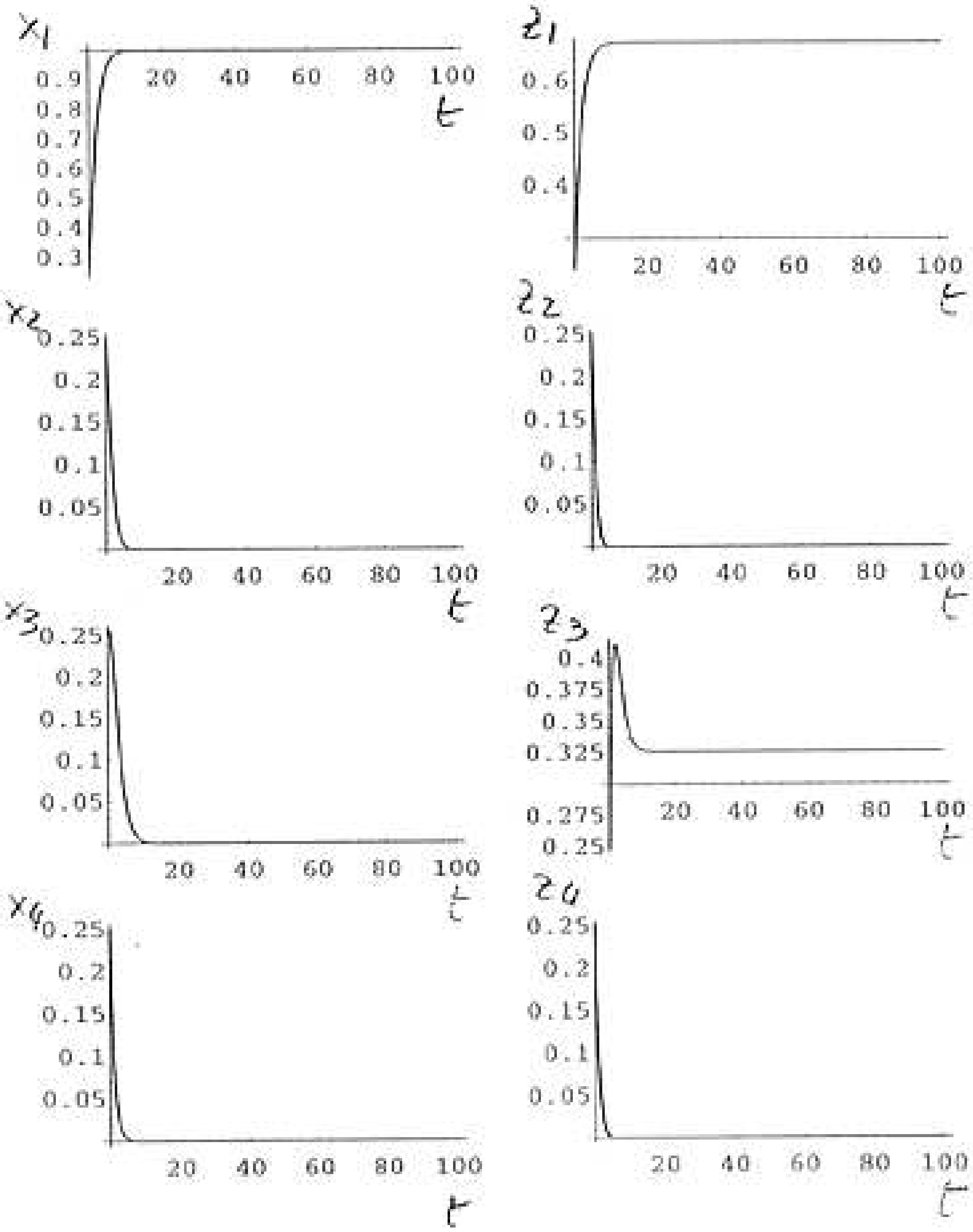

x1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20 40 60 80 100
t
z1
0.6
0.5
0.4
20 40 60 80 100
t
x2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z2
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x3
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z3
0.4
0.375
0.35
0.325
0.275
0.25
20 40 60 80 100
t
x4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t

fig. 2f

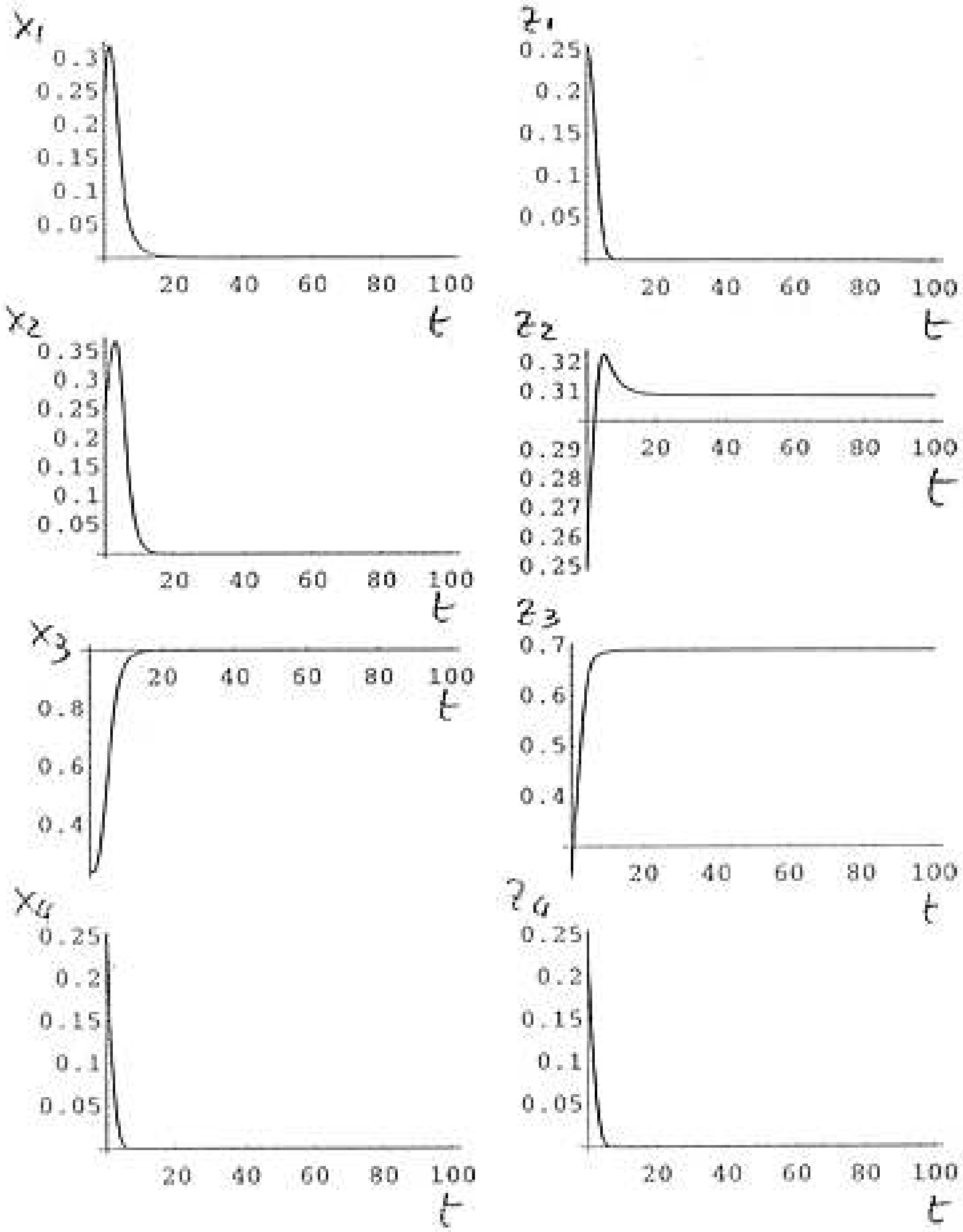
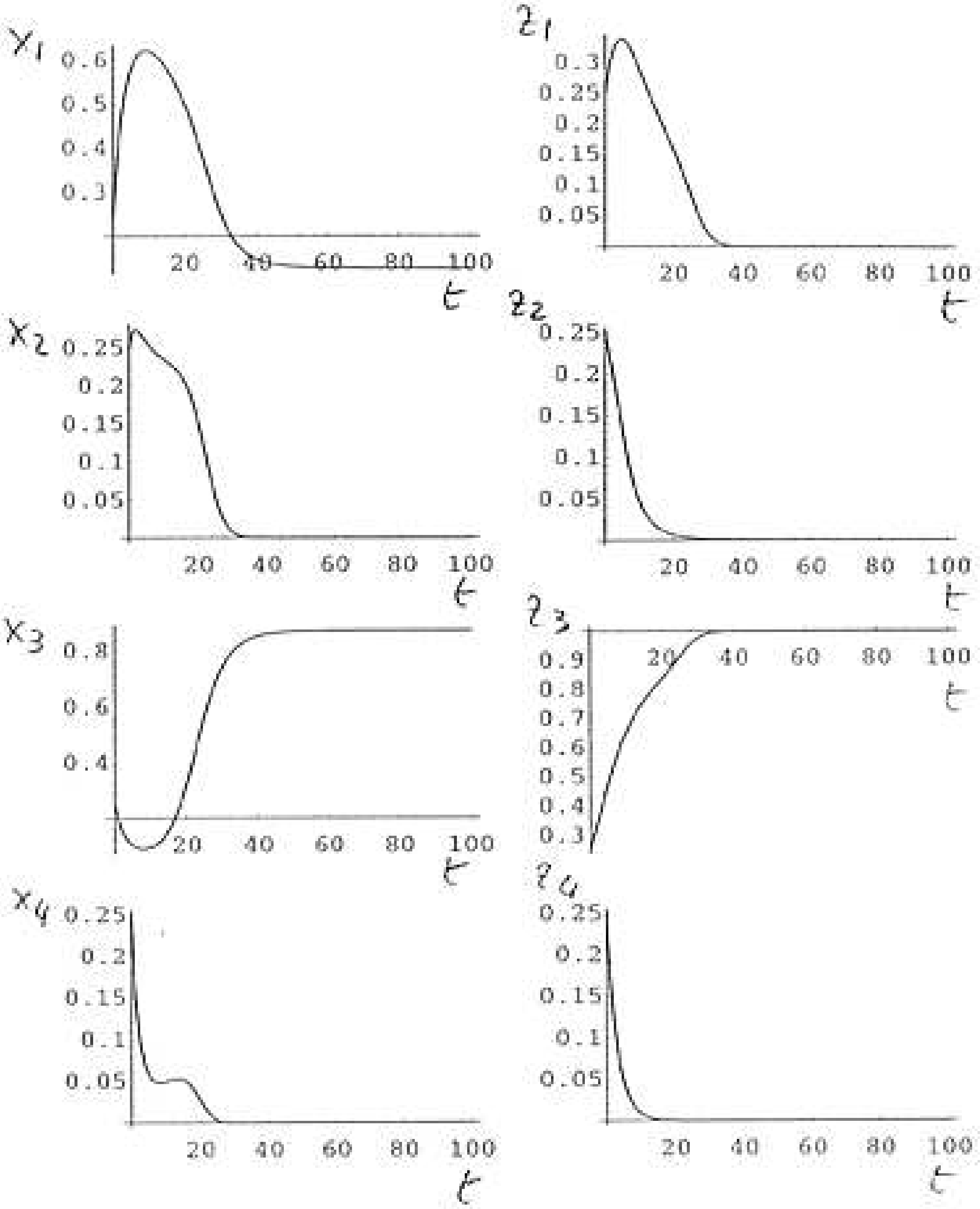

x1
0.3
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z1
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
x2
0.35
0.3
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z2
0.32
0.31
0.29
0.28
0.27
0.26
0.25
20 40 60 80 100
t
x3
20 40 60 80 100
t
0.8
0.6
0.4
z3
0.7
0.6
0.5
0.4
20 40 60 80 100
t
x4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t
z4
0.25
0.2
0.15
0.1
0.05
20 40 60 80 100
t

fig. 2g



x1
0.6
0.5
0.4
0.3
20
40
60
80
100
t
z1
0.3
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
x2
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
z2
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
x3
0.8
0.6
0.4
20
40
60
80
100
t
z3
0.9
0.8
0.7
0.6
0.5
0.4
0.3
20
40
60
80
100
t
x4
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t
z4
0.25
0.2
0.15
0.1
0.05
20
40
60
80
100
t

fig. 3a

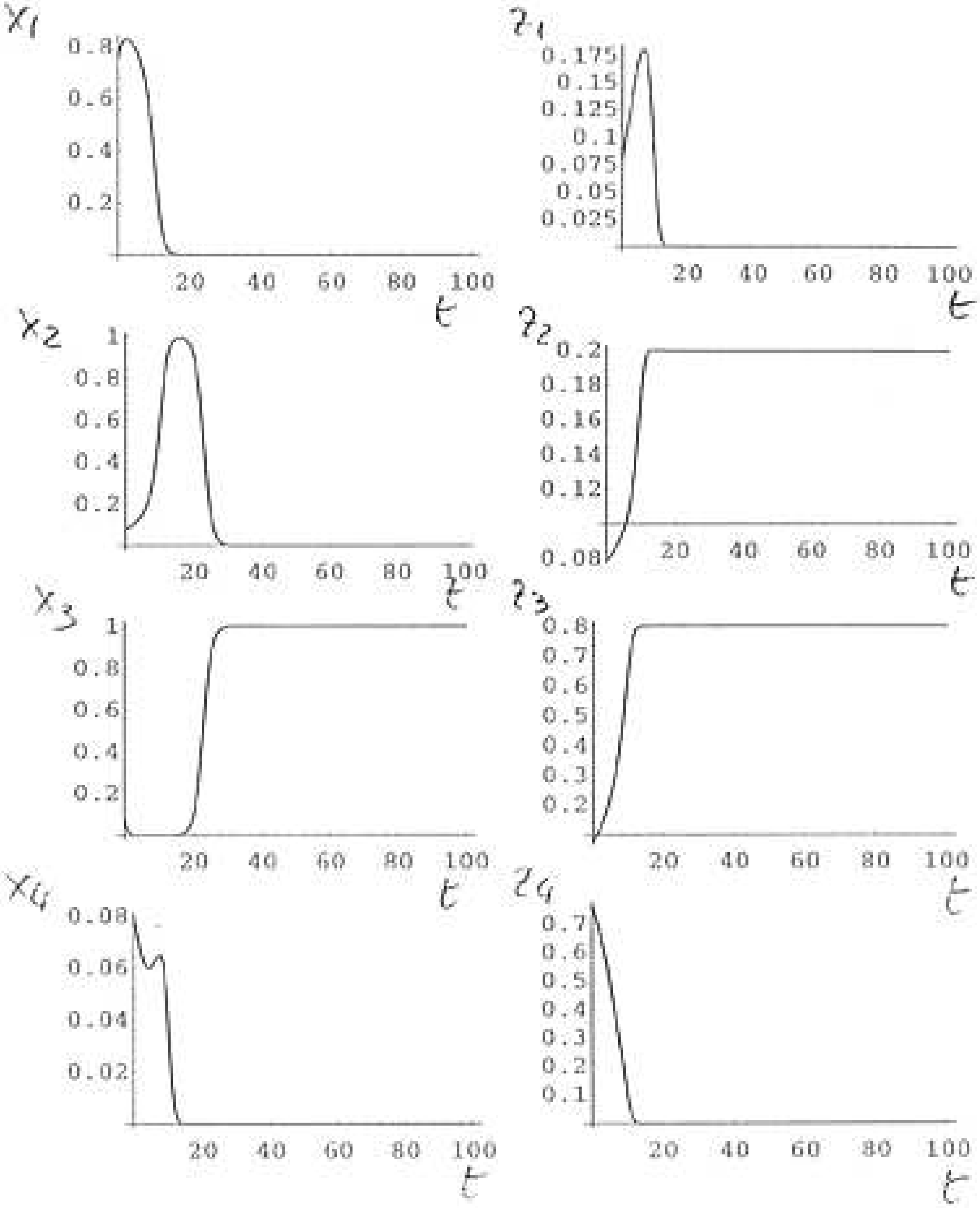
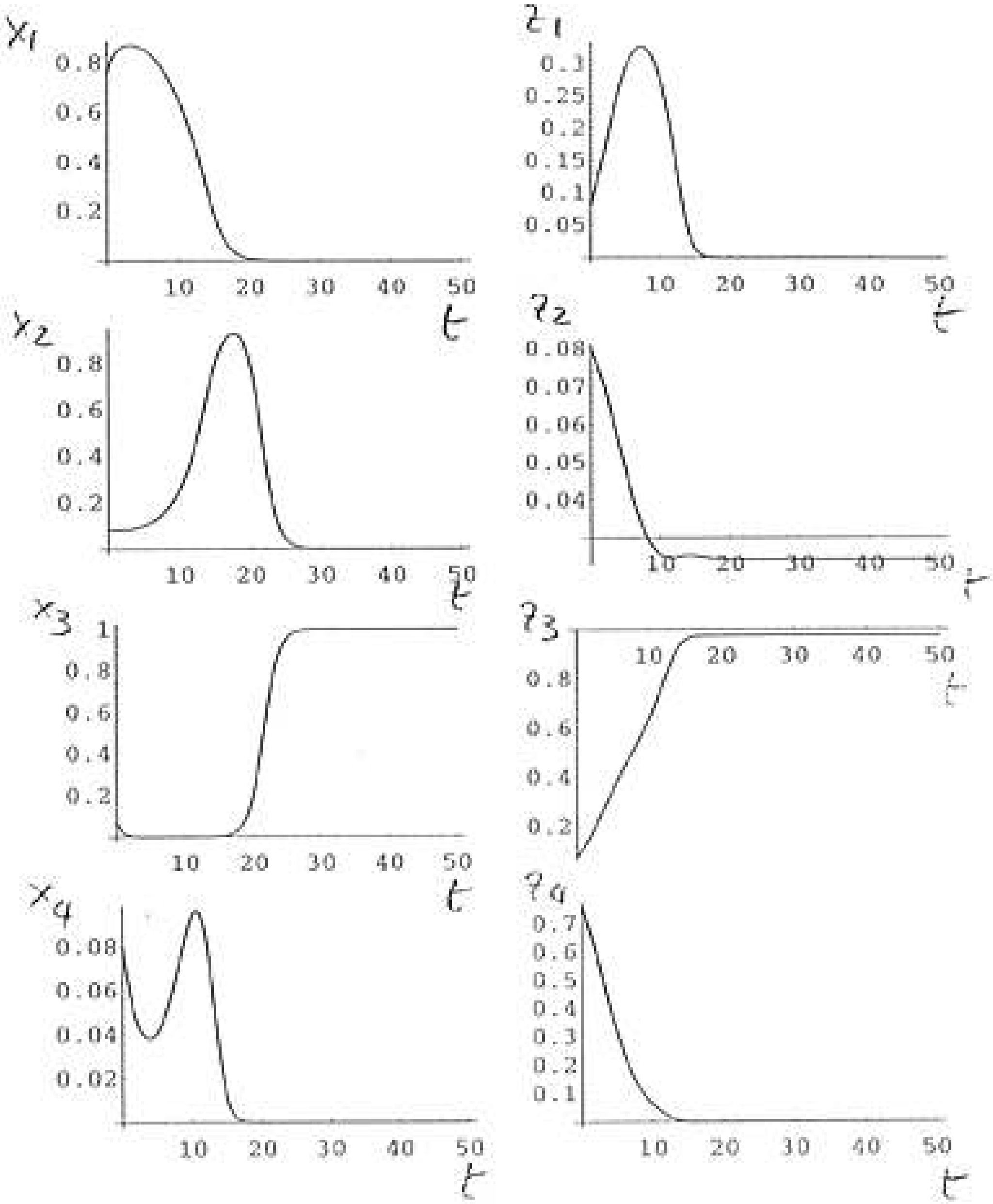
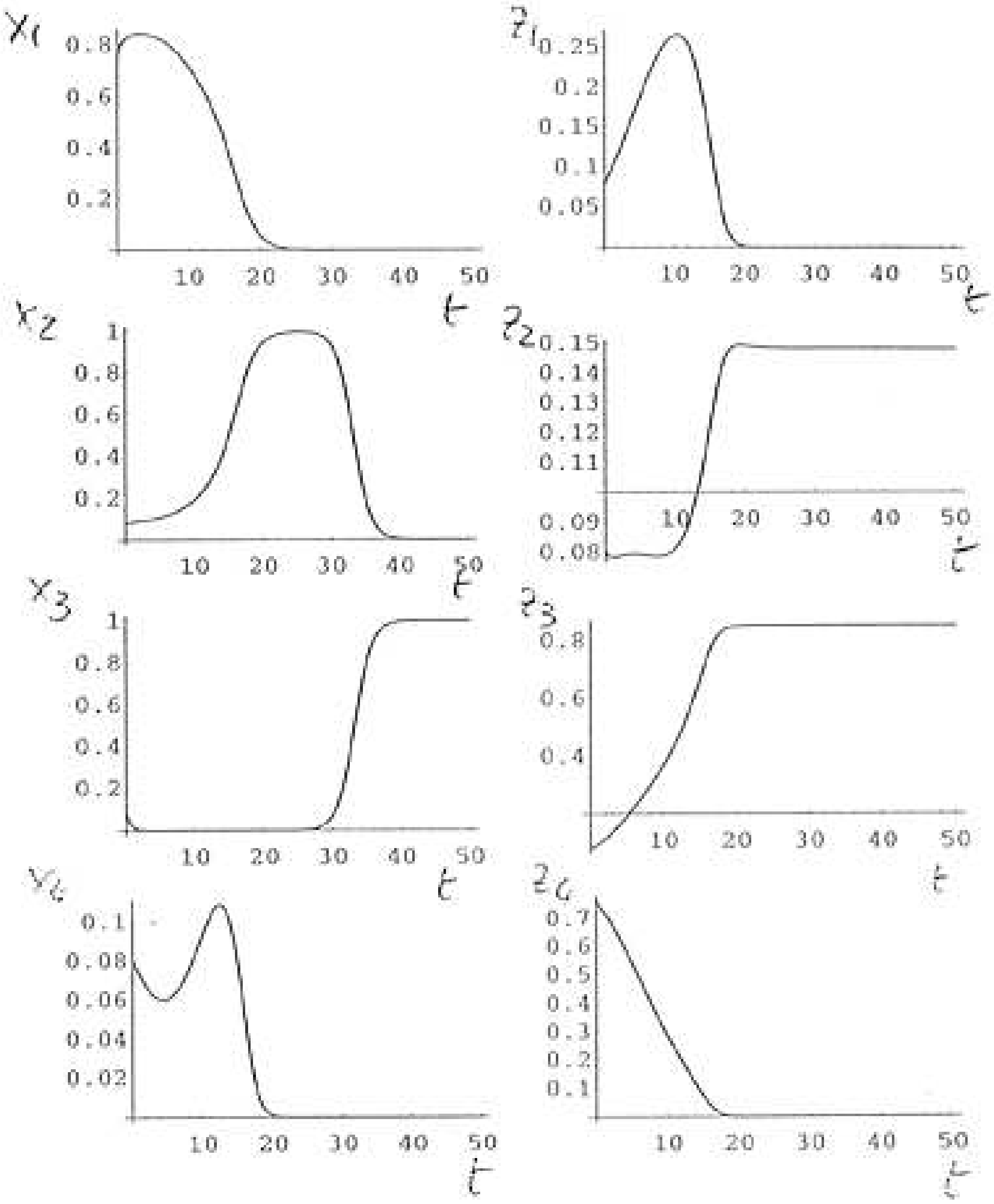


fig. 3b



x1
0.8
0.6
0.4
0.2
10 20 30 40 50
t
x2
0.8
0.6
0.4
0.2
10 20 30 40 50
t
x3
1
0.8
0.6
0.4
0.2
10 20 30 40 50
t
x4
0.08
0.06
0.04
0.02
10 20 30 40 50
t
z1
0.3
0.25
0.2
0.15
0.1
0.05
10 20 30 40 50
t
0.08
0.07
0.06
0.05
0.04
10 20 30 40 50
t
0.8
0.6
0.4
0.2
10 20 30 40 50
t
0.7
0.6
0.5
0.4
0.3
0.2
0.1
10 20 30 40 50
t

fig. 3c



x1
0.8
0.6
0.4
0.2
10
20
30
40
50
t
z1
0.25
0.2
0.15
0.1
0.05
10
20
30
40
50
t
x2
1
0.8
0.6
0.4
0.2
10
20
30
40
50
t
z2
0.15
0.14
0.13
0.12
0.11
0.09
0.08
10
20
30
40
50
t
x3
1
0.8
0.6
0.4
0.2
10
20
30
40
50
t
z3
0.8
0.6
0.4
10
20
30
40
50
t
x4
0.1
0.08
0.06
0.04
0.02
10
20
30
40
50
t
z4
0.7
0.6
0.5
0.4
0.3
0.2
0.1
10
20
30
40
50
t

fig. 3d

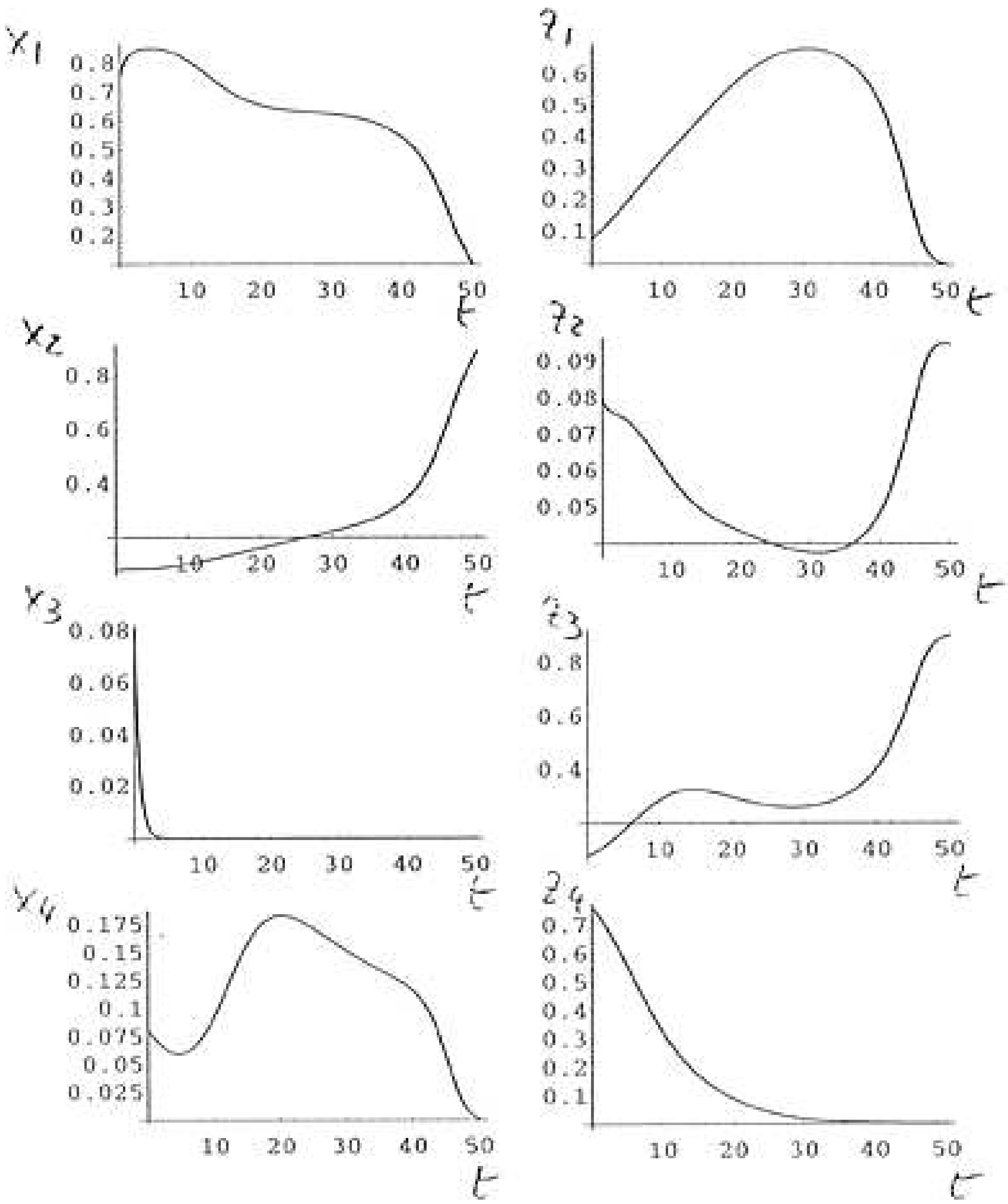

0.8 0.7 0.6 0.5 0.4 0.3 0.2
10 20 30 40 50
0.6 0.5 0.4 0.3 0.2 0.1
10 20 30 40 50
0.8 0.6 0.4
10 20 30 40 50
0.09 0.08 0.07 0.06 0.05
10 20 30 40 50
0.08 0.06 0.04 0.02
10 20 30 40 50
0.8 0.6 0.4
10 20 30 40 50
0.175 0.15 0.125 0.1 0.075 0.05 0.025
10 20 30 40 50
0.7 0.6 0.5 0.4 0.3 0.2 0.1
10 20 30 40 50

fig. 3e

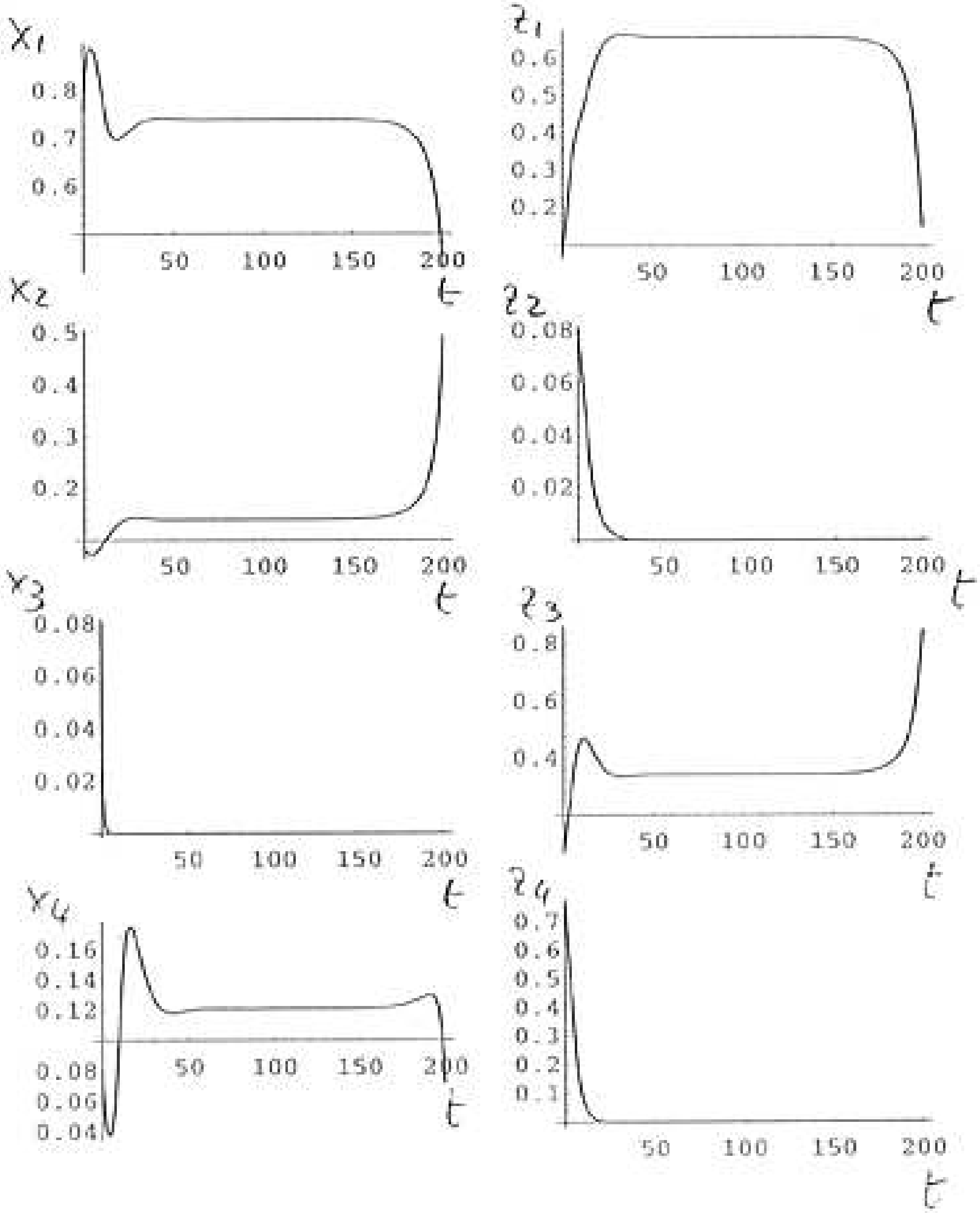

x1
0.8
0.7
0.6
50 100 150 200
t
z1
0.6
0.5
0.4
0.3
0.2
50 100 150 200
t
x2
0.5
0.4
0.3
0.2
50 100 150 200
t
z2
0.08
0.06
0.04
0.02
50 100 150 200
t
x3
0.08
0.06
0.04
0.02
50 100 150 200
t
z3
0.8
0.6
0.4
50 100 150 200
t
x4
0.16
0.14
0.12
0.08
0.06
0.04
50 100 150 200
t
z4
0.7
0.6
0.5
0.4
0.3
0.2
0.1
50 100 150 200
t

fig. 3f

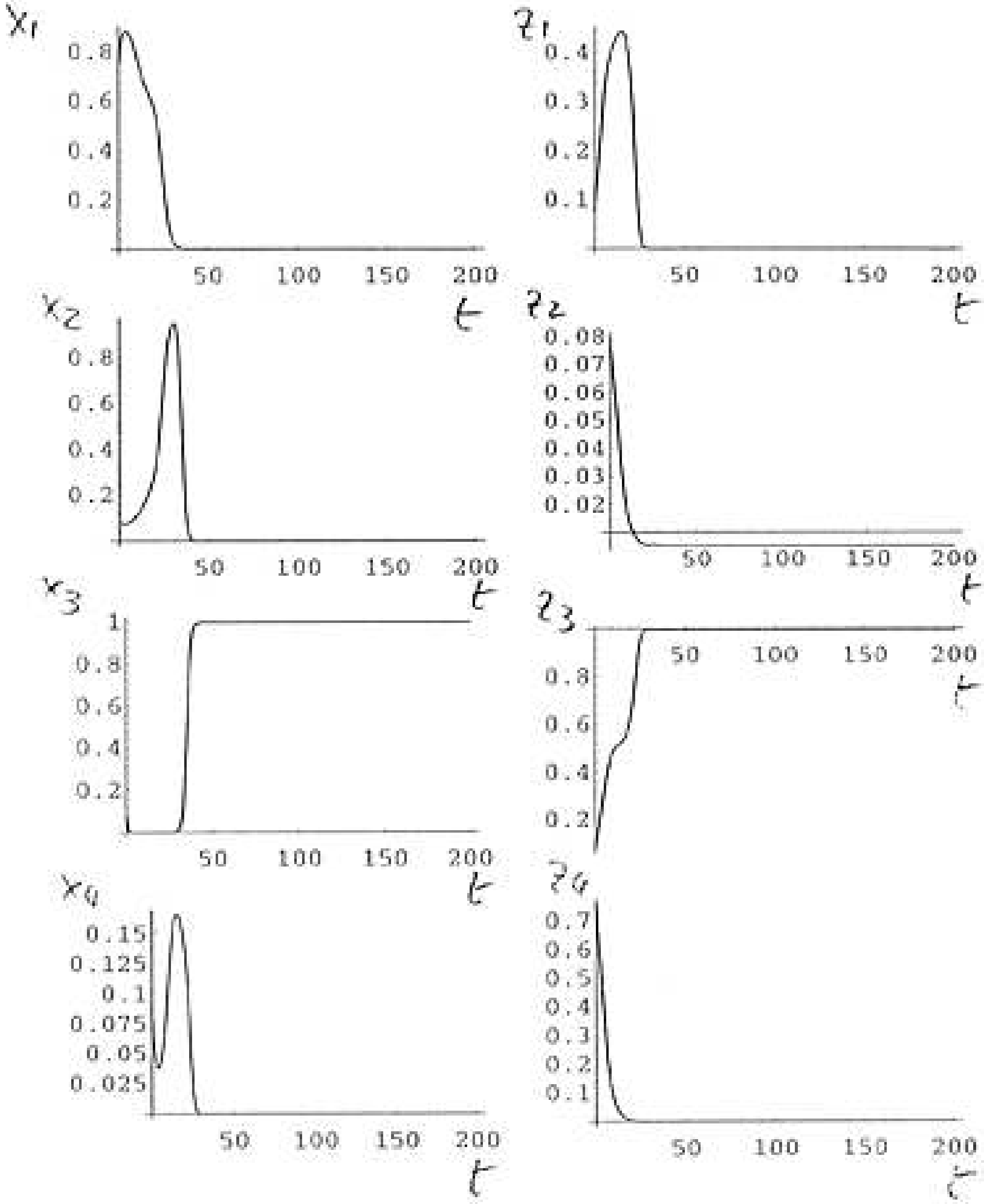

x1
0.8
0.6
0.4
0.2
50 100 150 200
t
z1
0.4
0.3
0.2
0.1
50 100 150 200
t
x2
0.8
0.6
0.4
0.2
50 100 150 200
t
z2
0.08
0.07
0.06
0.05
0.04
0.03
0.02
50 100 150 200
t
x3
1
0.8
0.6
0.4
0.2
50 100 150 200
t
z3
50 100 150 200
t
0.8
0.6
0.4
0.2
x4
0.15
0.125
0.1
0.075
0.05
0.025
50 100 150 200
t
z4
0.7
0.6
0.5
0.4
0.3
0.2
0.1
50 100 150 200
t

fig. 3g

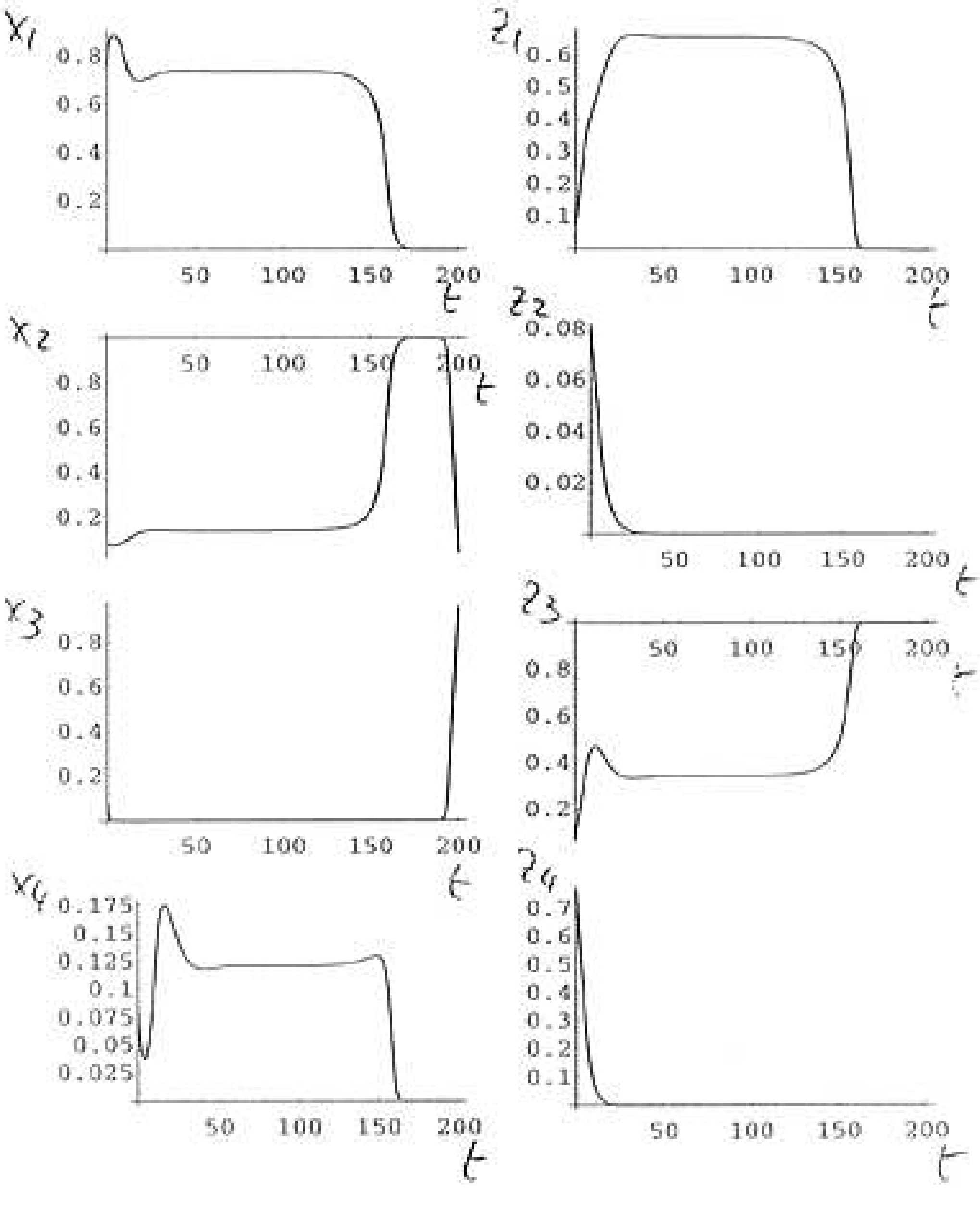

x1
0.8
0.6
0.4
0.2
50 100 150 200
t
z1
0.6
0.5
0.4
0.3
0.2
0.1
50 100 150 200
t
x2
0.8
0.6
0.4
0.2
50 100 150 200
t
z2
0.08
0.06
0.04
0.02
50 100 150 200
t
x3
0.8
0.6
0.4
0.2
50 100 150 200
t
z3
0.8
0.6
0.4
0.2
50 100 150 200
x4
0.175
0.15
0.125
0.1
0.075
0.05
0.025
50 100 150 200
t
z4
0.7
0.6
0.5
0.4
0.3
0.2
0.1
50 100 150 200
t

fig. 3h

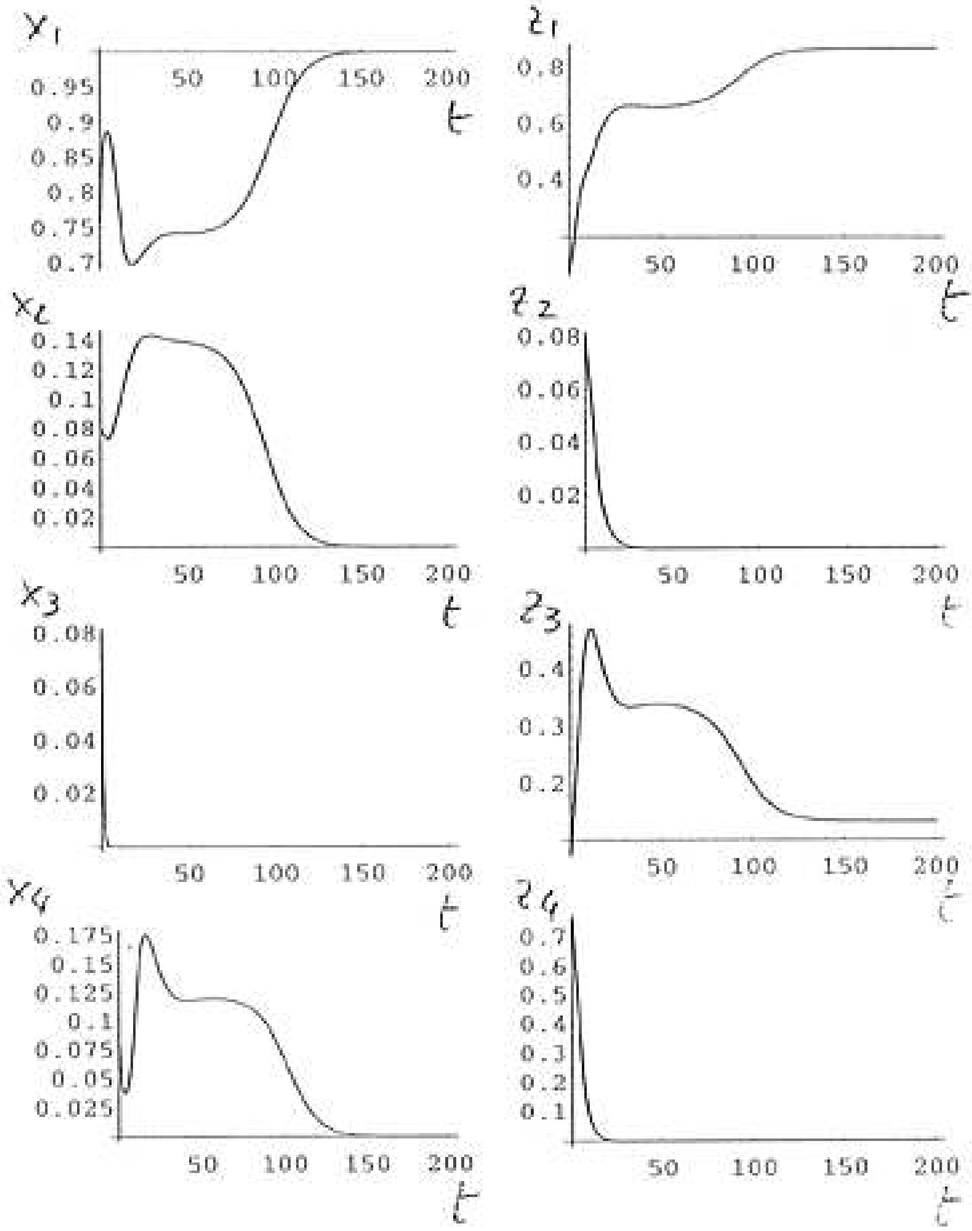


fig. 3i

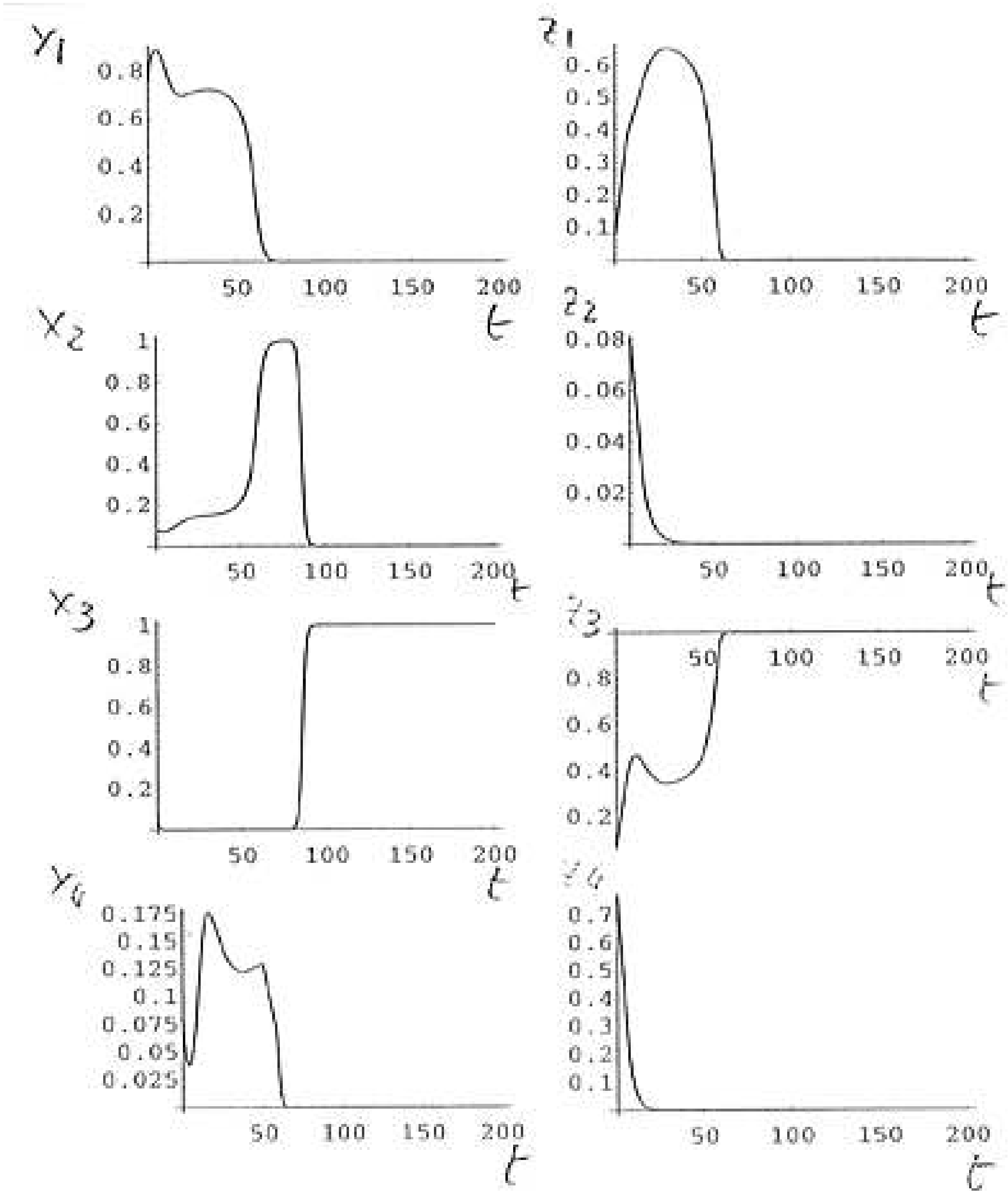
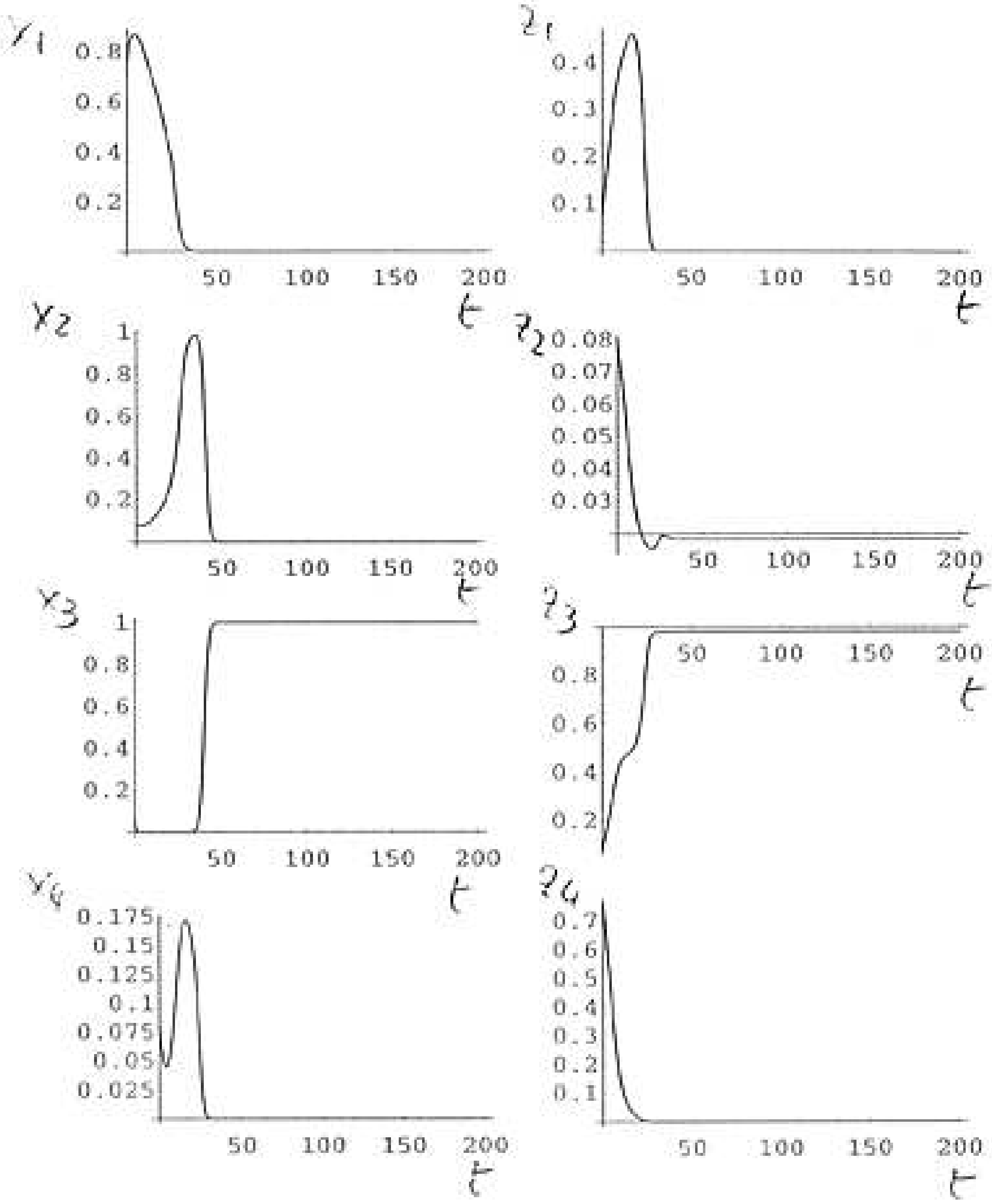
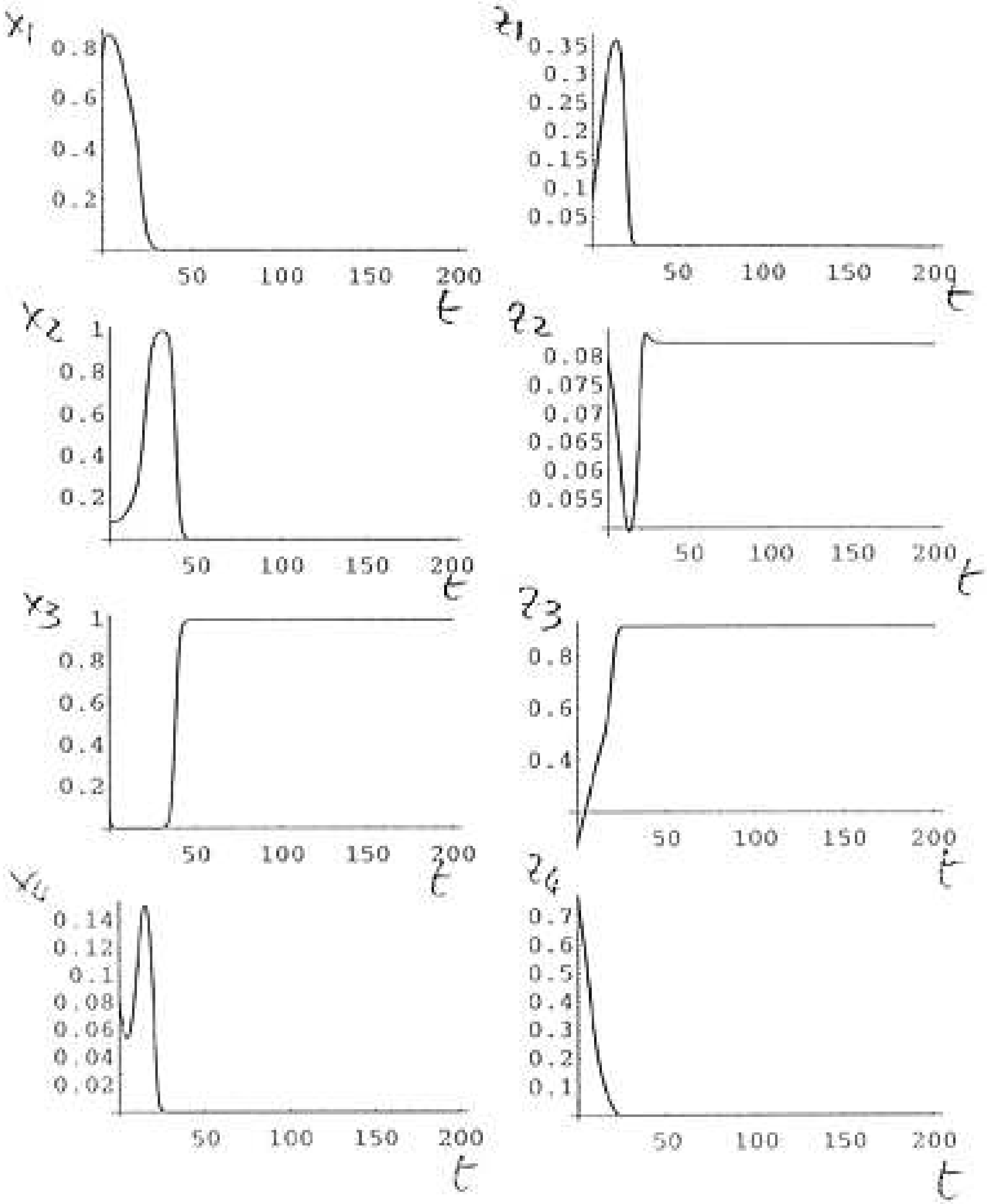


fig. 3l



y1
0.8
0.6
0.4
0.2
50 100 150 200
t
z1
0.4
0.3
0.2
0.1
50 100 150 200
t
y2
1
0.8
0.6
0.4
0.2
50 100 150 200
t
z2
0.08
0.07
0.06
0.05
0.04
0.03
50 100 150 200
t
y3
1
0.8
0.6
0.4
0.2
50 100 150 200
t
z3
50 100 150 200
0.8
0.6
0.4
0.2
t
y4
0.175
0.15
0.125
0.1
0.075
0.05
0.025
50 100 150 200
t
z4
0.7
0.6
0.5
0.4
0.3
0.2
0.1
50 100 150 200
t

fig. 3m



x1
0.8
0.6
0.4
0.2
50
100
150
200
t
z1
0.35
0.3
0.25
0.2
0.15
0.1
0.05
50
100
150
200
t
x2
1
0.8
0.6
0.4
0.2
50
100
150
200
t
z2
0.08
0.075
0.07
0.065
0.06
0.055
50
100
150
200
t
x3
1
0.8
0.6
0.4
0.2
50
100
150
200
t
z3
0.8
0.6
0.4
50
100
150
200
t
x4
0.14
0.12
0.1
0.08
0.06
0.04
0.02
50
100
150
200
t
z4
0.7
0.6
0.5
0.4
0.3
0.2
0.1
50
100
150
200
t