EIEF WORKING PAPER SERIES

EINAUDI INSTITUTE FOR ECONOMICS AND FINANCE

EIEF Working Paper 13/09

December 2009

Equilibrium Trust

by

Luca Anderlini

(Georgetown University)

Daniele Terlizzese

(EIEF and Bank of Italy)

Equilibrium Trust*

Luca Anderlini
(Georgetown University)

Daniele Terlizzese
(EIEF and Bank of Italy)

December 2009

Abstract. We build a simple model of trust as an *equilibrium phenomenon*, departing from standard "selfish" preferences in a minimal way. Agents who are on the receiving end of an offer to transact can choose whether to cheat and take away the entire surplus, taking into account a "cost of cheating." The latter has an idiosyncratic component (an agent's type), and a socially determined one. The smaller the mass of agents who cheat, the larger the cost of cheating suffered by those who cheat.

Depending on the parameter values, the model can have a unique equilibrium level of trust (the proportion of transactions not cheated on), or two equilibria, one with high and the other with low trust. Thus, differences in trust levels across societies can reflect different fundamentals or, for the same fundamentals, a switch across multiple equilibria. Surprisingly, we find that these two possibilities are partially identifiable from an empirical point of view.

Our model can also be reinterpreted as one with standard selfish preferences and an enforcement agency with limited resources that are used to catch and fine a subset of those who cheat. Lastly, we carry out a robustness exercise in which agents learn in a simple way from experience about how many agents cheat in society. Our results indicate that when there are multiple equilibria the high trust equilibrium is less robust than the low trust one.

JEL Classification: D80, D89, C79.
Keywords: Trust, Social Norms, Multiple Equilibria, Enforcement.
Address for correspondence: Luca Anderlini, Department of Economics, Georgetown University, 37^{th} and O Streets, Washington DC 20007, USA. luca@anderlini.net

*We are grateful to Helena Aten, Larry Blume, Hülia Eraslan, Leonardo Felli, Dino Gerardi, Guido Kuersteiner, Roger Lagunoff, George Mailath, Aleh Tsyvinski and seminar participants at Georgetown and Johns Hopkins for comments. Luca Anderlini thanks EIEF for their generous hospitality. All errors remain our own.

1. Introduction

1.1. Motivation and Overview

Not many transactions are carried out using rotating security trays with money on one side and the purchased good on the other. Yet, without these devices or some equivalent arrangement the standard selfish agents that populate modern economic models would be unable to trade, barring repeated interaction or binding contracts. It is instead self-evident that trade among agents flourishes way beyond what this hypothetical world would look like.

The lubricant that makes so many transactions take place is *trust*. This is the belief that economic agents hold that the other side of the transaction will not behave in a completely opportunistic way, thus impeding mutually advantageous exchange.[1]

Our purpose here is to build a simple model of trust as an *equilibrium* phenomenon. The agents' beliefs about the probability of not being cheated — their level of trust as we just defined it — should be endogenously determined in equilibrium, and hence *correct*. The reason we focus on beliefs is simple. There is a basic tension between "trusting beliefs," and consequent "trusting behavior," and the incentives to cheat of other agents in society. Trusting beliefs can be exploited. However, a trustful agent should not be cheated often; if she is, she should change her belief and start trusting less. This tension seems to warrant a close examination of what an equilibrium model of trust can generate.

To model trust we place a "cost of cheating" in the agents' utility function (this is our sole departure from the standard purely-selfish agents paradigm). The cost of cheating has two components. One which is an exogenous characteristic of each agent, and another which is socially determined by the behavior of others. The less common cheating behavior is in society, the higher is the cost of cheating for individual agents. This feedback component of the cost of cheating is a central ingredient of our analysis, and we return to it at length below.

While we insist that trust should emerge as an equilibrium phenomenon in our model, we do not intend to dismiss "behavioral traits" and non-equilibrium factors in the explanation of trust as a belief. Rather, by investigating whether a successful model of trust in a society can be built along standard lines we hope to shed light on whether, and if so which, departures from the standard paradigm in the agents' belief formation are needed to address the issue

[1]A large literature exists on the sources and effects of the presence of trust. For ease of exposition, we postpone any discussion of it until Subsection 1.2 below.

of trust. This may eventually lead to an understanding of how important the new behavioral traits are in interpreting reality.

Our model is deliberately kept simple in the extreme. In particular, our model is *not* dynamic. Although repeated interactions have been used very successfully to generate "co-operative behavior" of many kinds, our initial motivation as above is to have a model of trust that applies to situations that would seem to call for "swivel-tray trading" if trust were not present. Hence, we stay deliberately clear of reputational issues and more generally of repeated interaction ingredients in our set up. As well as making our analysis more transparent, this way of proceeding makes our results immune from the price that many dynamic models have to pay. Multiplicity of equilibria (although it features prominently in our analysis) is not an issue for us. While dynamic models are often plagued by a staggering multiplicity of equilibria that dramatically curtails their predictive ability, we are able to proceed with only a very mild simplifying assumption in this respect.

In spite of its extreme simplicity, our set up provides a rich enough framework to address, in an interesting and novel way, the well documented diversity of levels of trust in different societies.[2] Indeed, depending on the configuration of preference and other parameters, our model either generates a unique equilibrium (with a single equilibrium level of trust), or two equilibria, one with a higher and another with a lower level of trust. Moreover, by varying the parameters of the model, higher or lower levels of equilibrium trust can be obtained without switching across different equilibria. Given this rich set of possible equilibrium outcomes, the model allows us to frame in a natural way an important question on the observed diversity of levels of trust in different societies. Are two societies, one with a high level of trust and another with a lower one, *different because their fundamental parameters differ*, or because, given *same fundamental parameters*, they *happen to be in different equilibria*? Obviously these two possibilities have different policy implications, as it is likely easier to shift from one equilibrium to another for given fundamentals than it is to induce a change in the fundamental parameters.

A key insight from our model is that the two cases are in fact *partially identifiable* in terms of the outcomes they generate. A stark prediction of our analysis is that if different levels of trust result from multiple equilibria, then the level of trust must be negatively correlated with the size of individual transactions. A positive correlation can only emerge if different levels

[2]As opposed to heterogeneity of behavior *within* a society, which is not our direct focus of attention, although not in contrast with our results.

of trust were to result from differences in the parameters of the model.[3] To our knowledge, the possibility to empirically disentangle multiple vs. unique equilibrium regimes is new.

A second set of results concerns the relationship between levels of trust and measures of economic performance. While the level of trust correlates positively with a measure of overall activity in the economy, the link between level of trust and welfare (inclusive of the perceived costs associated with cheating) is ambiguous. Though our model is much too simple to take welfare measures as other than qualitative, we find this difference interesting and worth exploring in more detailed models.

A third batch of results concerns the fact that the social feedback on the cost of cheating outlined above can be re-interpreted as resulting from an enforcement technology whose effectiveness depends on the average behavior. We find that, in the multiple equilibria regime, an infinitesimal increase in the resources devoted to enforcement can yield a discontinuous increase in the level of total activity in the economy (the effect on welfare is ambiguous, due to our third result). In the single equilibrium regime, instead, the level of activity changes continuously with the resources spent in the enforcement technology.

A fourth batch of results that emerges from our analysis is that, in the multiple equilibria regime, the low-trust equilibrium is more robust than the high-trust one in a well specified sense. Roughly, small deviations from the high-trust equilibrium are much more likely to destroy it. High-trust is thus "more difficult to sustain."

1.2. Related Literature

We are certainly not the first to point out that without security swivel-trays an element of trust is needed for most transactions to take place. Arrow (1972, p. 357) notes that "Virtually every commercial transaction has within itself an element of trust, certainly any transaction conducted over a period of time." In the absence of "instantaneous exchange," an element of trust is required.

Arrow (1972) goes on to comment on the path-breaking study by Banfield (1958) of the devastating effects of the lack of trust on a "backwards" small community in southern Italy.

[3]Since the term "partial identification" has been used before (Phillips, 1989), it is useful to be precise as to the meaning we give it here. The identification is *partial* in the sense that we cannot rule out that two equilibria corresponding to two sets of parameters entail a negative correlation between trust levels and size of individual transactions. Therefore, while a positive correlation excludes the possibility of a switch across multiple equilibria with unchanged parameters, the observation of a negative correlation is inconclusive.

Following Banfield (1958) a large literature has blossomed on the roots and effects of the lack, or presence, of trust.

The literature is way too large and varied to attempt even a reasoned outline here, let alone a survey. We confine ourselves to recalling how Putnam (1993) documents the heterogeneous levels of "social capital" in different regions of Italy and its role in fostering growth. A couple of years later, Fukuyama (1995) published an influential monograph concerning the positive role of trust in large firms and hence economic growth.[4]

We also selectively recall the contributions by Knack and Keefer (1997), La Porta, Lopez-de-Silanes, Shleifer, and Vishny (1997), and more recently Guiso, Sapienza, and Zingales (2004), Sapienza, Toldra, and Zingales (2007) and Butler, Giuliano, and Guiso (2009).[5] These studies all document in a variety of ways how the presence of trust is correlated with desirable economic outcomes.

The departure from standard preferences that we postulate here can also be traced back a long way. Vernon Smith (1998), citing experimental evidence draws out the distinction that goes back to Adam Smith (1759, 1776) between "moral" and "selfish" preferences. It is interesting to go back this far since in this reading of Smith (1759), "moral sentiments" seem to fit well the idea of an "extra entry" in agents' utilities that capture their regard for the "fortune of others." In essence, we take the same approach here. Arrow (1972) contains an illuminating discussion of the different ways in which non-selfish motives can enter agents' preferences.

Much more recently Feddersen and Sandroni (2006) explore the effects of "ethical" social feed-back mechanisms not unlike the one we consider here, and their effects on the equilibria of voting models.[6] Horst and Scheinkman (2006) are concerned with the general theoretical problems of models with social feed-back variables, particularly with the (far from trivial) issues that arise in proving the existence of equilibrium in general in this class of models.

Dixit (2003) and Tabellini (2008) are both theoretical contributions to the literature on

[4]It is important to note at this point that the literature documenting the positive impact of "social capital" or "trust" on income, wealth and growth rates also includes some notable skeptics. To our knowledge, the most prominent one is Solow (1995), who in turn cites evidence from Kim and Lau (1994) and Young (1994). For a survey of much of the literature on trust and social capital we cite here, including an account of the debate we have just mentioned, see Sobel (2002).

[5]Glaeser, Laibson, Scheinkman, and Soutter (2000), Irlenbusch (2006) and Butler, Giuliano, and Guiso (2009) focus on experimental set-ups.

[6]Blume (2004) investigates the effect of "stigma" in a dynamic model.

trust. Their main focus is on the differential effects of distance and society's size on the sustainability of trust. In both cases the possibility of equilibrium phenomena that address the issue of trust is due to the dynamic, repeated nature of the interaction between agents. In our model, play only takes place once, and trusting and trustworthy equilibrium behavior can be traced directly back to our non-standard preferences generating the social feed-back we have described above.

Finally, the theoretical literature on repeated interactions, from which we purposedly stay away, is also vast. We simply refer the interested reader to the recent monograph by Mailath and Samuelson (2006) which also has a comprehensive and up-to-date list of references.

1.3. Plan of the Paper

The rest of the paper is structured as follows. In Section 2, we describe the basic model in detail and make precise what constitutes an equilibrium in our set up. In Section 3 we characterize the set of possible equilibria of the model. In Section 4 we highlight how high and low trust equilibria may arise from either differences in the fundamental parameters of the model, or a switch across multiple equilibria supported by the same set of parameter values. In this section, we also spell out the identifiable characteristics of high and low trust equilibria in these two cases, and we proceed to characterize transaction volumes and welfare properties of the different equilibria. Section 5 provides a re-interpretation of the socially generated component of cheating costs as stemming from an enforcement technology with limited resources available. In Section 6 we carry out a robustness analysis of low and high trust equilibria in the multiple equilibria regime. Finally, Section 7 briefly concludes.

For ease of exposition, all proofs have been relegated to an Appendix.

2. Set-Up

2.1. The Model

There is a continuum of risk-neutral players of mass 2 uniformly distributed on $[0, 2]$. All players $i \in [0, 2]$ face equal and independent chances of playing on the offer side, and on the receiver side. We refer to the former as $\mathcal{O}$ agents and the latter as $\mathcal{R}$ agents. So there is a unit mass of both $i \in \mathcal{O}$ and $i \in \mathcal{R}$ agents after the realization of this first draw.

The $\mathcal{O}$ and $\mathcal{R}$ agents are then randomly matched to form a unit mass of pairs. The only thing that is of consequence here is that an $\mathcal{O}$ agent should not know the "cost of cheating" (to be defined very shortly) of the agent she is matched with.[7]

Each agent $\mathcal{O}$ makes an offer $x \in [0,1]$ to the $\mathcal{R}$ agent in her match. The offer generates a total surplus of $2x$ to be split equally between $\mathcal{O}$ and $\mathcal{R}$ if the transaction goes through without "cheating" on the part of $\mathcal{R}$.[8] So, if $\mathcal{R}$ does not cheat, an offer of x generates a payoff of x for both $\mathcal{O}$ and $\mathcal{R}$.

It is the $\mathcal{R}$ agent in the match who decides whether to cheat or not. After receiving an offer x from the $\mathcal{O}$ agent in the match, $\mathcal{R}$ may decide to cheat and grab the entire surplus $2x$ instead of abiding by what the splitting procedure suggests. However, if she cheats, $\mathcal{R}$ will also suffer a cost c.[9] Therefore, $\mathcal{R}$ will cheat if $2x - c > x$ or equivalently $x > c$, and will not cheat otherwise.[10,11]

The total cost of cheating c has two components. One depends on the exogenously given "type" of the $\mathcal{R}$ agent and the other is determined by the behavior of other agents in the model.

For simplicity, we assume that there are just two types of $\mathcal{R}$ agents, "high" (H) and "low" (L).[12] The exogenous component of the cost of cheating is $t_L \in (0,1)$ for type L and $t_H \in (0,1)$ for type H, with $t_L < t_H$. The proportion of type H is denoted by $p \in (0,1)$ throughout.

The component of c that is "socially determined" is the same for all players.[13] Let s be the proportion of $\mathcal{R}$ agents who do *not* cheat. We simply set the social component of the

[7]Since we do not consider repeated interaction, the other details of the matching process are completely inessential. It should also be noted that, given the simplicity of our set up, it is easy to model our matching process in an effective way avoiding all the well known technical problems that can arise in the random drawing and matching of continuous populations of players. We omit the details. For a recent contribution and a substantial set of references see Duffie and Sun (2007)

[8]The fact that the surplus is split equally simplifies our calculations, but is completely inessential.

[9]Note that we are assuming that $\mathcal{R}$ has the choice of whether to cheat or not even when $x = 0$. If she cheats after an offer equal to zero, her payoff will therefore be $-c$. However, when $x = 0$, there is, so to speak, nothing to grab. Hence, an alternative would be to assume that $\mathcal{R}$ does not have a choice of whether to cheat or not when $x = 0$. Proceeding as we do simplifies the analysis but does not impact the results.

[10]We are normalizing any "production cost" for the $\mathcal{O}$ agent to be equal to zero. If by cheating the $\mathcal{R}$ agent is interpreted as appropriating the entire value — surplus, plus production cost — then c must be interpreted as *net* of any production cost.

[11]Our implicit assumption that when $\mathcal{R}$ is indifferent she will necessarily not cheat simplifies the analysis but is in fact without loss of generality. In equilibrium, the cheating set defined here and in (2) below must be open even if it were allowed in principle to be closed. The reason is that if it were not, then the optimal offer of $\mathcal{O}$ agents could not be defined because the acceptance set would have to be open.

[12]The overall flavor of our results easily generalizes to an world with any finite number of types. Many of our results also have analogues in a world with a continuum of types. We proceed in this way since one of our main aims is to keep the set up as simple as we possibly can.

[13]Again, this is the simplest way of proceeding. One could imagine heterogeneous cost "sensitivities" to the behavior of others, and this could easily be accommodated in our set up.

cheating cost to equal s and we take the two components of the cheating cost to combine in a linear way.[14] For an $\mathcal{R}$ agent of type $\tau \in \{L, H\}$, the total cost of cheating if a proportion s of transactions go through without cheating is given by

$$c(\tau, s, \alpha) \;=\; \alpha\, t_\tau \;+\; (1 - \alpha)\, s \tag{1}$$

with $\alpha \in (0, 1)$ a parameter that measures (inversely) the social sensitivity of the agents' cost of cheating.[15]

As we have already remarked, the social component s of c is a critical ingredient of our model. We think of it as embodying the influence of social norms on individual behavior. When fewer people in society cheat, those who do are in some sense further away from the social norm, and this has a "moral cost."

The effect of s on c can also be re-interpreted as a cost stemming from an enforcement technology. As fewer people cheat, for given resources devoted to enforcement, the probability that a cheater is "caught" increases, thus increasing the expected cost of cheating. We pursue this interpretation more formally in Section 5 below.

2.2. Equilibrium Definition

An equilibrium in our model is just a Nash equilibrium of the game we have described: a strategy profile σ^* describing for every player which offer she makes in her $\mathcal{O}$ role, and which offers she cheats and does not cheat on when she is in her $\mathcal{R}$ role, and such that no player has an incentive to unilaterally deviate.

A strategy profile σ assigns (in a measurable way) three numbers to every player $i \in [0, 2]$: the offer $x(i) \in [0, 1]$ that i will make if she is chosen to be an $\mathcal{O}$ agent, and a cheating cut-off value $z(i, \tau) \in [0, 1]$ indicating that, if she is of type $\tau \in \{L, H\}$, she will cheat if and only if she receives an offer strictly above $z(i, \tau)$ when she is chosen to be an $\mathcal{R}$ agent.[16] Notice

---

[14] As with our other modeling choices, this is just the simplest way of proceeding.

[15] Karni and Safra (2002) consider the overall preferences of an agent who is both "selfish" and "moral." The "moral" component in their paper is quite different from the social feedback we envisage here. Nevertheless, it is worth noting that in their analysis the two components combine *linearly* as a result of an actual *preferences representation theorem*.

[16] In principle, the players' cheating responses could be based on more than simple cut-off value. However, it is easy to show that we can restrict strategy spaces in this way without loss of generality using a standard weak-dominance argument.

that once a profile σ is given, a value of s is also given since the expected proportion of transactions that will not involve cheating is determined directly by σ.[17]

We begin the analysis of the players' maximization problem on the $\mathcal{R}$ side. Fix a σ, which, as we noted, also fixes a value of s. Consider an $\mathcal{R}$ agent of type $\tau \in \{L, H\}$. She will cheat if and only if the offer x she receives satisfies

$$x \;>\; c(\tau, s, \alpha) \;=\; \alpha t_\tau \;+\; (1-\alpha)s \tag{2}$$

For given σ, and hence s, using (2) we can compute the mass of agents $\mathcal{R}$ who will *not* cheat on any given offer $x \in [0,1]$. Let this be denoted by $P(x, s)$. Notice that we have

$$P(x,s) = \begin{cases} 0 & \text{if} \quad x \in (c(H,s,\alpha), 1] \\ p & \text{if} \quad x \in (c(L,s,\alpha), c(H,s,\alpha)] \\ 1 & \text{if} \quad x \in [0, c(L,s,\alpha)] \end{cases} \tag{3}$$

For given σ, we can therefore write the expected payoff of an $\mathcal{O}$ agent offering x as $x\,P(x,s)$ (recall that offering x, she gets a payoff of x whenever she is not cheated). Hence she will choose an x that solves

$$\max_{x\in[0,1]} \; x\,P(x,s) \tag{4}$$

The solution to (4) is immediate to characterize. Since $c(L,s,\alpha) < c(H,s,\alpha) < 1$, the solution to (4) depends on the comparison between

$$c(L,s,\alpha) \quad \text{and} \quad p\,c(H,s,\alpha) \tag{5}$$

If $c(L,s,\alpha) > pc(H,s,\alpha)$ then it is uniquely optimal to set $x = c(L,s,\alpha)$ — the largest level that ensures that no $\mathcal{R}$ agents cheat. If instead $c(L,s,\alpha) < p\,c(H,s,\alpha)$ the unique solution is to set $x = c(H,s,\alpha)$ — the largest level that ensures that only $\mathcal{R}$ agents of type L cheat. Finally if $c(L,s,\alpha) = p\,c(H,s,\alpha)$, then an $\mathcal{O}$ agent is indifferent between making an offer of $c(L,s,\alpha)$ and an offer of $c(H,s,\alpha)$, and hence both values solve the maximization problem (4).

[17]For each $i \in \mathcal{O}$ the distribution $z(\cdot,\cdot)$ of cut-off values across $\mathcal{R}$ agents and $x(i)$ determine the probability that i will be cheated by her partner. This is then averaged out across all $\mathcal{O}$ agents, to yield s.

Intuitively, increasing x increases (in jumps, because of the discrete nature of the types) the probability that the offer will be cheated on. The trade off between increased payoff conditional on not being cheated on and the increase in the probability of cheating is what determines the optimal behavior of $\mathcal{O}$ agents.

Before proceeding further, we introduce a simplifying assumption on the behavior of $\mathcal{O}$ agents.

Assumption 1. *Tie Break: Whenever $c(L, s, \alpha) = p\, c(H, s, \alpha)$, all $\mathcal{O}$ agents make an offer of $c(L, s, \alpha)$ which is not cheated on with probability one, rather than an offer of $c(H, s, \alpha)$ which is cheated on by all $\mathcal{R}$ agents of type L.*

For ease of exposition, form now on, when we say that "x solves (4)" we will mean a solution that complies with the tie-breaking rule posited here.

Assumption 1 makes the behavior of all $\mathcal{O}$ agents uniquely determined for any parameter values and any level of s, substantially simplifying the analysis.

Note that the behavior that Assumption 1 postulates can be interpreted as the result of "lexicographic" risk-aversion of the $\mathcal{O}$ agents (added to their basic risk-neutrality). Whenever expected values are equal (and only then), random variables with a lower risk are preferred. In particular, when $c(L, s, \alpha) = p\, c(H, s, \alpha)$, the sure payoff of $c(L, s, \alpha)$ will be preferred to a random payoff equal to 0 with probability $1 - p$ and to $c(H, s, \alpha)$ with probability p.

Assumption 1 simplifies our analysis since it rules out, for parameter configurations supporting multiple equilibria, a mixed equilibrium that is "intermediate" between the two that we will focus on. This equilibrium is "between" the two remaining ones, and its presence would not affect our qualitative conclusions in any way.[18] We will remark again informally on how the intermediate equilibrium would change some of the details as we go along.

We can now provide a working definition of what constitutes an equilibrium in our model. Note that, using Assumption 1, the equilibrium behavior of all $\mathcal{O}$ agents is summarized by a single number $x \in [0, 1]$ — the solution to (4), which is the offer they all make to the $\mathcal{R}$ agent they are each matched with. Of course, in equilibrium it must also be the case that the value of s that appears in (4) is the correct one, as determined by the behavior of the $\mathcal{R}$ agents given x. This justifies the following definition of equilibrium, which we will work with throughout the rest of the paper.

[18] Details are available from the authors on request.

Definition 1. *Equilibrium: An equilibrium is a pair (x, s) such that x solves (4) given s and such that*

$$P(x, s) \;=\; s \tag{6}$$

Given what we know about the behavior of $\mathcal{O}$ agents, it is also easy to check that only two possibilities are open for the value of s in any equilibrium. Setting an x such that both types of $\mathcal{R}$ agent cheat is clearly never optimal. Hence in equilibrium it must be that either $s = 1$ (and no cheating at all takes place), or $s = p$ (and all $\mathcal{R}$ agents of type L cheat, and all those of type H do not).

It follows easily that, in equilibrium, if $s = 1$ then $x = c(L, 1, \alpha) = \alpha t_L + 1 - \alpha$, and similarly if $s = p$ then $x = c(H, p, \alpha) = \alpha t_H + (1 - \alpha)p$.

At this point, it is useful to crystallize some terminology for future use.

Definition 2. *NC and LC Equilibria: An equilibrium (x, s) with $s = 1$ and $x = c(L, 1, \alpha)$ — in which no $\mathcal{R}$ agents cheat — is called a* No Cheating (NC) *Equilibrium. An equilibrium (x, s) with $s = p$ and $x = c(H, p, \alpha)$ — in which $\mathcal{R}$ agents of type L cheat — is called a* Low Cheating (LC) *Equilibrium.*

3. Equilibrium Characterization

3.1. NC Equilibrium

We can now work out the conditions under which the model has an NC equilibrium.

By definition in this case the equilibrium value of s is 1 — the probability of cheating is in fact 0. Therefore

$$c(L, 1, \alpha) \;=\; \alpha t_L + 1 - \alpha \quad \text{and} \quad c(H, 1, \alpha) \;=\; \alpha t_H + 1 - \alpha \tag{7}$$

We already know that in an NC equilibrium $x = c(L, 1, \alpha)$. Therefore, to confirm that $[c(L, 1, \alpha),\, 1]$ is an equilibrium, we just need to check that the parameters of the model are such that no $\mathcal{O}$ agent has an incentive to deviate unilaterally from offering $x = c(L, 1, \alpha)$ (which gives her a payoff of precisely $c(L, 1, \alpha)$ since no cheating takes place).

Deviating to an offer below $c(L, 1, \alpha)$ is never profitable since it yields a lower payoff conditional on no cheating taking place, but obviously cannot decrease any further the probability that cheating occurs. Deviating to an offer above $c(L, 1, \alpha)$, and hence accepting that $\mathcal{R}$ agents of type L will cheat, can yield at most a payoff of $p\, c(H, 1, \alpha)$. In fact this is what an $\mathcal{O}$ agent gets if she makes the largest offer that $\mathcal{R}$ agents of type H will not cheat upon, taking as given the equilibrium value of $s = 1$. Hence, using (7), a necessary and sufficient condition for an NC equilibrium to exist is that

$$\alpha t_L + 1 - \alpha \;\geq\; p\,[\alpha t_H + 1 - \alpha] \tag{8}$$

3.2. *LC Equilibrium*

The conditions under which an LC equilibrium exists can be worked out in a parallel way. By definition in this case, $s = p$. Hence

$$c(L, p, \alpha) \;=\; \alpha t_L + (1 - \alpha)p \quad \text{and} \quad c(H, p, \alpha) \;=\; \alpha t_H + (1 - \alpha)p \tag{9}$$

As we noted above, in equilibrium when $s = p$ it must be that $x = c(H, p, \alpha)$. To ensure that $[c(H, p, \alpha), p]$ is an equilibrium we then need to check that the parameters of the model are such that no $\mathcal{O}$ agent has an incentive to deviate unilaterally from offering $x = c(H, p, \alpha)$, which yields her an expected payoff of $p\,c(H, s, \alpha)$. With a logic that is by now familiar, without loss of generality we can consider only the deviation to offering $x = c(L, p, \alpha)$ — the largest offer that will induce no cheating from either type of $\mathcal{R}$ agent, taking as given the equilibrium value $s = p$. This deviation yields a payoff of $c(L, p, \alpha)$. Hence, using (9), a necessary and sufficient condition for the model to have an LC equilibrium is

$$p\,[\alpha t_H + (1 - \alpha)p] \;\geq\; \alpha t_L + (1 - \alpha)\,p \tag{10}$$

3.3. *Multiple and Unique Equilibria*

It is useful to sum up and sharpen our picture of the possible equilibria of the model as a function of the parameter quadruple (α, p, t_L, t_H). Purely for the sake of simplicity, from now on throughout the paper we restrict attention to quadruples away from the boundary of $[0, 1]^4$, satisfying $t_H > t_L$. This dispenses us from having to consider separately some of the boundary cases which would not add anything of interest to our results.

Proposition 1. *Equilibrium Set: The equilibrium set of the model is guaranteed to be non-empty, and can be of three types. A unique NC equilibrium, and in this case we say that we are in the NCU regime. A unique LC equilibrium, in which case we will say that we are in the LCU regime. Finally, there can be one LC and one NC equilibrium, and in this last case we will say that we are in the LCNC or simply "multiple equilibria" regime.*

If (8) is satisfied and (10) is not, then we are in the NCU regime. If (10) is satisfied and (8) is not then we are in the LCU regime. Finally, if (8) and (10) are both satisfied then we are in the NCLC multiple equilibria regime.

Although a formal proof of Proposition 1 does not require much more than using some of the observations we have already made, for the sake of completeness we present one in the Appendix. Note that in the multiple equilibria regime only two equilibria are possible because of our simplifying Assumption 1. As we mentioned above, without it we would get a third "intermediate" equilibrium, in which the proportion of transactions not cheated upon is strictly between p and 1.

The three equilibrium regimes of Proposition 1 are also all robust in the standard sense.

Proposition 2. *Parametric Conditions: The set of parameter quadruples (α, p, t_L, t_H) that yield the NCU regime contains an open set. The same is true for the set of quadruples yielding LCU, and for those yielding NCLC.*

A formal proof of Proposition 2 is in the Appendix. The simple argument behind it hinges on the fact that (8) and (10) can be jointly rewritten as

$$(1-\alpha)p(1-p) \leq \alpha(pt_H - t_L) \leq (1-\alpha)(1-p) \tag{11}$$

with the first inequality corresponding to (10) and the second to (8). Since $(1-\alpha)p(1-p) < (1-\alpha)(1-p)$ for all α and p in the interval $(0,1)$, we are guaranteed to find robust parameter configurations that support each of the three regimes.

We conclude this section with an observation, which we do not fully formalize purely for reasons of space.[19] Although they embody a level of trust s exactly equal to 1, the NCU regime and the NC equilibrium of the NCLC regime are less special than might seem at first sight. The fact that they yield $s = 1$ rather than just a "high" s is an artifact of our choice to consider only two types $\tau \in \{L, H\}$ of $\mathcal{R}$ agents in order to keep the model as simple as possible.

Suppose that a small proportion of a third "very low" type were introduced, with the exogenous component of the cost of cheating between 0 and t_L. Then it would be easy to check that the same parameters supporting an NC equilibrium in the two-types model would

[19]The details available from the authors upon request.

yield an equilibrium in which only the "very low" type of $\mathcal{R}$ agents cheat. This is clearly possible in a completely straightforward way whenever the model yields an NC equilibrium which is "strict" in an appropriate sense. With this observation in mind, we will often interpret the NC equilibria as "high trust" equilibria of a general kind, rather than focusing specifically on the fact that they display "full trust" ($s = 1$).

4. Societies With Different Levels of Trust

4.1. Differences in Fundamentals vs. Equilibrium Switch

As we noticed above, differences of trust levels across different societies are well documented. Moreover, they are often seen to be correlated with (and sometimes held responsible for) phenomena of primary economic importance like income levels and growth rates.

Our model can generate such differences in two conceptually distinct ways.

The first is when two societies are characterized by the *same* parameter quadruple generating the NCLC regime, the LC equilibrium is realized in one of the societies (the low trust one) while the NC equilibrium is realized in the other (the high trust society). We refer to this case as an *equilibrium switch*.

The second possibility is that different levels of trust across two societies reflect the fact that the two societies are characterized by *different* parameter quadruples, leading to different levels of equilibrium trust. This in turn can happen in several different ways, as the two quadruples can be in different regions of the parameter space supporting different equilibrium regimes, or they might support the same equilibrium regime but with different equilibrium trust levels. These cases are what we refer to as a *difference in the fundamentals.*

Among the many ways in which a difference in fundamentals can arise it is useful to single out the case in which only the probability mass of the two types of $\mathcal{R}$ agents changes, with all other parameters constant. This we refer to as a difference in the *distribution of types*. Specifically, in one society the parameter quadruple is $(\alpha, \overline{p}, t_L, t_H)$, in the other is $(\alpha, \underline{p}, t_L, t_H)$, and in both the LC equilibrium obtains.[20] In the first case the equilibrium trust level is $\overline{s} = \overline{p}$ and in the second it is $\underline{s} = \underline{p} < \overline{s}$.

The distinction between differences in the equilibrium level of trust that arise from an equilibrium switch or from a difference in fundamentals is more than a curiosity. These two possibilities would in fact call for distinct policy stances from a Government seeking to

[20]So, clearly, both parameter quadruples must give rise to either the LCU or the NCLC regimes.

intervene to rectify a low level of trust. If the low trust level is the result of an equilibrium switch, then the Government can attempt to engender a switch to a higher level of trust by re-focussing the *expectations* of the players. Convincing everyone in society that the level of trust is $\overline{s}$ instead of $\underline{s}$ will do the job since this is capable of becoming a self-fulfilling prophecy.

If instead the low level of trust is due to a difference in fundamentals then the only way to ensure a high level of trust is to engineer a change in the parameters of the model. This is conceptually and operationally different from a self-fulfilling change in beliefs, and likely harder to achieve, although one might imagine a range of tools that vary from economic incentives to educational programs based on "civic culture" that could be brought to bear.

The conceptual difference between the two hypotheses, difference in fundamentals versus equilibrium switch, might also be of interest to scholars in other disciplines, such as political science or sociology. In essence, a difference in fundamentals points to different norms of behavior being rooted in "anthropological" factors, while an equilibrium switch points in the direction of random events triggering changes that become long-lasting because of self-reinforcement mechanisms at work in society.

4.2. Trust and Transaction Levels

Suppose that we observe two societies, one with a high and the other with a low level of trust. Can we identify whether the differing trust levels are due to an equilibrium switch or to a difference in the fundamentals?

Somewhat surprisingly, the answer is a qualified *yes*. As we mentioned in the introduction, *partial* identification is in general possible, as there are some observations that allow us to rule out an equilibrium switch.[21] Full identification can be achieved if the range of possible alternatives can be further restricted in an appropriate way. In particular, if we know that the only differences in the fundamentals that need to be considered are variations in the distribution of types, then full identification becomes possible.

We begin with the general case, in which all possible variations in parameter quadruples are considered.

The key to identification — partial or full, as the case may be — is the equilibrium level of x, the offer made in equilibrium by all the $\mathcal{O}$ agents, which is also the equilibrium individual transaction level.

[21] See footnote 3 above.

The following two propositions make our claim precise for the general case

Proposition 3. *Individual Transaction Levels — Equilibrium Switch: Let a parameter quadruple supporting the NCLC regime be given. Then the level of x in the associated NC equilibrium is* lower *than the level of x in the associated LC equilibrium.*

It follows that if the difference in equilibrium trust between two societies is due to an equilibrium switch *in the sense of Subsection 4.1, then the equilibrium level of x is* negatively correlated *with the equilibrium trust level.*

When it comes to a difference in fundamentals, in principle we should consider any pair of equilibria, of the NC and/or LC type. For completeness, the following proposition examines four cases. The bottom line is that when we allow for unrestricted differences in fundamentals, the relationship between differences in the equilibrium trust level and the level of individual transaction cannot be pinned down.

Proposition 4. *Individual Transaction Levels — Change in Fundamentals: Fix arbitrarily a parameter quadruple supporting an NC equilibrium. Then we can find parameter quadruples that support LC equilibria with individual transaction levels that can be both higher and lower than in the given NC equilibrium.*

Fix again arbitrarily a parameter quadruple supporting an NC equilibrium. Then we can find parameter quadruples that support NC equilibria with individual transaction levels that can be both higher and lower than in the given NC equilibrium.

Fix arbitrarily a parameter quadruple supporting an LC equilibrium. Then we can find parameter quadruples that support NC equilibria with individual transaction levels that can be both higher and lower than in the given LC equilibrium.

Finally, fix again arbitrarily a parameter quadruple supporting an LC equilibrium. Then we can find parameter quadruples that support LC equilibria displaying a higher level of trust and a lower level of individual transactions than in the original LC equilibrium, or a lower level of trust and a higher level of individual transactions than in the original LC equilibrium.

The picture yielded by Propositions 3 and 4 — namely, partial identification — changes considerably when we consider the case in which parameter differences are restricted to be differences in the distribution of types in the sense mentioned above. In this case the ambiguity highlighted by Proposition 4 no longer holds and a definite conclusion can be reached about the correlation between trust and the size of individual transactions following the difference in parameter values.

Proposition 5. *Transaction Levels — Change in Distribution of Types: Consider two parameter quadruples $(\alpha, \overline{p}, t_L, t_H)$ and $(\alpha, \underline{p}, t_L, t_H)$ with $\overline{p} > \underline{p}$, each giving rise to an LC equilibrium with high ($s = \overline{p}$) and low ($s = \underline{p}$) trust levels respectively.*

Then the equilibrium level of x associated with $(\alpha, \overline{p}, t_L, t_H)$ is higher than the equilibrium level of x associated with $(\alpha, \underline{p}, t_L, t_H)$.

It follows that if the difference in equilibrium trust between two societies is due to a difference in fundamentals in the narrower sense of a difference in the distribution of types, *then the equilibrium level of x is* positively correlated *with the equilibrium trust level.*[22]

Formal proofs of Propositions 3, 4 and 5, which, again, consist of fairly simple algebra, appear in the Appendix. Here, we elaborate on the intuition behind our results, beginning with Proposition 4.

Recall that in any NC equilibrium the level of individual transaction is $\alpha t_L + 1 - \alpha$ while in any LC equilibrium it is given by $\alpha t_H + (1 - \alpha)p$. Proposition 4 is then the result of the following observations. If we are allowed to vary all the parameters at will, the necessary and sufficient conditions (8) for an NC equilibrium to exist are compatible with any value of $\alpha t_L + 1 - \alpha$ in $(0, 1)$. Similarly, if we are allowed to vary all the parameters at will, the necessary and sufficient conditions (10) for an LC equilibrium to exist are compatible with any value of $\alpha t_H + (1 - \alpha)p$ in $(0, 1)$.

The statement of Proposition 5 is an almost immediate consequence of the fact that in any LC equilibrium the level of individual transactions is given by $\alpha t_H + (1 - \alpha)p$. Hence, since all other parameters are kept constant, it must increase as p increases.

The intuition behind Proposition 3 requires some intermediate steps. Recall that in this case we are concerned with a *single* parameter quadruple supporting the NCLC regime. Call the level of individual transactions in the LC equilibrium $x_{LC} = \alpha t_H + (1 - \alpha)p$. In the LC equilibrium an $\mathcal{O}$ agent gets an expected payoff of $p\, x_{LC}$ since with probability $1 - p$ he is cheated by the $\mathcal{R}$ agent he meets and ends up with a payoff of zero. She could however deviate and make the largest offer that will induce *no* cheating from any type of $\mathcal{R}$ agents. Denote this by x^D_{LC}, and observe that it must equal $\alpha t_L + (1 - \alpha)p$. By deviating, she would get a payoff equal to her offer for sure, hence her incentive constraint tells us that it must be the case that $p\, x_{LC} \geq x^D_{LC}$.

[22]Note that once we fix all parameters except for p, all NC equilibria are the same. Hence, since the effect of a switch from LC to NC is already characterized by Proposition 3, the only relevant comparison is the one treated here — two LC equilibria. All the other cases we treated in Proposition 4 can be ignored.

Now consider the NC equilibrium associated with the given parameter quadruple. The equilibrium level of individual transactions now is $x_{NC} = \alpha t_L + 1 - \alpha$, since this is the largest offer that will keep *all* $\mathcal{R}$ agents from cheating. The only difference between x_{LC}^D and x_{NC} is in fact given by the higher level of equilibrium trust reflected in x_{NC}.[23] This makes it immediate to check that we have $x_{LC}^D > p\ x_{NC}$.

The two steps we have outlined give that $p\ x_{LC} \geq x_{LC}^D$ and $x_{LC}^D > p\ x_{NC}$ respectively, which together immediately yield the claim of Proposition 3, namely $x_{LC} > x_{NC}$.

Before moving on, it is important to remark on the fact that the results in this subsection are much less special than they seem at first sight. While the algebra used here of course rests on the specific features of our model, including for instance the linear form of the cost of cheating, the basic logic behind the arguments generalizes considerably.

In particular, the logic of Proposition 3 relies on the fact that the set of types who cheat in the two equilibria is different — it is larger in the low trust equilibrium than in the high trust one. The incentive constraint of $\mathcal{O}$ agent in the low trust equilibrium is always going to entail a comparison with what she can get if the offer is lowered so as to induce a larger set of types not to cheat — a lower transaction (and hence payoff) with a higher probability (x_{LC}^D above). This lower transaction in general is going to differ from her transaction level in the high trust equilibrium (x_{NC} above) only because of the effect of higher equilibrium trust, via the socially determined component of the cost of cheating, but not the idiosyncratic one. The critical comparison then becomes the one between the payoff to the $\mathcal{O}$ agent in the high trust equilibrium, but multiplied by the probability of not cheating in the *low* trust equilibrium ($p\ x_{NC}$ above), and the deviating payoff (x_{LC}^D above). The idiosyncratic component of the cost of cheating is the same in both cases, while the lower trust level multiplies both components in the former (as a probability), but only enters the socially determined component of the cost of cheating in the latter. So long as this is sufficient to establish the analogue of $x_{LC}^D > p\ x_{NC}$, the conclusion of Proposition 3 remains valid.

4.3. Trust and Aggregate Transactions

The key result of Subsection 4.2 above is that the level of trust and transaction level x are negatively correlated in the case of an equilibrium switch.

On the other hand, a recurrent theme in the extant literature (see Subsection 1.2 above)

[23] The term $1 - \alpha$ is multiplied by p in the case of x_{LC}^D. It is not in the expression for x_{NC}.

is that of a positive relationship between trust and income levels and/or growth rates. Does it then follow that these findings exclude switches across multiple equilibria as the root of different trust levels in the societies that have been examined?

The answer is no. And this is because the transaction level x in our simple model is not the proper analogue of "GDP." The analogue of GDP in our simple model is clearly not the equilibrium level of x (an "intensity" of activity index), but rather the equilibrium level of x times the probability that no cheating occurs and hence that the surplus from the transaction does in fact materialize (so, x times an "extensiveness" measure).

Because of the extreme simplicity of our set up, the aggregate level of transactions is easily computed. In an LC equilibrium it will be the equilibrium value of x times p since a proportion $1 - p$ of transactions is cheated on. In an NC equilibrium, since all transactions produce surplus because no one cheats, the equilibrium value of x is the appropriate measure of aggregate transactions instead.

Our model predicts that the correlation between equilibrium trust and the aggregate level of transactions is positive in the case of an equilibrium switch.

Proposition 6. *Trust and Aggregate Transactions: Consider the equilibrium switch of Proposition 3, for a given parameter quadruple giving rise to the NCLC regime. Then, the equilibrium aggregate level of transactions is lower in the associated LC equilibrium than in the associated NC equilibrium.*

The proof of Proposition 6 is in the Appendix. As before, simple algebra is sufficient to formalize the argument.

4.4. Welfare

Despite its simplicity, in our model we can draw a meaningful distinction between "GDP" (the aggregate level of transactions of Subsection 4.3) and welfare measures. This is because our $\mathcal{R}$ agents cheat when they find it *optimal* to do so. Hence, more cheating does not automatically reduce aggregate welfare in society. A proper piece of welfare calculus is needed to draw a conclusion.

Of course, welfare comparisons between to societies with different fundamentals which — among other things — imply a variation in the players' preferences are not particularly meaningful. Therefore, we confine ourselves to comparing welfare across the two different equilibria when the model is in the NCLC regime.

Before proceeding any further we have to be precise about aggregate welfare. As seems natural in this context, we focus on the utilitarian benchmark. The expected utilities of $\mathcal{R}$ agents are weighted according to the mass of the two types L and H, and then added to the (degenerate average) expected utility of $\mathcal{O}$ agents. Recall that each transaction of amount x that is executed without cheating generates a payoff of x for each side; hence a total surplus of $2x$. Each offer x that is cheated upon by the $\mathcal{R}$ agent instead produces a payoff of 0 for the $\mathcal{O}$ agent and a payoff of $2x - c$ (with c the cost of cheating) for the $\mathcal{R}$ agent. It follows easily that aggregate welfare in an NC equilibrium where all $\mathcal{O}$ agents offer x_{NC} is equal to $2x_{NC}$. In an LC equilibrium the $\mathcal{R}$ agents of type L (who have mass $1 - p$) cheat while those of type H do not. Hence in an LC equilibrium in which all $\mathcal{O}$ agents make an offer of x_{LC}, aggregate welfare is equal to $p2x_{LC} + (1-p)[2x_{LC} - c(L, p, \alpha)]$. Using (9) aggregate welfare in this case can therefore also be written as $2x_{LC} - (1-p)[\alpha t_L + (1-\alpha)p]$.

The bottom line is that the natural welfare comparison between the high trust NC equilibrium and the low trust LC equilibrium is in fact ambiguous.

Proposition 7. *Aggregate Welfare in the NCLC regime: There exist parameter quadruples that give rise to the NCLC regime and such that aggregate welfare is larger in the associated NC equilibrium than in the associated LC equilibrium. There are also parameter quadruples that give rise to the NCLC regime and such that aggregate welfare is lower in the associated NC equilibrium than in the associated LC equilibrium.*

The Appendix contains a formal proof of Proposition 7. As before, simple formal manipulations are all that is required.

5. Enforcement

As we mentioned already, the component of the cost of cheating c that is socially determined — via s — can be reinterpreted as arising from an enforcement technology with limited resources for catching and punishing the $\mathcal{R}$ agents who cheat. This is quite different from the social norm interpretation we gave before, but the formalisms are surprisingly close in the two cases.

As previously, we seek to proceed in the simplest possible way.

Assume that there is an enforcement agency with resources $k \in [\underline{k}, 1]$.[24] The parameter k pins down the capacity of the enforcement mechanism in the sense that a mass k of $\mathcal{R}$

[24] We take k to be a number strictly between 0 and 1. The convenience of assuming that k be bounded away from 0 will be apparent shortly; see footnotes 26 and 29 below.

agents can be checked and fined. Note that, for simplicity, we assume that the enforcement is perfectly targeted; no resources are wasted on $\mathcal{R}$ agents that do not cheat.[25] A mass $\min\{k, 1-s\}$ of $\mathcal{R}$ agents who cheat is randomly caught and fined. A useful analogy is that of checks for speeding on the highway. Only cars that are actually above the speed limit are stopped, and when they are stopped they are fined. However, the number of cars that can in fact be stopped and fined is limited by the capacity of the Police to deploy its patrol cars.

The speeding analogy is also useful to see intuitively how the mechanics of the social feedback on the cost of cheating work in this case. Given that the Police have a fixed capacity for stopping and fining speeding cars, if a large percentage of the cars on the road actually speed it will be impossible for the Police to stop all of them. Other things equal, the probability of being caught and fined will be lower when more cars actually speed. Conversely, when very few cars actually speed, the fixed capacity of the Police will ensure that many of them will be caught. In fact, once the mass of speeding cars is equal or less than the Police capacity, the probability of being caught for those speeding will be one.

It is convenient to normalize the size of the fine to be equal to one. The probability of being caught and fined for an $\mathcal{R}$ agent who cheats is then given by

$$z \quad = \quad \min\left\{1, \frac{k}{1-s}\right\} \tag{12}$$

Since the fine is normalized to one, z is also the expected fine, or the expected cost of cheating, coming from the enforcement mechanism.[26]

As we did previously, we assume that z is combined in a linear way with a cost of cheating arising from the $\mathcal{R}$ agent's type.[27] We then obtain that the total cost of cheating now is

$$c(L, s, \alpha) = \alpha t_L + (1-\alpha)z \quad \text{and} \quad c(H, s, \alpha) = \alpha t_H + (1-\alpha)z \tag{13}$$

The logic of Section 3 applies to this reinterpretation of the model virtually unchanged. Only NC and LC equilibria are possible. Condition (8) is still necessary and sufficient for an NC equilibrium to exist.

[25] Our results easily generalize to the case in which the proportion of enforcement resources that are targeted towards the mass s of $\mathcal{R}$ agents that do not cheat is non-increasing in s.

[26] Note that since $k \geq \underline{k} > 0$, we get that $z = 1$ whenever $s = 1$. See footnote 24 above.

[27] As before, this is just the simplest way of proceeding. It also yields immediate comparability with the set up of Subsection 2.1.

When $s = p$, we must replace (9) with (13). Hence, condition (10) must be replaced by

$$p\left[\alpha t_H + (1-\alpha)\min\left\{1, \frac{k}{1-p}\right\}\right] \geq \alpha t_L + (1-\alpha)\min\left\{1, \frac{k}{1-p}\right\} \tag{14}$$

which is now necessary and sufficient for an LC equilibrium to exist.

Proposition 1 still holds provided that we replace condition (10) with condition (14).

Note next that if we set $k = p(1-p)$ the new condition (14) coincides with the old condition (10). Hence we can also conclude that, at least for an open interval of values of k around $p(1-p)$, Proposition 2 still holds.

It is then immediate to see that Propositions 3, 4, 5 and 6 are also still valid in this reinterpretation of the model.

The basic conclusion of Proposition 7 — that the effect of an equilibrium switch on aggregate welfare is ambiguous — is still valid in the case of enforcement we are considering here. The details are somewhat different however. The part of the cost of cheating that corresponds to the fine must now be treated differently from the socially generated cost of cheating in the social norms version of the model. The fine is a *transfer* from one agent to another (from the $\mathcal{R}$ agent who cheats to the enforcement agency, whatever shape and form it might take), and hence washes out of the aggregate welfare calculation, instead of generating a net decrease in aggregate welfare as in the case of the of social norms version of the model.[28] Hence, it is easy to see that aggregate welfare in an LC equilibrium is given by $2x_{LC} - (1-p)\alpha t_L$, with x_{LC} now equal to (15) below.

It is interesting to track what happens to the features of an LC equilibrium as k changes. We begin with the obvious observation that in the model with enforcement in an LC equilibrium the transaction level is equal to

$$\alpha t_H + (1-\alpha)\min\left\{1, \frac{k}{1-p}\right\} \tag{15}$$

since this is the largest offer that will induce the $\mathcal{R}$ agents of type H not to cheat. Hence, as we track the LC equilibrium, an increase in k guarantees both an increase in the individual transaction level, and an increase in aggregate transactions since the latter is just equal to (15) multiplied by p.

[28]The idiosyncratic part of the cost of cheating obviously remains as before

At this point, it would be tempting to use the model to investigate what is the optimal level of resources devoted to enforcement.[29] Following standard procedure, the latter would be determined by comparing the marginal cost of increasing k with its marginal benefit, i.e. the marginal increase in the aggregate welfare of the LC equilibrium as k raises.[30] However, given the simplicity of our model, it would be hard to specify a sufficiently reliable marginal cost function for k. Moreover, the measure of aggregate welfare we defined, though consistent with the model, clearly reflects our extremely stylized approach to the problem. On both counts, it would be reckless to base detailed policy conclusion on the exercise.

There is however a case in which the conclusion that can be drawn is sufficiently strong to deserve special attention in our view. This is what we turn to next.

The argument involves again the idea of an equilibrium switch that we discussed above.

Consider a parameter quadruple (α, p, t_L, t_H) such that condition (8) that guarantees the existence of an NC equilibrium is satisfied as a strict inequality. Note that this is equivalent to $\alpha(pt_H - t_L)/(1-\alpha) < 1-p$. Hence condition (14) that is necessary and sufficient for an LC equilibrium to exist is satisfied if and only if k is such that

$$k \leq \mathbf{k}(\alpha, p, t_L, t_H) = \frac{\alpha(p\, t_H - t_L)}{1-\alpha} \tag{16}$$

Now consider the following policy question. We are given a quadruple (α, p, t_L, t_H) such that condition (8) is satisfied strictly. We are also told that society is in the LC equilibrium and that k is equal (or below and very close) to the threshold level $\mathbf{k}(\alpha, p, t_L, t_H)$ given in (16). We are also told that the marginal cost of an increase in k is *finite*. Lastly, we are in a setting in which the policy-maker is interested in the aggregate transaction level — the GDP measure of Subsection 4.3 above.[31] The question then is whether we recommend a *local* policy change — a small increase in k.

The answer must be "yes." The reason is simply that after the increase in k, condition (14) will be violated. It then follows that the policy will force an equilibrium switch. The small increase in k will ensure that society switches from the LC equilibrium to the only

[29]Note that in an NC equilibrium the question is not a very interesting one. Since $s = 1$, the enforcement-generated component of the cost of cheating z in an NC equilibrium is equal to 1. Hence the optimal k is simply $\underline{k}$. See also footnote 24.

[30]Given (15) and the fact that aggregate welfare is computed as we discussed just above, the marginal welfare increase is $2(1-\alpha)/(1-p)$ if $k < 1-p$ and zero if $k \geq 1-p$.

[31]Notice that Proposition 6 still holds in the model with enforcement we are considering here.

remaining one — namely the NC equilibrium. Since the marginal cost of k is finite, a "small" increase in k must carry a correspondingly "small" cost. However, by Proposition 6 there will be a discrete jump up in the aggregate level of activity. Hence the policy change must be worthwhile.

6. Robustness

Our next goal is to investigate whether the high and low trust equilibria have different robustness attributes — different degrees of resilience to small changes in the environment in which they emerge.

There are many ways to proceed, as there are obviously many "robustness tests" that one might devise. We proceed with the "myopic belief revision" analysis below for at least three reasons. The first is that in our view it fits well our motivation of investigating the tension between "trusting beliefs" and "trusting behavior" that we mentioned above in the introduction to the paper. The second is that it seems more novel in nature than some of the alternatives — for instance payoff-based adaptive dynamics (we return to this point briefly below). And last, but certainly not least, is that the analysis can be carried out in a very simple way.

We return to the social norms version of the model, and, throughout this section, we assume that the parameter quadruple is one that sustains the NCLC regime, so that both equilibria are in fact possible. The question is whether one is more "likely" than the other in some coherent sense of the word.

Imagine that the population of players is divided into two sets. A fraction $q \in (0, 1)$ of the players believes "myopically" that the NC equilibrium prevails.[32] A fraction $1-q$ instead believes that the LC equilibrium prevails. These subsets are decided before any draws that determine the players' $\mathcal{O}$ and $\mathcal{R}$ roles, or their type H or L. Hence there are fractions q and $1-q$ of players with these beliefs in each possible role and type. For short, we will say that a player or an agent "believes NC" or "believes LC," as appropriate.

The players' beliefs about which equilibrium prevails affect the socially determined component of their cheating cost. The idiosyncratic component is given by their type as before. Hence the $\mathcal{O}$ agents, can be partitioned in four sets: the H types who believe NC, the H

[32]The reason we call (as is standard) these beliefs "myopic" is that, as will be apparent shortly, they are not correct.

types who believe LC, the L types who believe NC and the L types who believe LC. For these sets, we get the following shares of the population with corresponding cheating costs

$$\begin{array}{ll} \text{Population Share} & \text{Cheating Cost} \\ q\,p & \alpha t_H + 1 - \alpha \equiv c_1 \\ (1-q)\,p & \alpha t_H + (1-\alpha)\,p \equiv c_2 \\ q\,(1-p) & \alpha t_L + 1 - \alpha \equiv c_3 \\ (1-q)(1-p) & \alpha t_L + (1-\alpha)\,p \equiv c_4 \end{array} \tag{17}$$

Since we are in the NCLC regime, (10) must be satisfied. Using this together with $t_H > t_L$, we immediately get that

$$c_1 \; > \; c_2 \; > \; c_3 \; > \; c_4 \tag{18}$$

Now consider an $\mathcal{O}$ agent who believes NC. Given her beliefs, she will find it optimal to make an offer of $x_{NC} = \alpha t_L + 1 - \alpha$, just as in Subsection 3.1 above. Moreover, an $\mathcal{O}$ agent who believes LC, just as in Subsection 3.2 above, will find it optimal to make an offer $x_{LC} = \alpha t_H + (1-\alpha)\,p$.

Using (17) and (18) we now can work out the probabilities that each of the offers x_{NC} and x_{LC} will be cheated on.

Since $x_{NC} = c_3$, the $\mathcal{R}$ agents with costs of cheating equal to c_1, c_2 or c_3 will not cheat upon receiving the offer x_{NC}. On the other hand the $\mathcal{R}$ agents with cost of cheating c_4 will find it optimal to cheat on it. In sum, the $\mathcal{R}$ agents who believe LC and are of type L will cheat upon receiving the offer x_{NC}, while the other $\mathcal{R}$ agents will not. This pins down the probability that an $\mathcal{O}$ agent who believes NC and hence offers x_{NC} will be cheated on to be equal to $(1-q)\;(1-p)$.

Analogously, since $x_{LC} = c_2$, the $\mathcal{R}$ agents with costs of cheating equal to c_1 or c_2 will not cheat upon receiving the offer x_{LC}. On the other hand the $\mathcal{R}$ agents with costs of cheating equal to c_3 or c_4 will find it optimal to cheat on it. In sum, the $\mathcal{R}$ agents who are of type L, irrespective of whether they believe NC or LC will cheat upon receiving the offer x_{LC}, while the other $\mathcal{R}$ agents will not. Hence, the probability that an $\mathcal{O}$ agent who believes LC and hence offers x_{LC} will be cheated on is equal to $(1-p)$.

The next step is to compare the actual probabilities that the $\mathcal{O}$ agents will be cheated with the myopic beliefs we assigned them to begin with. A natural mechanism to imagine here is that if the former is larger the the latter, then the share of players with the given myopic belief will go down while it will go up if the reverse is the case. In other words, if the "empirical frequency" of cheating is different from what players expect, then in the aggregate the share of the population that holds the given belief will change: it will go down if the empirical frequency of cheating is larger than what the players expected and it will go up if it is lower. The details of the adjustment rule are largely unimportant, so long as it is "monotone" in the sense we have described, but to fix ideas we will postulate a very simple one.

Start with a proportion q_0 of players who believe NC and a corresponding proportion of players $1 - q_0$ who believe LC. Time is discrete and runs as $0, 1, \ldots, t, \ldots$ with the share of NC believers at t denoted q_t. Now suppose that all the players who believe NC who are in fact cheated switch their belief to LC. This is an extreme assumption, but, after all, they expect to be cheated with probability zero, and they are actually cheated by the $\mathcal{R}$ agent they are matched with.[33] In this extreme case we get

$$q_{t+1} \;=\; q_t \left[1 \,-\, (1-q_t)\,(1-p)\right] \tag{19}$$

From (19) it is immediate that as t becomes large q_t converges to zero. Hence, in this simple case, the fraction of the population of players who believe NC shrinks through time, approaching zero as t goes to infinity.

Before commenting further on the result, we reiterate that the same will clearly be true for a very large class of dynamics that are monotone in the difference between what the players expect (believe NC or believe LC) and the empirical frequency of cheating.

The fact that $q_t \to 0$ as t becomes large pins down the NC equilibrium as less robust than the LC one. To see intuitively the general phenomenon that we are capturing here recall that, critically, the players beliefs about the behavior of others affect their cost of cheating. In other words, those players who believe in LC will have a lower cost of cheating than those who believe in NC. In general, those players who believe in lower trust equilibria will have a lower cost of cheating than those who believe in higher trust equilibria.

[33]To see that it is admissible to only consider the switch away from believing NC as driving the dynamics, recall that the players who believe LC expect to be cheated with probability $1 - p$ and are in fact cheated with frequency precisely equal to $1 - p$.

Now imagine that we "perturb" a high trust equilibrium by adding a small fraction of players who believe that society is in a low trust equilibrium. Since the "invaders" have a lower cost of cheating, they will cheat in situations in which the players who believe in the high trust equilibrium do not. Hence, the overall frequency of cheating in society will go up. Under "monotonic" belief dynamics, the population share of players who believe in the low trust equilibrium will go up. This will set off another round of increase in cheating, and so on.

The case of a low trust equilibrium being perturbed with an invasion by a small mass of players who believe in a high trust equilibrium is a very different one. Notice that in this case the proportion of $\mathcal{R}$ agents who cheat upon receiving the offer x_{LC} (made by those $\mathcal{O}$ agents who believe LC) does *not* depend on q — the fraction of players who believe NC. The reason is that in the LC equilibrium the offer x_{LC} — as is familiar by now — is the largest offer that keeps the $\mathcal{R}$ agents of type H from cheating, given that they believe LC. Switching some $\mathcal{R}$ agents of type H to believe in NC, will *raise* their cost of cheating, and hence will not introduce any new cheaters. After the "invasion" by a small population of agents who believe NC, the empirical frequency of cheating will remain unchanged, since the empirical frequency of being cheated still equals the one expected by the $\mathcal{O}$ agents.

Before moving on, we mention briefly another robustness exercise that we do not report in any detail for reasons of space, but which also points in the direction of high trust equilibria being less robust than low trust ones. The idea is to start again with some players who myopically believe NC and some others who myopically believe LC, but to base the dynamics of their respective shares of the population on the expected *payoffs* that accrue to each combination of type (H or L) and myopic belief.[34] In this case too, the results point in the direction of low trust equilibria being more robust than low trust ones. Intuitively, the players who believe LC have, other things equal, lower cheating costs and this gives them an advantage in the payoff based dynamics over the players who believe NC.

[34]There is a vast literature on learning/evolutionary game theory that concerns dynamics related to the ones we are sketching out here, and is also related to our previous exercise. This is clearly not the place to even attempt a survey, but we refer the interested reader to the two monographs by Weibull (1995) and Samuelson (1997) and the many references cited there.

7. Summary and Conclusions

Our main purpose was to explore the tension between trusting beliefs and consequent trusting behavior, and the incentives to cheat by requiring that the level of trust in society be the fraction of transactions that are not cheated upon *in equilibrium*.

Our results demonstrate that a simple social feed-back mechanism is sufficient to generate a rich pattern of possible equilibria that shed light on the issue of equilibrium trust. We set up a simple static model in which one side of a transaction can cheat and walk away with the entire surplus, but must suffer a cost of cheating with a socially generated as well as an idiosyncratic component. The socially generated component captures the idea that social norms will make it more costly to cheat when a smaller fraction of the population engages in cheating behavior.

Our findings point to the fact that there are two main possible sources of differences in trust levels across societies. One is the *multiplicity* of equilibria, that can arise for given *fundamental* parameters. The other is a difference in the fundamental parameters themselves. Surprisingly, the two regimes are empirically partially identifiable in general, and fully identifiable under some further restrictions. The key variable to identification is the equilibrium size of individual transactions, which must be negatively correlated with the equilibrium level of trust when the source of different trust levels are multiple equilibria.

The positive relationship between trust and aggregate income levels (our static model has little to say about growth of course) that has been claimed in the literature is *not* negated by the negative correlation associated with a switch across multiple equilibria. The reason is that, in this case, when individual transaction levels go down, the frequency of transactions increases.

Interestingly, since our agents cheat if and only if they find it optimal to do so, the relationship between aggregate income and aggregate welfare is not unambiguous in our model. As society switches across multiple equilibria with high and low trust, aggregate welfare could go up as well as down.

Our model can also be reinterpreted as one with standard selfish preferences, in which an enforcement agency uses limited resources to catch and punish those agents who cheat. Because of the limited resources devoted to enforcement, when the fraction of agents who cheat is higher, it is less likely that cheating agents are caught. This produces a social feed-back mechanism that is the same as the one in the social norms case, and hence all the same results are valid in this reinterpretation of the model.

Finally, we report explicitly on one robustness exercise which indicates that high-trust equilibria are more vulnerable than low-trust ones. Introducing a small fraction of agents who believe in the low-trust equilibrium in a population of agents who believe and play according to the high-trust one can destroy the high-trust equilibrium. The reason is that the invaders have a lower cost of cheating precisely they believe in the low-trust equilibrium — a less demanding social norm. This pushes up the fraction of transactions that are cheated on, which in turn causes the agents who believe in the high-trust equilibrium to revise down their beliefs.

Appendix

Proof of Proposition 1: To see that the equilibrium set is always not empty, suppose first that

$$t_L \;\geq\; p\, t_H \tag{A.1}$$

In this case, inequality (8) is clearly satisfied and hence an NC equilibrium exists. Next, suppose that (A.1) is violated. Then the middle term of (11) is positive. Since the first term in (11) is positive and strictly less than the third term in (11), one or both of the inequalities in (11) must be satisfied. Since the first inequality in (11) is the same as (10) and the second is the same as (8), we must have that either an LC or an NC or both equilibria exist. Hence the equilibrium set is always not empty.

To see that only the three NCU LCU and LCNC regimes are possible, we only need to argue that there are no equilibria other than the LC and the NC ones.

As we noted in the text, the solution to (4) consistent with Assumption 1 is unique and can only take the values $x = c(L, s, \alpha)$ or $x = c(H, s, \alpha)$. Together with (6), this is sufficient to prove the claim. ■

Proof of Proposition 2: We have already noted that the first inequality in (11) is the same as (10) and the second is the same as (8).

Hence, any parameter quadruple such that

$$\alpha\,(pt_H - t_L) \;<\; (1-\alpha)\,p\,(1-p) \tag{A.2}$$

must yield the NCU regime. Hence the set of parameter quadruples that yield the NCU regime must contain an open set.

Moreover, any parameter quadruple such that

$$(1-\alpha)\,(1-p) \;<\; \alpha\,(pt_H - t_L) \tag{A.3}$$

must yield the LCU regime. Hence the set of parameter quadruples that yield the LCU regime must contain an open set.

Lastly, clearly

$$(1-\alpha)\,p\,(1-p) \;<\; (1-\alpha)\,(1-p) \tag{A.4}$$

and hence for an open set of parameter quadruples we can be sure that

$$(1-\alpha)\,p\,(1-p) \;<\; \alpha\,(pt_H - t_L) \;<\; (1-\alpha)\,(1-p) \tag{A.5}$$

Therefore, we can be sure that the set of parameter quadruples that yield the NCLC regime also contains an open set. ■

Proof of Proposition 3: By assumption, the given parameter quadruple supports the NCLC regime. Hence, (11) must be satisfied. The level of individual transactions in the NC equilibrium is $\alpha t_L + 1 - \alpha$, while in the LC equilibrium it is $\alpha t_H + (1-\alpha)\,p$. Hence it suffices to show that (11) implies

$$\alpha\, t_L \;+\; 1 \;-\; \alpha \;<\; \alpha t_H \;+\; (1\;-\;\alpha)\,p \tag{A.6}$$

Consider the first inequality in (11). Divide both sides by p and rearrange to obtain

$$\frac{\alpha\, t_L}{p} \;+\; 1 \;-\; \alpha \;\leq\; \alpha\, t_H \;+\; (1\;-\;\alpha)\,p \tag{A.7}$$

which, given that $0 < p < 1$ immediately yields (A.6). ■

Lemma A.1: *Fix any arbitrary $x \in (0,1)$. Then there exist parameter quadruples that satisfy (8). This ensures that an NC equilibrium exists, and such that the individual transaction level in the associated NC equilibrium (given by $x_{NC} = \alpha t_L + 1 - \alpha$) is in fact equal to the arbitrarily given x.*

Proof: The claim is obvious by inspection of (8). Simply fix α and t_L so as to ensure that $x_{NC} = x$, as required. Then fix an arbitrary $x_H \in (x_L, 1)$, and finally pick p suitably small so that (8) is satisfied. ■

Lemma A.2: *Fix any arbitrary $x \in (0,1)$. Then there exists parameter quadruples that satisfy (10). This ensures that an LC equilibrium exists, and such that the individual transaction level in the associated LC equilibrium (given by $x_{LC} = \alpha t_H + (1-\alpha)p$) is in fact equal to the arbitrarily given x.*

Proof: The claim is obvious by inspection of (10). One way to see this is to observe that we can pick α and p arbitrarily close to 1, and ensure that $x_{LC} = x$ as required by choosing the appropriate level of t_H. Picking x_L sufficiently small is then sufficient to ensure that (10) is met. ■

Proof of Proposition 4: To avoid ambiguity we will refer to the first second, third and fourth paragraphs of the statement of Proposition 4 as A, B, C, and D respectively.

A is an immediate consequence of Lemma A.2. B and C are immediate consequences of Lemma A.1.

To prove D, proceed as follows. Let an arbitrary parameter quadruple (α, p, t_L, t_H) such that (10) is satisfied and let the individual transaction level in the associated LC equilibrium be $x_{LC} = \alpha t_H + (1-\alpha)p$.

We need to show that we can find two parameter quadruples as follows. A quadruple $(\alpha', p', t'_L, t'_H)$ such that (10) is satisfied, $p' > p$ and $x'_{LC} = \alpha' t'_H + (1-\alpha')p' < x_{LC}$, and finally a quadruple $(\alpha'', p'', t''_L, t''_H)$ such that (10) is satisfied, $p'' < p$ and $x''_{LC} = \alpha'' t''_H + (1-\alpha'')p'' > x_{LC}$.

To construct $(\alpha', p', t'_L, t'_H)$ starting from (α, p, t_L, t_H) we can increase p by a small amount $\varepsilon > 0$, while decreasing t_H by a small amount $2\varepsilon(1-\alpha)/\alpha$, so that overall individual transaction level decreases as required. If we then decrease t_L by the same amount as t_H, it is immediate that the new quadruple $(\alpha', p', t'_L, t'_H)$ must in fact satisfy (10), as required.

To construct $(\alpha'', p'', t''_L, t_H''\prime)$ starting from (α, p, t_L, t_H), we can decrease p by an an arbitrarily small amount $\varepsilon > 0$, and set t''_H arbitrarily close to 1. By inspection, for ε sufficiently small and t''_H sufficiently close to 1, the individual transaction level will increase and (10) will be satisfied, as required. ■

Proof of Proposition 5: In the low trust equilibrium the individual transaction level is $\alpha t_H + (1-\alpha)\underline{p}$. In the high trust equilibrium the individual transaction level is $\alpha t_H + (1-\alpha)\overline{p}$. The claim is then immediate from the fact that $\overline{p} > \underline{p}$. ■

Proof of Proposition 6: Since the parameter quadruple gives rise to the NCLC regime, (11) must hold. The second inequality in (11), together with the fact that $p < 1$ immediately gives that

$$p\,[\alpha\, t_H \,+\, (1\,-\,\alpha)\,p] \;<\; \alpha\, t_L \,+\, 1 \,-\, \alpha \tag{A.8}$$

Recall that $x_{LC} = \alpha t_H + (1-\alpha)p$ and $x_{NC} = \alpha t_L + 1 - \alpha$, and that the aggregate transaction levels in the LC and NC equilibria are $p\, x_{LC}$ and x_{NC} respectively. Hence (A.8) proves the claim. ■

Proof of Proposition 7: Aggregate welfare in the NC equilibrium is given by

$$W_{NC} \;=\; 2\,(\alpha\, t_L \,+\, 1 \,-\,\alpha) \tag{A.9}$$

while aggregate welfare in the LC equilibrium is given by

$$W_{LC} \;=\; 2\,[\alpha\, t_H \,+\, (1\,-\,\alpha)\,p] \;-\; (1\,-\,p)\,[\alpha\, t_L \,+\, (1\,-\,\alpha)\,p] \tag{A.10}$$

Since the parameter quadruples we are concerned with all give rise to the NCLC regime, (11) must hold.

Notice next that, provided that t_H and t_L are suitably close to each other, (11) is compatible with quadruples that have a value of p arbitrarily close to 1. By inspection of (A.9) and (A.10) in this case we must have that $W_{LC} > W_{NC}$.

It remains to show that for some parameter quadruples that satisfy (11) we can have that $W_{NC} > W_{LC}$. Notice that, again provided that t_H and t_L are suitably close to each other, (11) is compatible with quadruples that have a value of p arbitrarily close to 0. By inspection of (A.9) and (A.10) in this case we must have that $W_{NC} > W_{LC}$. ■

References

Arrow, K. J. (1972): "Gifts and Exchanges," *Philosophy and Public Affairs*, 1, 343–362.

Banfield, E. C. (1958): *The Moral Basis of a Backward Society.* New York: The Free Press.

Blume, L. (2004): "Stigma and Social Control," Cornell University, *mimeo.*

Butler, J., P. Giuliano, and L. Guiso (2009): "The Right Amount of Trust," Working Paper 15344, National Bureau of Economic Research, http://www.nber.org/papers/w15344.

Dixit, A. K. (2003): "Trade Expansion and Contract Enforcement," *Journal of Political Economy*, 111, 1293–1317.

Duffie, D., and Y. Sun (2007): "Existence of Independent Random Matching," *Annals of Applied Probability*, 17, 386–419.

Feddersen, T., and A. Sandroni (2006): "The Calculus of Ethical Voting," *International Journal of Game Theory*, 35, 1–25.

Fukuyama, F. (1995): *Trust: The Social Virtues and the Creation of Prosperity*. New York: The Free Press.

Glaeser, E. L., D. I. Laibson, J. A. Scheinkman, and C. L. Soutter (2000): "Measuring Trust," *The Quarterly Journal of Economics*, 115(3), 811–846.

Guiso, L., P. Sapienza, and L. Zingales (2004): "The Role of Social Capital in Fiancial Development," *The American Economic Review*, 94, 526–556.

Horst, U., and J. A. Scheinkman (2006): "Equilibria in Systems of Social Interactions," *Journal of Economic Theory*, 130, 44–77.

Irlenbusch, B. (2006): "Are Non-Binding Contracts Really Not Worth the Paper?," *Managerial and Decision Economics*, 27, 21–40.

Karni, E., and Z. Safra (2002): "Individual Sense of Justice: A Utility Representation," *Econometrica*, 70, 263–284.

Kim, J.-I., and L. J. Lau (1994): "The Sources of Economic Growth of the East Asian Newly Industrialized Countries," *The Journal of the Japanese and International Economies*, 8, 235–271.

Knack, S., and P. Keefer (1997): "Does Social Capital Have an Economic Payoff? A Cross-Country Investigation," *The Quarterly Journal of Economics*, 112(4), 1251–1288.

La Porta, R., F. Lopez-de-Silanes, A. Shleifer, and R. W. Vishny (1997): "Trust in Large Organizations," *The American Economic Review*, 87(2), 333–338.

Mailath, G., and L. Samuelson (2006): *Repeated Games and Reputations: Long Run Relationships*. Oxford: Oxford University Press.

Phillips, P. C. B. (1989): "Partially Identified Econometric Models," *Econometric Theory*, 2, 181–240.

Putnam, R. D. (1993): *Making Democarcy Work: Civic Traditions in Modern Italy*. Princeton, NJ: Princeton University Press.

Samuelson, L. (1997): *Evolutionary Games and Equilibrium Selection*. Cambridge, MA: MIT Press.

Sapienza, P., A. Toldra, and L. Zingales (2007): "Understanding Trust," *SSRN eLibrary*, http://ssrn.com/paper=1014335.

Smith, A. (1759): *The Theory of Moral Sentiments*. Printed in London for A. Millar in the Strand and A. Kincaid and J. Bell in Edinburgh, Second Edition (1761), available at http://books.google.com/books?id=xVkOAAAAQAAJ&printsec=titlepage.

——— (1776): *The Wealth of Nations*. Dublin: Printed by N. Kelly for P. Wogan, Gilbert and Hodges, W. Porter, J. Moore and B.Dornin. Sixth Edition (1801), available at http://books.google.com/books?id=9RsUAAAAYAAJ&printsec=titlepage.

Smith, V. L. (1998): "The Two Faces of Adam Smith," *Southern Economic Journal*, 65(1), 2–19.

Sobel, J. (2002): "Can We Trust Social Capital?," *Journal of Economic Literature*, 40, 139–154.

Solow, R. M. (1995): "But Veriify," *The New Republic, September 11*, pp. 36–35, Review of Fukuyama's *Trust: The Social Virtues and the Creation of Prosperity.*

Tabellini, G. (2008): "The Scope of Cooperation: Values and Incentives," *The Quarterly Journal of Economics*, 123, 905–950.

Weibull, J. W. (1995): *Evolutionary Game Theory*. Cambridge, MA: MIT Press.

Young, A. (1994): "Lessons From The East Asian NICS: A Contrarian View," *The European Economic Review*, 38, 964–973.