Robust Social Norms in Bargains and Markets

Eric J. Friedman*
Department of Economics, Rutgers University
New Brunswick, NJ 08903.

March 30, 1999

Abstract

We consider the social norms of repeated matching games in the presence of finite probability trembles and show that such norms must be subgame perfect along the equilibrium path but need not be subgame perfect off the equilibrium path.

This is consistent with the well known experimental results by Roth et. al. (1991) in which subjects play subgame perfect equilibria in the market game but play non-subgame perfect equilibria in the ultimatum game, providing a simple and intuitive explanation for this behavior in terms of societal norms, where societal norms are simply the dominant play induced by the sequential equilibria arising in societal games.

Our analysis provides a fully rational explanation of behavior that has been typically analyzed as arising from boundedly rational play. It also emphasizes the importance of studying the effects of finite tremble probabilities and population sizes directly, for example *both* limits, very large, and very small populations yield misleading predictions for the intermediate case.

*email: friedman@econ.rutgers.edu, www: http://econ.rutgers.edu/home/friedman

1 Introduction

Experiments in extensive form games have led to many results which challenge the foundations of rational game theoretic analysis. Play often contradicts the standard theory and many authors have attempted to create a framework in which to understand these results.

For example, Roth et. al. (1991) ran a series of experiments on two related games: the ultimatum game and a market game. Interestingly, the market game shows robust and rapid convergence to the subgame perfect outcome, while in the ultimatum game non subgame perfect outcomes seem to be robust. These results for the ultimatum game have been verified in a variety of settings (e.g., Guth and Tietz, 1990, and Slonim and Roth, 1998, among others).

Various explanations have been suggested. Perhaps the most commonly accepted is the idea of learning in the medium term (Roth and Erev, 1995, Erev and Roth, 1998). However, another commonly cited explanation is the idea of social conventions (e.g., Kandori, 1992, Camerer and Thaler, 1995, among many others). In this view, people develop notions of fairness and proper behavior in their daily life which they then apply in experimental situations, thus leading to behavior that appears irrational to the experimenters, but is rational in the context of typical societal interactions.

In this paper we take this latter approach, and show that the results in these experiments are consistent with the social norms of a larger "societal game' (see also, Kandori, 1992 and Ellison, 1994). This provides a fully rational explanation for behavior that is commonly considered to arise from boundedly rational play. In particular, we show that in anonymous repeated matching games in which players tremble with 'reasonable' probability the 'robust' sequential equilibria of these larger games leads to behavior in the stage games which is consistent with experiments. This provides an alternative, and we believe complementary[1], explanation of experimental behavior.

This paper is organized as follows: in the next section we review the experimental results and describe the games under study. Then in Section 3 we present our model of societal interactions, while Section 4 presents our main results. We conclude in Section 5 with a discussion.

2 Experiments on Extensive Form Games

Roth et. al. (1991) performed experiments in four different countries on two seemingly similar games, the market game and the ultimatum game. In all four countries play in the market game converged rapidly and robustly to the subgame perfect Nash equilibrium, while play in the ultimatum game converged rapidly to a nonsubgame perfect equilibrium in each country, but each was slightly different in the different countries.

In the market game there are groups of 11 players, 10 buyers and 1 seller, where the buyers bid on an object worth 1000 units (about US$10 or US$30), in 5 unit increments and then the seller decides whether to accept the highest bid, or reject all the bids. If the seller accepts a bid of b, then the player who made that bid receives $1000 - b$, the seller gets b and all the other players get 0. (Ties are broken randomly.) If the seller rejects the bid then all

[1]As we discuss in Section 6, we believe that there is a deep connection between robust social equilibria and learning. A well known intermediate connection has been offered by Young (1993) and others which study the evolution of social conventions in processes analogous to learning.

players get 0. Iin their experiments, the play very quickly converged to 995 or even 1000. They describe the robustness of this result as "striking."

In order to simplify the analysis we study a simpler version of this game. We assume that there are only 2 buyers who are restricted to 2 possible offers, low (L) or high (H). If a low offer is accepted the seller gets 2 utils and the buyer 3 utils, while if a high offer is accepted the payoffs are reversed. This simplification does not affect either the equilibria of the game or our results significantly. (Although, an important subtlety in the analysis is discussed in Section 4.3) In particular, the only subgame perfect strategy in this game is for all the buyers to offer H and for the seller to accept any offer.

The second game studied was the ultimatum game. In this game there are 2 players, the proposer and the responder. This game can be viewed as the market game with a single buyer, the proposer. In this game the only subgame perfect Nash equilibrium is for the proposer to offer 0 (or 5) and the responder to accept any offer. However, in the experiments (Roth et. al., 1990) play converged rapidly to the case where the proposer made a fair offer (typically between 400 and 600 units) which the responder accepted, while the responder would often (40-90% of the time) reject the low offer (typically less than 400). This equilibrium is Nash, but not subgame perfect and was considered the 'fair' outcome. Interestingly, the value of this medium offer varied slightly from country to country, about 500 units in the U.S. and Yugoslavia and 400 in Japan and Israel. Notice that all of these are still Nash equilibria and significantly different from the subgame perfect outcome. In fact, given the seller's behavior, Roth et. al. (1990) show that the buyer is in fact choosing the optimal offer, while rejections by the seller are clearly suboptimal in the one shot game, since rejecting any offer causes a loss in payoff.

In this case our simplified game has two offers (L, F) where F is the 'fair' bid. If the proposer offers L then she gets a payoff of 9 and the responder gets 1, if the responder accepts, while if she offers H and the responder accepts the payoffs are 5 for the proposer and 5 for the responder. Once again, both players get 0 if the responder rejects.

As we will demonstrate in the following, these results have precise analogues in the social equilibrium when embedded in a larger societal game; in particular, the nonoptimal behavior observed in the ultimatum game arises quite naturally as equilibria in social settings.

3 Robust Social Equilibria and Norms

We consider the following model of societal interactions with a simple model of information transfer. For ease of presentation we will discuss the set up for the market game, although the set up for the ultimatum game is analogous. Consider a population of M players, where for simplicity assume that M is divisible by 3 and divided into 2 groups, buyers and sellers. There are twice as many buyers as sellers. There are an infinite number of periods, $T = 0, 1, 2, \ldots$ in which players are divided up at random to play the market game, with one seller and 2 buyers in each game. At the end of each period, the players receive their payoffs for the game and then are rematched randomly in the following period. This model is a standard one so far (see, e.g. Rosenthal 1979); however we add 2 nonstandard parts.

First, we want a simple model of information transfer, so we assume that the history of play is public knowledge, but anonymous. Thus each player knows precisely what actions were chosen in every game in every previous period, but does not know which players chose those actions, since the information is anonymous. We will discuss this assumption in more detail in the context of our results in the next section.[2]

Second, we want our model to be robust in the sense that if not all players play perfectly, the analysis is still holds. To this end, we assume that players occasionally deviate from the equilibrium strategy. In particular, we assume that each time a player chooses an (intended) action, the player trembles with probability $\epsilon > 0$ and chooses a different (unintended) action, randomly from the remaining actions.[3] The main complications in our analysis will arise because, in contrast to the standard analyses with trembles we do not consider the limit as ϵ approached zero[4], but will instead be interested in "reasonable" values of ϵ. Note also, that one can also interpret ϵ as a fraction of "stupid" or "malicious" players.

Lastly, to simplify the presentation we will assume that there is a public signal on which players can coordinate their actions, q^t. Where $\{q^t\}$ is a sequence of i.i.d. variables which are each uniformly distributed on the interval $[0, 1]$.

Let h_i^t be player i's history at time t, which includes both the public history of actions, public signals and the player's personal history which includes her own play and that of the opponents against who she has been matched. A strategy in the societal game is a mapping s_i from histories to mixtures over actions. A player's payoff in this game is $U_i(s_i, s_{-i}) =$

[2]This assumption is also discussed in more detail in Friedman (1998) and Friedman and Resnick (1998). The latter paper discusses examples from the Internet, in which such information is often kept by the computer system and is publicly available in the form of history file or publicly available archives.

[3]For example, evolutionary analyses of the prisoner's dilemma are dramatically altered by the presence of trembles as in Nowak and Sigmund (1993).

[4]For example, trembling hand perfect equilibria (Selton , 1975) or in a closely related model to the one in this paper (Ellison, 1994)

$E[\sum_{t=0}^{\infty} \delta^t \pi_i^t | s, \epsilon]$ where the expectation is over both mixtures and trembles.

Definition 1 *A set of strategies in the societal game* $< \pi, A, \delta, \epsilon, M >$ *which are part of a sequential equilibrium will be denoted a* social equilibrium*.*

As experimental observers, we are interested in the play in the stage game as part of a social equilibrium. When play in a social equilibrium leads to a dominant mode of play in the stage games we say that this is a "social norm."

Definition 2 *For any* $\gamma \in [0, 1]$ *a set of strategies for the stage game is a* γ*-social norm if those strategies are played with probability greater than* γ *in the stage games, where this probability is taken with respect to trembles, player randomizations, and the public signal.*

Note that since $\epsilon > 0$ every play path is possible, and thus our analysis is robust with respect to the standard refinements.

4 Results

We now study the relationship between the parameters of the societal game $< \pi, A, \delta, \epsilon, M >$ and the set of social equilibria and social norms. Our main argument will be that for "reasonable choices of parameter values" we get social norms which agree with those found in the experiments previously discussed; in particular we show that for reasonable parameter values the only social norms of the societal market game is one in which players play the subgame perfect equilibrium of the (one-shot) market game, while for similar parameter values the observed "fair equilibrium" of the ultimatum game (which is not subgame perfect) is a social norm.

4.1 The market game

The basic idea of our analysis, is that the set of social equilibria and norms depends crucially on the expected total number of trembles per period, $M\epsilon$. This is because when players are anonymous, societal punishments must be meted out to the entire population, which is quite costly, and in equilibrium will be triggered accidentally by trembles.[5]

For example, consider the following collusive set of strategies, which is the only other Nash equilibrium of the stage game. In the normal phase, the buyers all offer L and the seller accepts all offers. If in any period some buyer deviates and offers the high amount H, a myopically beneficial deviation, then the players switch to a punishment phase, in which all buyers make high offers and the seller still accepts all offers. The punishment phase continues until the first period when the exogenous signal q^t is less than some threshold, denoted by $\hat{q}$.

Now we study the parameter values for which these strategies form a social equilibrium. Since the seller never has any incentive to deviate we focus on the buyers. Let V (resp. W) be the expected future payoffs beginning in a normal (resp. punishment) phase. When a player decides to deviate, she must weigh the immediate benefit, which is approximately $2 - 3/2 = 1/2$ (the expected payoff from bidding H minus the payoff from bidding L and winning half of the time) against the expected loss which is roughly $\delta(V - W)P$ where P is the probability that no other buyer, in the entire society, has deviated. Note that

[5] This is closely related to the paradox of trigger strategies in repeated games with imperfect monitoring (Green and Porter, 1984) which is forcefully brought out in Rubinstein's (1979) model of criminal sanctions in a world of rational citizens with imperfect monitoring. In equilibrium it is necessary to punish people who appear to have violated the law, but are known (by the assumptions of rationality and equilibrium) to be innocent.

$P = (1-\epsilon)^{M'-1}$, where $M' = 2M/3$ is the number of buyers in the entire society. Note that when $M\epsilon$ is large P is approximately $e^{-M'\epsilon}$ and thus this probability depends roughly on the value of $M'\epsilon$.

First we compute all these quantities then we carefully choose $\hat{q}$ so that buyers will be indifferent between deviating and note that this leads to the most collusive outcome in which

$$\hat{q}(M,\epsilon) = \frac{(1+\epsilon)(1-\epsilon)^{-2M/3}}{\delta(2+\epsilon)}.$$

(The detailed computations are in the appendix.) Since $\hat{q}$ must be less than 1, there can be no equilibrium when $\hat{q}(M,\epsilon) \geq 1$. Fixing ϵ we solve for $M_{\max}$ which satisfies $\hat{q}(M_{\max},\epsilon) = 1$, which yields

$$M_{\max} = \left(\frac{3}{2}\right) \frac{\ln(\frac{1+\epsilon}{2+\epsilon}) - \ln(\delta)}{\ln(1-\epsilon)}.$$

This is the largest value for M such that the punishment strategies yield an equilibrium.

For example, when $\epsilon = .05$ and $\delta = .95$ the equilibrium does not exist for $M > M_{\max}(.05) = 18.06$ a fairly small population size. Even when $M < M_{\max}$ play may still not be far from the subgame perfect outcome and colluusive play not a social norm. For example, when $M = 12$ then $\hat{q}(10, .05) = .81$ which implies that punishment stages last about 5.3 periods on average. Since the probability of beginning a punishment period is approximately 0.34 we see that the expected time between punishment periods is approximately 2 periods. This implies that play is in a normal phase only about 27% of the time and thus play is still predominantly at the subgame perfect outcome. Thus, for any $\gamma > 0.27$ the collusive equilibrium is not a γ-social norm. Note also that these calculations are only weakly dependent on the discount factor, δ. For example, when $\delta = .99$, then $M_{\max}(.05) = 19.3$

which is very close to the value when $\delta = .99$.

In order to understand these relationships, we take the Laurent expansion of $M_{\max}$ near $\epsilon = 0$. This yields $M_{\max} \approx (3/2)\ln(2\delta)/\epsilon$, and thus we see that $M\epsilon$ is the crucial parameter which determines whether there is play that is not part of the subgame perfect outcome of the market game.

Given these numerical results, we believe that it is quite reasonable to claim that for reasonable tremble probabilities and population sizes the only reasonable outcome is the subgame perfect equilibrium of the market game.[6] Formally, we have shown the following:

Theorem 1 *Consider the societal market game* $< \pi, A, \delta, \epsilon, M >$*. Then:*

1) for $\epsilon < .1$ *and* $M\epsilon < \ln(2\delta)$ *there exists a social equilibrium in which players do not play the market equilibrium in every period.*

2) for any $\gamma \in [0, 1)$ *there exists a* $\beta > 0$ *such that for all* $\epsilon < .1$ *and*$M < \beta/\epsilon$ *the collusive stage game strategies are a* γ*-social norm.*

The obvious weakness of the preceding argument is that we have only considered a specific set of strategies, and that some other set of strategies might yield a non-subgame perfect outcome, i.e, we have not proven the converse of the above theorem. However, the following theorem shows that for $M\epsilon$ sufficiently large the only social norm is the market equilibrium.

Theorem 2 *Consider the societal market game* $< \pi, A, \delta, \epsilon, M >$*. Then:*

for any $\gamma > 0.5$*, there exists* $\beta > 0$ *such that in any game with* $\epsilon < .1$ *and* $M > \beta/\epsilon$ *the only*

[6]Friedman (1998) provides a similar analysis for the prisoner's dilemma; however in that game not only is the static Nash equilibrium (defect, defect) the only social equilibrium, but the strategy defect always is a dominant strategy in *for the repeated game,* and thus the only rationalizable strategy.

γ-social norm is the market equilibrium.

4.2 The Ultimatum Game

Based on the analysis in the previous section, it might appear that the only social norms for the ultimatum game consists of playing the subgame perfect outcome of the stage game. This is true in the following sense:

Theorem 3 *Consider the societal ultimatum game, $< \pi, A, \delta, \epsilon, M >$. Then: in any game with $\epsilon < .1$ there exists some $\overline{M} > 0$ such that for $M > \overline{M}$ the unique social equilibrium is one in which the subgame perfect equilibrium of the stage game is (intended to be) played in every period.*

This immediately implies the analogous result for social norms:

Corollary 1 *Consider the societal ultimatum game, $< \pi, A, \delta, \epsilon, M >$. Then: in any game with $\epsilon < .1$ there exists some $\overline{M} > 0$ such that for $M > \overline{M}$ the unique γ-social norm is the subgame perfect equilibrium of the stage game.*

Thus, for fixed $\epsilon > 0$ there exists some population size for which we cannot get the fair equilibrium as a social norm. However, this result is misleading as the following example suggests.

Once again consider a trigger equilibrium, but this time instead of triggering when the proposer for makes an 'unfair' proposal, we only trigger when the responder accepts such a proposal. Note that we are triggering on an action which is off the equilibrium path. This will make a significant difference in the dependence of this strategy on the population size.

Formally, define the trigger strategy as follows: In a normal phase the proposer offers F and the responder accepts F but declines L. If the proposer trembles and offers L *and* the responder trembles and accepts that offer a punishment phase begins in which the proposer offers L and the responder accepts all offers. The punishment phase ends in the first period in which $q^t < \hat{q}$. As we will see the key aspect of this strategy is that it takes 2 mistakes to trigger a punishment phase.

Note that this strategy is consistent with the observed play in Roth et. al. (1990), in which most proposers make the fair offer, but occasionally proposers make lower offers, which are often rejected by the responder. In fact, as mentioned earlier, they show that the proposers are making the optimal offers conditional on the responder's (off equilibrium in our modeling) behavior.

Thus, in a normal period if neither player actually deviates, the responder gets 5; if only the proposer deviates then they both get 0; if only the responder deviates then both get 0, while if they both deviate the responder gets 1. As in the previous example, we can compute the conditions for these strategies to be an equilibrium. Note that the key difference is that the probability of accidentally triggering a punishment phase is now $(1-\epsilon^2)^{M/2}$ which is approximately $e^{-M\epsilon^2/2}$ and now the crucial quantity is $M\epsilon^2$ instead of $M\epsilon$. Detailed calculations yield (in the appendix):

$$\hat{q} = \frac{(1+\epsilon)(1-\epsilon^2)^{-M/2}}{(5-3\epsilon)\delta}$$

and

$$M_{\max} = 2\frac{\ln(\frac{1+\epsilon}{\delta(5-3\epsilon)})}{\ln(1-\epsilon^2)}.$$

As before we compute the value of $M_{\max}$ for various parameters. For example, for the societal ultimatum game $M_{\max} = 1181$ when $\delta = .95$ and $\epsilon = .05$. This was only 14 for the market game, with the same parameters. Thus we see that for reasonable error probabilities and fairly large populations the nonsubgame perfect equilibrium of the ultimatum game can arise as a social equilibria. Also, notice that the dependence on δ is weak.

Now consider some 'reasonable parameters'. Let $M = 300$, $\delta = .95$ and $\epsilon = .05$. In this case $\hat{q} = .26$ and the fair outcome is an 80%-social norm. While for $\delta = .95$ and $\epsilon = .01$ for any $M < 3000$ the fair outcome is an 80%-social norm.

Once again consider the first term of the Laurent expansion of $M_{\max}$. This is $M_{\max} \approx 2\ln(5\delta)/\epsilon^2$ which implies that this equilibria can be sustained for $(M\epsilon)\epsilon \leq 2\ln(5\delta)$, while the analogous condition was $M\epsilon \leq (3/2)\ln(2\delta)$ for the market game. This is because the strategies are triggering on accidental *off equilibrium* deviations which are far less likely than accidental *on equilibrium* deviations.

To summarize, we have the following analogues to Theorems 1 and 2:

Theorem 4 *Consider the societal ultimatum game* $< \pi, A, \delta, \epsilon, M >$*. Then:*

1) for $\epsilon < .1$ *and* $M < \ln(2\delta)/\epsilon$ *there exists a social equilibrium in which players do not play the subgame perfect equilibrium in every period.*

2) for any $\epsilon < .1$ *and* $\gamma > 0$ *there exists a* $\beta > 0$ *such that for* $M < \beta/\epsilon^2$ *the "fair equilibrium" strategies are a* γ*-social norm.*

Theorem 5 *Consider the societal ultimatum game* $< \pi, A, \delta, \epsilon, M >$. *Then:*

for any $\gamma \in [0,1]$, *there exists* $\beta > 0$ *such that in any game with* $\epsilon < .1$ *and* $M > \beta/\epsilon^2$ *the only* γ*-social norm is the subgame perfect equilibrium of the stage game.*

If we think of the proposer as the salesperson naming a price and the receiver deciding whether or not to purchase, we see that this equilibrium corresponds to a social convention in which salesman who overcharge are not socially sanctioned, but consumers who purchase overpriced goods are, in the sense that salesmen learn that they can "get away" with making low offers.

Lastly, note that this argument is unchanged for the ultimatum game when the offer can be anything between 0 and 20 units. In particular, any offer greater than 0 can be sustained as a social equilibria. Thus, it is not surprising that the modal offer in different countries may be different, as this is the case of selection of a social convention for which there is no distinction between the different conventions.[7] However, in every case the proposer's offer must be optimal with respect to the responders expected action, which need not be optimal, as seen in Roth et. al. (1991).

4.3 A comment on the market game

At first it seems possible that a construction such as this could be used to construct a nonsubgame perfect social norm in the market game. For example one possible approach is to have the buyers offer L and have the seller only accept low offers. This strategy is triggering on off equilibrium behavior and thus could be robust. However, the difficulty is

[7]For more on the selection of social conventions see, e.g., Young (1993, 1995), and references therein.

that the buyer is receiving the lowest possible payoff, and thus there is no way to punish him for a deviation when this is a γ-social norm with $\gamma > .5$.

One might think that this is related to the fact that we are considering a reduced action set where there are only 2 possible offers. If there were 3, such as $L < F < H$ then the buyers could collude on F and switch to L if the seller accepts a high offer. While this enforces the sellers behavior in the normal phase, there is no way to maintain the punishment phase in a γ-social norm for $\gamma > .5$, as there is no further punishment to enforce the seller's behavior.

Notice that this result holds for any number of possible actions due to an unraveling argument. It even holds when offers are continuous, since each punishment phase must be a finite distance below the previous offer and even though the size of these steps can become arbitrarily small as $M\epsilon$ gets large, the unraveling argument still leads to the conclusion that no non-market outcome is possible as a social norm.

5 Extensions

One can formalize the concept of distance from the equilibrium path in terms of the number of deviations needed to reach a subtree of the stage game, and show that stage game equilibria in which the only profitable deviations are k deviations away from the equilibrium path depend on the quantity, $M\epsilon^k$, and thus typically can be supported as social norms for large population sizes. For example, any nonsubgame perfect equilibrium in the market game, has a profitable deviation on the equilibrium path and thus $k = 0$, while the fair equilibrium in the ultimatum game has $k = 1$ and is more 'robust.'

For another example, consider the alternating bargaining game (Rubinstein, 1982). If there are r alternating offers then there are nonsubgame perfect strategies for which the only profitable deviations are distance $k = 2r - 1$ from the equilibrium path. For example, when $r = 2$ and the payoffs are chosen to be identical to the ultimatum game we studied above, nonsubgame perfect equilibria can exist as γ-social norms for quite large ϵ and M, e.g. when $\epsilon = .01$, $M_{\max} \approx 400,000$ while even for $\epsilon = .1$, $M_{\max} \approx 4,000$.

In fact, in the infinite stage alternating bargaining game, for any $\epsilon < .1$, $M > 0$ and $\delta > .9$ there exist nonsubgame equilibria of the stage game which are γ-social norms for *any* size population and value of $\gamma \in [0, 1)$; however, these norms only differ slightly (along the equilibrium path) from the unique subgame perfect equilibrium.

6 Summary and Concluding Comments

To summarize we present the following table which compares the parameters under which non subgame perfect outcomes can arise, for $\delta = .95$ and $\gamma = 0.8$.

		Market Game				Ultimatum Game			
			Non SPE exist?				Non SPE exist?		
ϵ	M	$\hat{q}$	Equil.	Norm	% non SPE	$\hat{q}$	Equil.	Norm	% non SPE
.05	30	-	No	No	-	.23	Yes	Yes	98
.05	300	-	No	No	-	.26	Yes	Yes	82
.05	3000	-	No	No	-	-	No	No	-
.01	30	.57	Yes	Yes	84	.21	Yes	Yes	99.9
.01	300	-	No	No	-	.21	Yes	Yes	99
.01	3000	-	No	No	-	.3	Yes	Yes	93
.01	30000	-	No	No	-	.38	Yes	No	43

Table 1: Existence and properties of nonsubgame perfect outcomes from trigger strategy equilibria.

Thus, it seems reasonable to claim that the market outcome should be observed in the market game, while the fair outcome is reasonable in the ultimatum game. Thus, the distinction between enforcement on and off the equilibrium path is an important factor in robust social equilibria, and perhaps other types of robust solution concepts.[8].

On important general insight from this analysis, is that limiting behavior of models may be misleading. In Kandori's (1990) analysis, $\epsilon = 0$, $\delta \in (0,1)$ was fixed and the limit of $M \to \infty$ was taken, while in Ellison (1994), $M > 0$ and $\delta \in (0,1)$ were chosen carefully, while the limit $\epsilon \to 0$ was taken. In the early stages of this work we considered the case where $\epsilon > 0$ and $\delta \in (0,1)$ were fixed and the limit $M \to \infty$ was taken. All three would lead to misleading conclusions on at least one of these two games.

Only by carefully analyzing the actual parameter values, before taking limits, do we get the answers presented in this paper, which we believe are the reasonable outcomes of the model and also agree with the behavior observed in experiments, and perhaps more generally in society.

A Appendix

A.1 Theorems 1 and 4

In this section we provide a sketch of the analysis used to compute the equilibrium conditions in the examples.[9] The main complication is the explicit consideration of trembles.

[8]For example, this distinction has been noticed in models of learning and appears to be closely related to the idea of self-confirming equilibria (Fudenberg and Kreps, 1995. Fudenberg and Levine, 1993).

[9]The detailed computations were done using Maple© and are available directly from the author upon request.

First note that in a normal period, the immediate expected payoff is

$$\pi_{norm} = (1-\epsilon)\left((1-\epsilon)\left(3/2 - 3/2\,\epsilon\right) + \epsilon\ (2-\epsilon)\right),$$

where we have conditioned on all possible combination of trembles. Similarly in a punishment period we have

$$\pi_{punishment} = (1-\epsilon)\left((1-\epsilon)\,(1+\epsilon) + 3/2\,\epsilon^2\right).$$

In a normal period the probability of there being a deviation which starts a punishment phase is $P_{dev} = 1-(1-\epsilon)^{2M/3}$ and the probability of a punishment phase ending is $1-q$. If we let v be the expected normalized discounted payoffs (payoffs multiplied by $1-\delta$) starting in a normal phase and w the analogous quantity starting in a punishment phase, then these quantities must satisfy the following equations:

$$v = (1-\delta)\pi_{norm} + (1-P_{dev})\delta v + P_{dev}\delta w$$

$$w = (1-q)v + q((1-\delta)\pi_{punishment} + \delta w),$$

This gives us equations for v, w in terms of M, δ, ϵ, q.

Now we choose q to be as small as possible while maintaining the strategies as an equilibrium. The gain from choosing to deviate, for a buyer, is $\pi_{norm} - \pi_{punishment}$ while the loss from future payoffs is $v - w$ multiplied by the probability of their deviation triggering a punishment phase, which wouldn't have been triggered by another player, which is $P_{trig} = (1-\epsilon)^{(2M/3)}$. Thus, the equilibrium condition is

$$\pi_{norm} - \pi_{punishment} = (1-\epsilon)^{(2M/3)}(v-w)$$

from which we solve for $\hat{q}$.

Then, the condition $\hat{q} = 1$ is solved to find $M_{\max}$. Lastly we compute the expected time in a normal phase as $\tau_n = 1/(1 - \hat{q})$, and the expected time in a normal phase which is $\tau_p = 1/P_{dev} - 1$ and compute the percentage time in the normal phase as $\tau_n/(\tau_n + \tau_d)$, completing the analysis.

The identical analysis is used for the bargaining game, but the values of $\pi_{norm}, \pi_{punishment}$ and the various probabilities differ.

A.2 Proofs of Theorem 2, 3 and 5

First we consider Theorem 2:

Let $\beta = c * \delta/(1-\delta)$, where c will be defined below and assume that $\epsilon < .1$ and $M > \beta/\epsilon$. Let s be a social equilibria of this societal market game. Given any history h^t let $\psi(h^t)$ be the expected fraction of players choosing action L conditional on h^t. First we prove the following:

Lemma 1 *For every history h^t, $\psi(h^t) > .9$ or $\psi(h^t) < .1$.*

Proof: Assume that $\psi(h^t) \in (.1, .9)$. Then it is never optimal for a seller to reject all offers. This is because the immediate loss from rejecting an offer is at least 1 while the gain is at most the expectation of future payoffs conditional on declining the offer minus the expectation of future payoffs conditional on declining the offer. However, as shown in Friedman and Resnick (1998) Lemma 1, for c sufficiently large this quantity can be made to be less than 1, since the probability distribution of the public history (and most private histories) are very close to each other under defection and non-defection (see also, Sabourian, 1990). Since sellers are accepting all offers, it is always optimal for a buyer to bid H, for the same reason, the immediate gain outweighs the expected loss, for sufficiently large c. However, this violates the assumption that $\psi(h^t) \in (.1, .9)$, proving the lemma by contradiction. $\diamond$

Lemma 2 *For every history h^t:*

1) if $\psi(h^t) > .9$ then $\psi(h^t) = 1$ and sellers accept the high offer.

2) if $\psi(h^t) < .1$ then $\psi(h^t) = 0$ andplay at stage t has both buyers (intending to) bid L and the seller only accepting the low offer.

Proof: If $\psi(h^t) > .9$ then the probabilistic argument from the previous lemma implies that sellers must accept offers of the type L. Combining this with the probabilistic argument requires that sellers must also reject offers of the type H; otherwise, the buyers would deviate and offer H, and then this implies that all buyers offer L. Similarly if $\psi(h^t) > .9$ then the probabilistic argument implies that sellers must accept H offers , and then all buyers must offer H. $\diamond$

Proof of Theorem 2: From the above results, it must hold that in any equilibrium, play in every period consists either of all buyers offer H or all offer L. Now we note that in a L-period (in which all buyers offer L) it *is possible* for the strategies to condition on whether a seller accepts a high offer, as these need not be rare under the assumptions, i.e., although M/ϵ is 'large', M/ϵ^2 could be small. However, the future payoffs if a seller so deviates must be sufficiently less than those when no seller deviates, in order to prevent a seller from deviating on a buyer's tremble and increasing her immediate profit. However, since we are assuming that $\gamma > .5$ this implies that more than half of the periods must be L-types. However, the only way to prevent sellers from deviating in a L-period is to reward them with an H-period, but this implies that there must be more than one H-period following every L-period (since the gain from deviating is 2, while the difference in payoffs from an L-period and an H-period is 1 for the seller) contradicting the assumption that L type periods are a .5 social norm. □

The proofs of Theorems 3 and 5 are analogous.

References

[1] C. Camerer and R. Thaler. Ultimatums, dictators and manners. *Journal of Economic Perspectives*, 9(2):209–219, 1995.

[2] G. Ellison. Cooperation in the prisoner's dilemma with anonymous random matching. *Review of Economic Studies*, 61:567–588, 1994.

[3] I. Erev and A. Roth. Predicting how people play games: Reinforcement learning in experimental games with unique, mixed strategy equilibria. *American Economic Review*, 88(4):848–81, 1998.

[4] E. J. Friedman. On social norms in noisy environments. mimeo, 1999.

[5] E. J. Friedman and P. Resnick. Cooperating on the Internet: the social cost of cheap pseudonyms. mimeo, 1998.

[6] D. Fudenberg and D. Kreps. Learning in extensive-form games. I. self-confirming equilibria. *Games and Economic Behavior*, 8(1):20–55, 1995.

[7] D. Fudenberg and D. Levine. Self-confirming equilibrium. *Econometrica*, 61(3):523–45, 1993.

[8] E. Green and R. Porter. Noncooperative collusion under imperfect price information. *Econometrica*, 52:87–100, 1984.

[9] W. Guth and R. Tietz. Ultimatum bargaining behavior: A survey and comparison of experimental results. *Journal of Economic Psychology*, 11:417–449, 1990.

[10] M. Kandori. Social norms and community enforcement. *Review of Economic Studies*, 59:63–80, 1992.

[11] M. Nowak and K. Sigmund. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the prisoner's dilemma game. *Nature*, 364(64):56, 1993.

[12] R. Rosenthal. Sequences of games with varying opponents. *Econometrica*, 47(6):1353–1366, 1979.

[13] A. Roth and I. Erev. Learning in extensive form games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, 8:164–212, 1995.

[14] A. Roth, V. Prasnikar, M. Okuno-Fujiwara, and S. Zamir. Bargaining and market behavior in Jerusalem, Ljubljana, Pittsburgh and Tokyo: An experimental study. *American Economic Review*, 81:1065–1098, 1991.

[15] A. Rubinstein. An optimal conviction policy for offenses that may have been commited by accident. In A. Schotter S. Brahms and G. Schwodiauer, editors, *Applied Game Theory*. Physica-Verlag, 1979.

[16] A. Rubinstein. Perfect equilibrium in a bargaining model. *Econometrica*, 50(1):97–109, 1982.

[17] H. Sabourian. Anonymous repeated games with a large number of players and random outcomes. *Journal of Economic Theory*, 51(1):92–110, 1990.

[18] R. Selton. A reexamination of the perfectness concept for equilibrium points in extensive form games. *International Journal of Game Theory*, 4:25–55, 1975.

[19] R. Slonim and A. Roth. Learning in high stakes ultimatum games: An experiment in the Slovak republic. *Econometrica*, 66(3):569–96, 1998.

[20] H. P. Young. The evolution of conventions. *Econometrica*, 61(1):57–84, 1993.

[21] H. P. Young. The economics of conventions. *Journal of Economic Perspectives*, 10(2):105–22, 1995.