

# Cultivating Trust: Norms, Institutions and the Implications of Scale\*

Chris Bidner<sup>†</sup>

Patrick Francois<sup>‡</sup>

May 13, 2010

## Abstract

We study the co-evolution of norms and institutions in order to better understand the conditions under which potential gains from new trading opportunities are realized. New trading opportunities are particularly vulnerable to opportunistic behavior and therefore tend to provide fertile ground for cheating. Cheating discourages production, raising equilibrium prices and therefore the return to cheating, thereby encouraging further cheating. However, such conditions also provide institutional designers with relatively high incentives to improve institutions. We show how an escape from the shadow of opportunism requires that institutional improvements out-pace the deterioration of norms. A key prediction from the model emerges: larger economies are more likely to evolve to steady states with strong honesty norms, even though larger economies need not have better institutions. This prediction is tested using a cross section of countries; population size is found to have a significant positive relationship with a measure of trust, even when controlling for standard determinants of trust and institutional quality.

**Keywords:** Trust, Institutions, Norms, Population Size

**JEL Codes:** Z13, O10

---

\*This is a substantially revised version of a paper previously entitled: “Norms and the Dynamics of Institution Formation”. We thank the Editor and two anonymous referees for very helpful suggestions, as well as seminar participants at Brown, UBC, Simon Fraser, Calgary, Dalhousie, Laval, Northwestern, West-Coast Macro Conference, Bristol, LSE, AEA meetings Chicago, and the Canadian Institute For Advanced Research. We are especially grateful to Daron Acemoglu, Siwan Anderson, Tim Besley, Maitreesh Ghatak, Paul Grout, Peter Howitt, Ashok Kotwal, Kiminori Matsuyama, Nathan Nunn, Lars Osberg, Andrea Prat and Ken Shepsle for helpful discussions.

<sup>†</sup>University of New South Wales, School of Economics. [c.bidner@unsw.edu.au](mailto:c.bidner@unsw.edu.au)

<sup>‡</sup>University of British Columbia, Dept. of Economics, CEPR and CIFAR. [francois@interchange.ubc.ca](mailto:francois@interchange.ubc.ca)

# 1 Introduction

A large and growing body of empirical analysis is making the case that trust affects economic outcomes.<sup>1</sup> Rather than being an innate behavioral disposition, trust instead seems to arise where people tend to be trustworthy.<sup>2</sup> But why are people more trustworthy in some places than others?

A set of explanations, recently exemplified by [Bisin \*et al.\* \(2004\)](#) and [Tabellini \(2008\)](#), emphasizes behavioral sources of trustworthiness, where regional differences are thought to be the products of deep cultural distinctions across locales. For historical reasons, it is argued, some regions or countries succeeded in inculcating higher levels of civicness, altruism, generalized morality, or social capital than others and thus have higher trust today.<sup>3</sup>

Another set of explanations, prominent examples of which are [Dixit \(2004\)](#) and [Greif \(1994\)](#), hinge on the quality of institutions, and the consequent effects these have on incentives. The argument these authors make is that good institutions create trustworthy people by detecting opportunistic acts, fairly administering well defined laws, and punishing those who break them.<sup>4</sup>

Even accepting a clear distinction between these two sources of trustworthiness, it is unlikely that they would develop independently through time. As far back as Aristotle, it has been argued that repeated exposure to good institutions that make even the opportunistic act decently, may

---

<sup>1</sup>The main source of evidence for the analysis of trust is the World Values Survey (WVS) “generalized trust” question that we will analyze in some detail further on. [Tabellini \(2010b\)](#) discusses a large part of the recent literature that has established the seemingly important influence that trust and other cultural factors have on economic outcomes. A survey of some of the related work is also provided by [Guiso \*et al.\* \(2006\)](#). [Algan and Cahuc \(2009\)](#) is a noteworthy recent contribution to this literature in that it is the first paper to identify effects of trust on economic outcomes while controlling for country fixed effects.

<sup>2</sup>A number of authors have explored the meaning of affirmative answers to the generalized trust question. [Fehr \*et al.\* \(2003\)](#), using the German Socio-Economic Panel, show that positive answers to this question are correlated with trusting (sender) behavior in the [Berg \*et al.\* \(1995\)](#) game. [Sapienza \*et al.\* \(2007\)](#) look at how answers to the trust question correspond with play in a modified version of the “Trust” game where senders are also asked directly for information on their beliefs regarding receiver behavior. They find trust to be highly correlated with a respondents’ beliefs about the trustworthiness of a paired playing partner in the game. However [Glaeser \*et al.\* \(2000\)](#) found positive answers to be a better predictor of trustworthiness (receiver behavior in the trust game) though [Sapienza \*et al.\* \(2007\)](#) provide a reconciliation of these findings. For our purposes here, resolving these differences is not important. As will be seen, both interpretations are consistent with the theory we present.

<sup>3</sup>This is the approach in [Bisin \*et al.\* \(2004\)](#), for example, and in [Tabellini \(2008\)](#). It is also consistent with the evidence in [Nunn and Wantchekon \(2009\)](#), and the experimental work of [Barr \(2003\)](#) and [Fehr \(2008\)](#).

<sup>4</sup>The relevant institutions are those that underpin economic transactions. According to [Dixit \(2008\)](#), these comprise three categories: protection of property rights against theft, enforcement of voluntary contracts, and provision of the physical and regulatory infrastructure to facilitate economic activity and the functioning of the first two categories. The literature has explored numerous aspects of these, for example the causes of transitions from informal to formal institutions ([Li \(2003\)](#), [Leukeart \(2005\)](#), [Dixit \(2004\)](#) (chs. 2 and 3), [Dhillon and Rigolini \(2006\)](#) and [Greif \(2002\)](#)), and the effects of variations in institutional quality on economic outcomes both with political ([Acemoglu \*et al.\* \(2001\)](#)) and [Acemoglu \*et al.\* \(2002\)](#)) and legal ([La Porta \*et al.\* \(1998\)](#)) institutions.

eventually breed inherent trustworthiness, in either them or their descendants.<sup>5</sup> Also, as [North \(1990\)](#) has argued, if it tends to be the case that decent, fair and trustworthy individuals make it easier for poor quality institutions (i.e., where there is minimal oversight, few resources devoted to detection, and small chance of punishment) to work, then what we have called the behavioral component of trustworthiness above, may actually help the institutional component to work.<sup>6</sup>

In this paper we develop the first theoretical framework that explicitly models the evolution of behavioral norms and institutions. Since we are primarily interested in understanding how good institutions and the norms that support them arise, we explicitly allow for complex dynamic interactions between behavioral norms and institutions, like those discussed above, to occur.<sup>7</sup> The types of institutions that are key, and are our focus here, are those that facilitate trade. The legal institutions that provide and enforce contractual commitments and allow timely settlement of grievances, and the political institutions that recognize and protect private property from the state and from other agents.

The environment we study is a new trading situation where, initially, trader vulnerability to moral hazard is severe. We assume that the institutions needed to effectively enable this trade are not born ready-made. It takes time to find the structures needed to enable, monitor, enforce, and punish actions that are detrimental to trade. But, even before institutions are functional, if enough people are inherently honest, trade may still occur. This is because traders will risk the losses incurred when meeting opportunists if the chances of meeting such people are low enough. Some degree of pre-existing honesty norms in a population can thus allow trade to get started. However, once started, the value of such trade itself provides the spur for institutional improvement. We show that when the value of trade is high, incentives for institutional improvement are great, and institutions improve quickly. However, when low, institutions improve more slowly.

---

<sup>5</sup>According to Aristotle's *Nicomachean Ethics* "Lawgivers make the citizen good by inculcating habits in them, and this is the aim of every lawgiver; if he does not succeed in doing that, his legislation is a failure." See [Leukeart \(2005\)](#) (p.2) for this quote and others suggesting a similar process. This is also the position taken by [Tabellini \(2010a\)](#) where it is posited that a long term and persistent legacy of good institutions in previous centuries is high levels of civiness, and trust, today.

<sup>6</sup>[North \(1990\)](#) argues that transaction costs - the costs of policing the trading arrangements - are lower when individuals share beliefs, norms, ideology and a common sense of fairness, which can limit their opportunism. This will imply that lower levels of institutional quality will suffice in environments where individuals share such values. Legal scholars have also studied the role of norms in helping the law to work – see [Cooter \(2000\)](#) and [McAdams and Rasmusen \(2007\)](#).

<sup>7</sup>[North \(2005\)](#) also draws a distinction between a type of internal enforcement that makes individuals follow prescribed behavior, and more formal enforcement that arises from explicit punishments. The focus of his analysis is primarily historical and is not directly concerned with the dynamic interaction of these two features, which is our concern here. [Greif \(2006\)](#) is concerned with the process of institution formation in European history and [Aghion et al. \(2004\)](#) look at the formation of political institutions. The set of concerns there, which are mainly distributional and relate to relative power of constituents, are distinct from the concerns regarding institutions to facilitate trade which are our focus here.

If institutions improve enough, then even opportunists find it incentive compatible to trade honestly. This is the point at which institutions become, in our terminology, “functional”. But functional institutions are not guaranteed. Along the transition path, before institutions are functional, it pays to be dishonest, so pre-existing honesty norms are continuously eroded. A ‘race’ between institutions and honesty norms ensues. If institutions do not develop quickly enough, honesty norms eventually erode away, leading to a breakdown in trade. This in turn stifles further institutional improvement. If fast enough, however, institutions become functional and then honesty norms are reinforced since honesty becomes incentive compatible.

Two types of long-run outcome are possible in our model: (1) functional institutions/high trust, where there is extensive trade; (2) dysfunctional institutions/low trust, where trading opportunities are stifled by opportunistic behavior.

Many contemporary countries fit the good outcome in our model, and also have high quality institutions. But the functionality of institutions that we refer to here is not equivalent to their quality. Low quality institutions can still be functional, i.e., sufficient to deter the self-interested from acting dishonestly, and consequently support high levels of trust and trade. A contemporary example would be China, a country which scores poorly on measures of individual property rights protection, rule of law, and intellectual property rights yet still has high levels of trust and extensive trade. The model makes clear the non-trivial relationship between the functionality of institutions, their quality, and the existence of behavioral types who are inherently honest.<sup>8</sup>

Our model predicts a number of country level characteristics that should favor the functional institutions/high trust outcome over the dysfunctional institutions/low trust one. Most of these are the somewhat standard correlations with economic development that have already been widely reported in the previous literature. However, a unique testable implication of the theory follows from the fact that in larger economies the returns to creating functional institutions are likely to be greater. In common with the argument made by [Demsetz \(1967\)](#), our model has the feature that the fixed costs inherent in the creation and running of institutions lead to a scale economy favoring their development in larger economies.<sup>9</sup> Through our theory, this leads to a positive relationship between the number of people in a country and its level of trust, controlling for all of the other

---

<sup>8</sup>Our work thus somewhat parallels a recent set of contributions concerned with endogenizing norms and government policies: [Aghion \*et al.\* \(2008\)](#), [Aghion \*et al.\* \(2009\)](#), [Carlin \*et al.\* \(2007\)](#), and [Alesina and Angeletos \(2005\)](#) who are concerned with labor laws, government regulation, and redistributive policy, respectively, and how these are simultaneously determined with societal norms or beliefs. A marked distinction is our focus on institutions, and our explicit characterization of the dynamics of these processes. Also, [Hauk and Saez-Marti \(2002\)](#), who provide an endogenous cultural underpinning for corruption, [Francois and Zabojnik \(2005\)](#), who model endogenous Social Capital, and [Tabellini \(2008\)](#), who models Banfield’s limited versus generalized morality, are papers that provide an endogenous treatment of norms as we do here. The main distinction between ours and all of these is our focus on the simultaneous endogenous development of institutions.

<sup>9</sup>This aspect was also key to the identification employed recently by [Mulligan and Shleifer \(2005\)](#) to test their model of regulation.

factors that have previously been found to be determinants of trust, and even when controlling for the quality of institutions in the country.<sup>10</sup>

Though our model's predictions about the relationship between scale and trust are straightforward, the way that institutions are affected by scale is more subtle. As already mentioned, larger countries have greater incentives to invest in institutional improvement. This tends to promote a positive relationship between scale and institutional quality. However, the faster pace of institutional improvement limits the extent to which behavioral norms can erode, thereby allowing institutions to become functional at lower quality levels. Since institutions are assumed to be costly to improve, their quality rises only to the point where they become functional. This tends to promote a negative relationship between scale and institutional quality. The upshot is that, larger countries should have higher levels of trust than smaller ones, but institutional quality need bear no particular relationship with either country size, nor with trust levels. This conclusion is established in greater detail in the dynamic model presented in section 2, and is tested using data from the World Values Survey in Section 3. Section 4 concludes.

## 2 The Model

### 2.1 The Trading Environment

There is an infinitely elastic supply of profit-maximizing, risk neutral entrepreneurs that either enter into an exogenously given trading opportunity - a technology - or consume their reservation utility, normalized to zero. If entering into the trading opportunity, the entrepreneur's output is sold on a competitive market. Total demand for output as a function of its price,  $p$ , is given by the demand function  $V(p) \equiv N \cdot D(p)$ , where  $N$  is population size, and  $D(p)$  denotes the amount demanded by a single individual at price  $p$ . For any  $p \geq 0$ ,  $D(p) \geq 0$  and  $D'(p) < 0$ , with  $\lim_{p \rightarrow \infty} D(p) = 0$ . Production requires the investment of a fixed, sunk cost  $C$ , to produce one unit of output. Production is one shot, and time is continuous. In summary, production yields something of value,  $p$  (endogenous) at a cost,  $C$  (exogenous and sunk).

#### 2.1.1 One-Sided Moral Hazard

Entrepreneurial production requires a trading partner. Working with the trading partner is essential for output to be produced but the trading partner can steal all of the output and sell it on the

---

<sup>10</sup>Strokova (2010) has recently compiled a set of indexes measuring various dimensions of property rights that we will utilize as our measures of country level institutional quality. These are discussed at length in the empirical part of the paper.

market at price  $p$ .<sup>11</sup> For simplicity, the costs of filling a trading partner's position are set to zero, which is also the opportunity cost of trading partners.

### 2.1.2 Institutions

Institutions detect and punish cheating, in particular, a trading partner's stealing of output. Institutional quality at time  $t$  is the probability that such theft is detected;  $I(t) \in [0, 1]$ . There are no false detections.<sup>12</sup> Institutions range from those that are perfect and therefore always detect cheating,  $I(t) = 1$ , to those that are completely ineffective and never do;  $I(t) = 0$ . Detected cheaters are punished an amount  $F$ , which we treat as a parameter of the institutional technology. A proportion  $\alpha \in [0, 1]$  of the output is returned to the entrepreneur, with the remaining proportion being consumed by the cheater.<sup>13</sup>

### 2.1.3 First-Best Output

Since the only cost borne in production is the cost  $C$ , the first best level of entry is that which induces  $p = C$ . If there are  $n$  producers, then the equilibrium price satisfies  $n = N \cdot D(p(n))$ . As such, the first best level of entry,  $n^*$ , is given by

$$n^* = N \cdot D(C).$$

### 2.1.4 Rational Agents and Norm Followers

A proportion  $r \in [0, 1]$  of the set of potential trading partners are rational agents. These agents always act to maximize their payoffs, taking into account expected costs and benefits. They will steal output if the cost benefit ratio is low enough. The remaining proportion,  $1 - r$ , are norm followers. There are two possible norms. These individuals can either follow a norm of honesty, which dictates that they always trade in good faith and do not steal output, or they can follow a cheating norm and always steal output. At time  $t$ , denote the fraction of norm followers following an honesty norm  $\beta(t) \in [0, 1]$ , with the remaining  $1 - \beta(t)$  being cheaters. In summary, if a proportion  $\sigma$  of rational agents cheat, then the proportion of the population that do not cheat is given

---

<sup>11</sup>This theft of output will be referred to as "cheating" in the paper from hereon. This theft is a type of breach of contract, which is probably the more common form of "cheating" in reality. Unlike theft, it is rare that all of the value of output can be obtained by simple breach of contract in reality. However, all that matters for the results here is that the value of such breach will be positively related to the value of output.

<sup>12</sup>Here institutional improvement simply improves the rate of detection when cheating, since there are no errors arising from punishing non-cheaters. A simple extension would include false guilty accusations, in which case institutional improvement would also amount to reducing these. This would not change any of the model's substantive results.

<sup>13</sup>The cost of operating the institution is not modeled here. Introducing this would not greatly impinge on the set-up here, except in requiring some means of taxing and financing the running. It would, however, not affect the comparative statics which are the main focus. Also, it has already been analyzed elsewhere, see Gradstein (2004), and is hence omitted.

by

$$\delta \equiv r \cdot (1 - \sigma) + (1 - r) \cdot \beta. \quad (1)$$

While  $r$  is exogenous,  $\beta$  (and therefore  $\delta$ ) is endogenously determined as described below. An individual's type - honest, cheater or rational - is known to the individual but not observable to others.

## 2.2 Instantaneous Equilibrium

Given the state of norms and institutions,  $(I(t), \beta(t))$ , an instantaneous equilibrium is a mutually consistent description of i) the entry decision of entrepreneurs, and ii) the action taken by rational types. Fixing the action of rational types fixes  $\delta$ , and given  $\delta$  profits are

$$\pi = p \cdot [\delta + (1 - \delta) \cdot \alpha \cdot I] - C. \quad (2)$$

Free entry of entrepreneurs ensures that the equilibrium price induces zero profits:

$$p = \frac{C}{\delta + (1 - \delta) \cdot \alpha \cdot I}. \quad (3)$$

Given an output price, the expected payoff from cheating is

$$(1 - I) \cdot p + I \cdot [(1 - \alpha) \cdot p - F], \quad (4)$$

implying that rational agents cheat if the output price is sufficiently high. Using (3) in (4) tells us the net benefit from cheating,  $u^c(\delta, I)$ , is

$$u^c(\delta, I) = \frac{1 - \alpha \cdot I}{\delta \cdot (1 - \alpha \cdot I) + \alpha \cdot I} \cdot C - I \cdot F. \quad (5)$$

By rearranging this and recalling that rational agents only cheat if  $u^c(\delta, I) > 0$ , we have that rational agents decide to cheat if

$$\delta < \hat{\delta}(I) \equiv \frac{C}{I \cdot F} - \frac{\alpha \cdot I}{1 - \alpha \cdot I}, \quad (6)$$

and do not cheat otherwise. Note that  $\hat{\delta}(I)$  is a strictly decreasing function of  $I$ , which approaches infinity as  $I$  approaches zero. It will be convenient to define  $I_1$  as the value that satisfies  $\hat{\delta}(I_1) = 1$  and  $I_2$  as the value that satisfies  $\hat{\delta}(I_2) = 0$ . It is straightforward to verify that  $I_1 \in (0, \alpha^{-1})$  and  $I_2 \in (I_1, \alpha^{-1})$ . The  $\hat{\delta}(I)$  curve,  $I_1$ , and  $I_2$  are depicted in the left panel of Figure 1.

An inherent complementarity in the decision to cheat among rational types means that instantaneous equilibria will generally not be unique. This complementarity arises because the more rational types that cheat, the lower is profitability and market entry, which raises the output price, thereby increasing the payoff to cheating (and vice versa). Thus, the incentives for rational agents to (not) cheat are heightened when more rational types are (not) cheating.

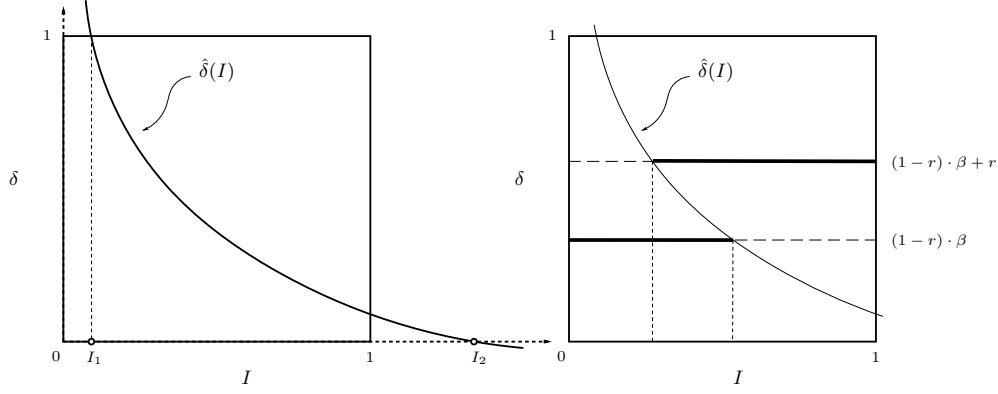


Figure 1: Instantaneous Equilibrium

The set of pure-strategy instantaneous equilibria is depicted in the right panel of Figure 1, where the dark horizontal lines represent equilibrium values of  $\delta$  for any given state of the economy  $(I, \beta)$ . We see that if  $I$  is relatively low, then rational types find it optimal to cheat even if they are the only rational agent to do so. On the other hand, if  $I$  is relatively high, then rational types find it optimal to act honestly even if they are the only rational agent to do so. For intermediate values of  $I$ , rational types will cheat if all others do, and will not cheat if all others do not. This is formalized in the following Proposition.

**Proposition 1.** *Given  $(I(t), \beta(t))$ , we either have:*

1. *Dysfunctional Institutions.* For  $(1-r) \cdot \beta + r < \hat{\delta}(I)$  the unique pure-strategy equilibrium has rational partners cheating.
2. *Functional Institutions.* For  $\hat{\delta}(I) < (1-r) \cdot \beta$  the unique pure strategy equilibrium has rational partners not cheating.
3. *An Ambiguous Institutional Environment.* For  $(1-r) \cdot \beta \leq \hat{\delta}(I) \leq (1-r) \cdot \beta + r$  there are two pure strategy equilibria: one in which rational partners cheat (dysfunctional institutions) and one in which they do not (functional institutions).

The corresponding equilibrium price is given by (3) where  $\delta = (1-r) \cdot \beta$  when institutions are dysfunctional and  $\delta = (1-r) \cdot \beta + r$  when institutions are functional.

*Proof.* See appendix. □

From the perspective of this static environment, we expect a positive relationship between institutional quality and trust. This arises for the simple reason that higher institutional quality makes it more likely that rational types choose to refrain from cheating. This predicted relationship *does not* remain once we consider the evolution of behavioral types.



### 2.2.1 The Ambiguous Region

In the ambiguous region, pure strategy equilibria either have functioning institutions and no cheating, or the entirely opposite - dysfunctional institutions and cheating. None of the qualitative results to follow depend on how this multiplicity is resolved (i.e. whether it is assumed that the cheating or the honest equilibrium is played). Intuitively, resolving this multiplicity allows us to pin down, for each  $\beta$ , a critical institutional quality beyond which rational players choose to act honestly. If we had instead assumed that the honest equilibrium is selected, then this critical institutional quality would simply be lowered (at each  $\beta$ ). Given the dynamics of this system described below, the cheating equilibrium will seem like the more focal equilibrium. This is because if we think of a situation in which institutions start out with low quality but improve over time, then initially cheating is the unique behavior consistent with instantaneous equilibrium. As conditions (costs and benefits of cheating) evolve over time, it may become the case that honest behavior also becomes consistent with instantaneous equilibrium. We say that cheating is the focal behavior in the face of this multiplicity in the sense that it is the “status quo” behavior.

With this focus, we have

$$\delta(I, \beta, r) = \begin{cases} (1-r) \cdot \beta & \text{if } (1-r) \cdot \beta \leq \hat{\delta}(I) \\ (1-r) \cdot \beta + r & \text{if } (1-r) \cdot \beta > \hat{\delta}(I) \end{cases} \quad (7)$$

As will be elaborated on below, notice that it is not possible to have  $\delta \in (\hat{\delta}(I), \hat{\delta}(I) + r]$ .<sup>14</sup> This is due to the fact that a positive measure of agents change their behavior discretely at certain critical points, and not due to the way in which we choose to deal with multiple equilibria.

## 2.3 Dynamics

At any time  $t$ , the economy is an instantaneous equilibrium as summarized by  $(I(t), \delta(t))$ . In this section we describe how the economy evolves over time by considering the impact of  $(I(t), \delta(t))$  on the dynamics of norms and institutions.

### 2.3.1 Dynamics of Norms

The most simple treatment of behavioral dynamics supposes that when violating norms pays well, individuals will be tempted to do so, leading to the erosion, and eventual breakdown of the norm in question. In contrast, when following a norm is profitable, the behavior is copied and reinforced. Here we adopt the [Bisin and Verdier \(2001\)](#) model of cultural selection which parsimoniously captures such a treatment. This model has become a widely used way of treating cultural

---

<sup>14</sup>This is because if  $(1-r) \cdot \beta \leq \hat{\delta}(I)$ , then  $\delta = (1-r) \cdot \beta$ , which implies  $\delta \leq \hat{\delta}(I)$ . Alternatively, if  $(1-r) \cdot \beta > \hat{\delta}(I)$ , then  $\delta = (1-r) \cdot \beta + r$ , which implies  $\delta > \hat{\delta}(I) + r$ . Thus, there are no values of  $(1-r) \cdot \beta$  such that  $\delta \in (\hat{\delta}(I), \hat{\delta}(I) + r]$ .

evolution in economics and has received a number of different applications: preferences for social status (Bisin and Verdier (1998)), corruption (Hauk and Saez-Marti (2002)), hold up problems (Olcina and Penarrubia (2004)), development and social capital (Francois and Zabojsnik (2005)), inter-generational altruism (Jellal and Wolff (2002)), labor market discrimination (Saez-Marti and Zenou (2007)), globalization and cultural identities (Olivier *et al.* (2008)) and work-ethics (Bisin and Verdier (2005)); see Bisin and Verdier (2008) for a full survey of the literature using this approach.<sup>15</sup>

At time  $t$ , each member of the current population socializes a single member of the successive population that will follow it at time  $t + \Delta$ . Denote the socializing agent a ‘parent’ and the agent socialized as an ‘off-spring’. In order to keep the population of norm followers constant at  $1 - r$ , we assume rational parents always produce rational offspring. For norm followers, there are two ways that socialization occurs in this model. The first, called direct socialization, happens when the parent socializes its offspring to be the same type as itself. The probability of direct socialization is denoted  $d^i$ , where  $i = h$  or  $c$ , for an honest or cheater type respectively. A parent does not have the capacity to directly socialize an off-spring in a way that is counter to type, but with probability  $1 - d^i$  the off-spring does not receive direct socialization from the parent. When off-spring are not directly socialized, then socialization is determined by some other individual in the population. The individual of influence is randomly drawn from the remaining population members. This has the appealing feature that, in a society where most people are type  $i$ , a person not directly socialized by a parent, will have a high chance of being socialized to be type  $i$  by someone else.<sup>16</sup> We are thus able to define the probability that a parent of type  $i$  will have an offspring of type  $j$  by the function  $q^{ij}$ . With proportion  $\delta(t)$  honest types at time  $t$ , this yields the following four transition probabilities:

$$q^{hh} = d^h + (1 - d^h) \cdot \delta(t) \quad (8)$$

$$q^{hc} = (1 - d^h) \cdot (1 - \delta(t)) \quad (9)$$

$$q^{ch} = (1 - d^c) \cdot \delta(t) \quad (10)$$

$$q^{cc} = d^c + (1 - d^c) \cdot (1 - \delta(t)). \quad (11)$$

---

<sup>15</sup>The reduced form of behavioral dynamics would be very similar if, instead of a model of cultural evolution, a model where parents make direct choices about childrens’ values were to be used. Such a model has been used by Lindbeck and Nyberg (2006) to analyze the effect of welfare provision on work norms.

<sup>16</sup>An additional layer of selection between rational and non-rational agents could be analyzed provided some costs to rationality were also modeled. This is not attempted here, though the evolutionary stability of pay-off maximizing behavior has been the subject of a substantial literature. For a recent contribution see Heifetz *et al.* (2007). Relatedly, Bénabou and Tirole (2006) analyze the role of incentives in motivating behavior when pro-social considerations are present. That analysis of individual decision making is much more sophisticated than here, as it is focused on the individual. Here, instead, the simpler treatment of individuals facilitates our focus on institution formation.

To calculate  $\dot{\delta}$ , consider  $\delta(t)$  and the transition to  $\delta(t + \Delta)$  in time interval  $\Delta$ . Thus we have:

$$\delta(t + \Delta) = \delta(t) \cdot q^{hh} \cdot \Delta + (1 - \delta(t)) \cdot q^{ch} \cdot \Delta.$$

Substituting from above, rearranging, and letting  $\Delta \rightarrow 0$  yields:

$$\dot{\delta}(t) = \delta(t) \cdot (1 - \delta(t)) \cdot (d^h - d^c). \quad (12)$$

Hence, the two key variables driving the dynamics are the probability that an honest norm follower directly socializes an off-spring to be honest,  $d^h$ , and the same probability for a cheater,  $d^c$ . The most reasonable assumption to make here is that consistent with a replicator dynamic: the higher relative returns to following an honesty (cheating) norm, the higher the probability of direct socialization to honesty (cheating). From (5), the net payoff associated with being honest is  $-u^c(\delta, I)$ , which motivates the following assumption.

**Assumption 1.** (*Replicator*): Given  $(\delta, I)$ , we have

$$d^h - d^c = \phi(-u^c(\delta, I)), \quad (13)$$

where  $\phi(0) = 0$ ,  $\phi'(\cdot) > 0$ , and  $u^c(\delta, I)$  is given by (5).

That is, under this assumption we have

$$\dot{\delta}(I, \delta) = \delta \cdot (1 - \delta) \cdot \phi(-u^c(\delta, I)). \quad (14)$$

One intuitive implication of this is that the sign of  $\dot{\delta}$  is the same as the sign of  $\delta - \hat{\delta}(I)$ . This, together with the boundary restriction that  $\delta \in [0, 1]$ , implies

$$\dot{\delta}(t) \begin{cases} > 0 & \text{if } \delta(t) > \hat{\delta}(I(t)) \text{ and } \delta(t) < 1 \\ < 0 & \text{if } \delta(t) < \hat{\delta}(I(t)) \text{ and } \delta(t) > 0 \\ = 0 & \text{otherwise} \end{cases} \quad (15)$$

The phase diagram associated with these dynamics is displayed in Figure 2. Honesty norms deteriorate for points below  $\hat{\delta}(I)$  and improve for points above  $\hat{\delta}(I) + r$ . As mentioned above, it is impossible to have  $\delta \in (\hat{\delta}(I), \hat{\delta}(I) + r)$  in an instantaneous equilibrium. If the economy did happen to find itself in such a position, it would *immediately* jump to  $\hat{\delta}(I) + r$  because the rational types would all immediately find that the net payoff to cheating is negative.

## 2.4 Dynamics of Institutions

Institutions change via the efforts of a risk neutral institutional designer. The most natural interpretation of the institutional designer is a political actor who is rewarded for i) direct improvements in institutional quality, as evidenced by improved detection, prosecution and punishment

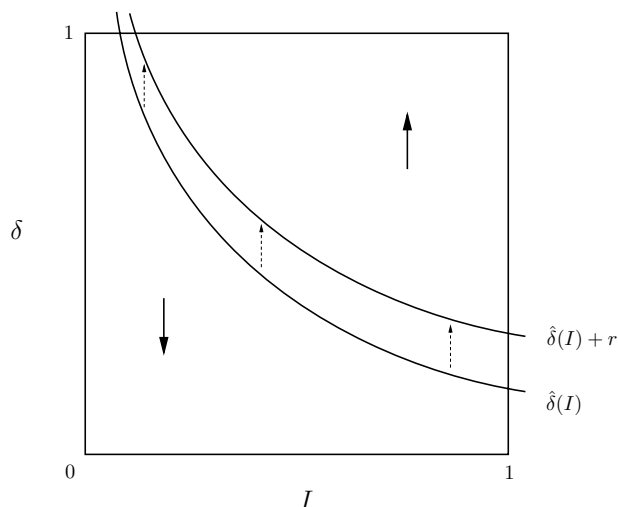


Figure 2: Dynamics of  $\delta$

of traders who violate agreements, and ii) improvements in consumer surplus that arise from such institutional improvements (free entry implies producer surplus is always zero). The rewards may take the form of increased political donations, a perception of competence on the part of constituents that increases re-election probability, some other benevolent career concerns, or even direct transfers to the designer. The designer's myopia may be literal, in the case of political agents with short-lived mandates and re-election concerns, or reflect difficulties in appropriability due to incomplete markets. For instance, as would occur when future beneficiaries of institutional development cannot be identified and taxed to reward the institution building efforts incurred today.<sup>17</sup>

This simple way of treating institutional change is a first pass at the problem which clearly abstracts from many important elements of reality. A richer political economy structure would allow for concerns regarding distributional impacts to play a role. As argued earlier, for the institutions of enforcement (except where the cheaters themselves are powerful enough to thwart improved oversight) the issue is not so much one of losers versus gainers but one of providing incentives for the sorts of institutional improvement that all participants agree are needed. The general forces that are highlighted here would still be at play in a richer political economic environment, but other distributional ones which are more likely to reflect local political considerations, would also emerge. The modeling strategy here is to abstract from these richer - but likely much more context specific - political considerations in order to explore more general effects.

<sup>17</sup>It is assumed that agents who would actively benefit from continued dysfunctional institutions - for instance those whose aim is to cheat good faith trading partners - have no capacity to influence the institution designer. Hoff and Stiglitz (2004) analyze a case when this can occur to explain the lack of institutional improvement in post-Soviet Russia. Agents whose aim it is to strip corporate assets, or to defraud share-holders for personal gain, will oppose (or buy off) institutional designers who aim to improve corporate governance and accountability. Since the costs of maintaining institutional quality are also not modeled in that paper, institutional dynamics are simple and monotonic.

The institutional designer decides how much effort to devote to improving institutional quality. An investment of  $e$  improves institutional quality according to a technology given by  $\dot{I}(e) = Z(e)$ , where  $Z$  is a strictly concave function with  $Z(0) = 0$  and  $Z'(0) = z \geq 0$ . An investment of  $e$  costs the designer  $c(e)$ , where for simplicity we assume  $c(e) = e$ . Given non-negative weights  $\gamma_1 \geq 0$  and  $\gamma_2 > 0$ , the institutional designer is rewarded according to

$$R = \gamma_1 \cdot \dot{I} + \gamma_2 \cdot \dot{\Psi},$$

where  $\Psi$  is consumer surplus:

$$\Psi \equiv N \cdot \int_p^\infty D(s) ds.$$

The fact that consumer surplus at time  $t$  depends on the price prevailing in the instantaneous equilibrium at time  $t$ , the marginal return to the designer's effort will depend not only on the technology available to them, but also on the state of the economy. In particular, the designer's return function can be written as a function of their effort and the state of the economy as follows (details in appendix):

$$R(e) = M(I, \delta) \cdot Z(e),$$

where  $M(I, \delta)$  is a term that shifts the designer's marginal return, defined as:

$$M(I, \delta) \equiv \gamma_1 + \gamma_2 \cdot N \cdot \left\{ D \left( \frac{C}{\delta + (1 - \delta) \cdot \alpha \cdot I} \right) \cdot \frac{C \cdot (1 - \delta) \cdot \alpha}{[\delta + (1 - \delta) \cdot \alpha \cdot I]^2} \right\}. \quad (16)$$

This expression has two key features. First, it is increasing in  $N$  - a reflection of the economies of scale inherent in institutional improvement. Second, it is decreasing in  $I$  if and only if the price elasticity of demand (in absolute value) is less than two. This is less obvious, and is proved in the appendix, but one intuition is as follows: when institutional quality is relatively low, the designer's return tends to be high to the extent that institutional improvements have a relatively large effect on the price, but tends to be low to the extent that price decreases have a relatively small effect on consumer surplus (since the price tends to be relatively high). If demand were highly elastic then the latter effect would outweigh the former and we would have a situation in which the marginal return to institutional improvement is *increasing* in the level of institutional quality. We consider this to be more of a technical curiosity<sup>18</sup> than a relevant feature of the forces we are highlighting, so we make the following assumption.

**Assumption 2.** (*Demand*): *The demand function is unit elastic:*

$$D(p) = s \cdot p^{-1},$$

for some scaling parameter  $s > 0$ .

---

<sup>18</sup>For instance we could have a development trap arising from the fact that consumer surplus is relatively insensitive to price reductions at high price levels, which lowers the designer's incentive to engage in institutional improvement, when ensures that high prices/low output persists. An analysis of such traps would be distracting and is therefore beyond the scope of this paper.

The substantive component of this assumption is that the elasticity is less than two, not that it is constant nor that it is equal to one per se. A generalization to other constant elasticity functions is provided in the appendix. Under this assumption, we have

$$M(I, \delta) = \gamma_1 + \gamma_2 \cdot N \cdot s \cdot \frac{\alpha \cdot (1 - \delta)}{\delta + \alpha \cdot (1 - \delta) \cdot I}$$

The fact that this expression is independent of  $C$  does not generalize to other constant elasticity demand functions (see the appendix).

Returning to the designer's problem, the optimal choice of  $e$  is that which solves

$$\max_{e \geq 0} R(e) - e.$$

The optimal investment is zero if the marginal benefit is negative at  $e = 0$ , and is positive otherwise. That is,

$$e^*(I, \delta) \begin{cases} = 0 & \text{if } M(I, \delta) \cdot z < 1 \\ \text{satisfies } M(I, \delta) \cdot Z'(e^*) = 1 & \text{otherwise} \end{cases} \quad (17)$$

Since  $M$  is strictly decreasing in  $I$ , institutions improve if and only if institutions are sufficiently weak at prevailing norms. Specifically, institutions improve if and only if  $I < \hat{I}(\delta)$ , where  $\hat{I}(\delta)$  satisfies  $M(\hat{I}(\delta), \delta) \cdot z = 1$ . Some manipulation reveals that

$$\hat{I}(\delta) = N \cdot s \cdot \frac{z \cdot \gamma_2}{1 - z \cdot \gamma_1} - \frac{\delta}{1 - \delta} \cdot \frac{1}{\alpha}. \quad (18)$$

Notice that  $\hat{I}$  is a strictly decreasing concave function where  $\hat{I}(0) = N \cdot z \cdot \gamma_2 / (1 - z \cdot \gamma_1) > 0$ , and  $\lim_{\delta \rightarrow 1} \hat{I}(\delta) = -\infty$ . The function  $\hat{I}$ , as well as the point  $\hat{I}(0)$ , are depicted in Figure 3.

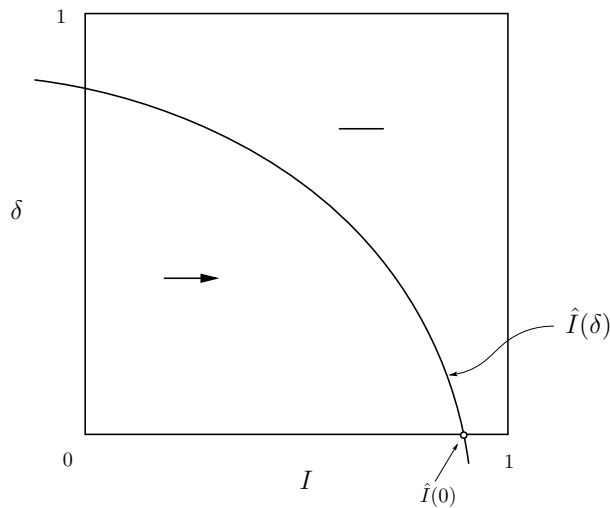


Figure 3:  $\hat{I}(\delta)$  and the Dynamics of  $I$

Institutional dynamics are summarized as  $\dot{I}(I, \delta) = Z(e^*(I, \delta))$ , where  $e^*(I, \delta)$  is given by (17). This, together with the boundary constraints, implies that the institutional dynamics are given by:

$$\dot{I}(t) \begin{cases} > 0 & \text{if } I(t) < \hat{I}(\delta(t)) \text{ and } I(t) < 1 \\ = 0 & \text{otherwise.} \end{cases} \quad (19)$$

These dynamics are also displayed in Figure 3: institutions improve for points strictly below  $\hat{I}(\delta)$ , but do not change for points above  $\hat{I}(\delta)$ .

## 2.5 Steady States

The system dynamics are captured by (14) and (19). As usual, a *steady state*, is defined as follows.

**Definition 1.** A point  $(I^*, \delta^*)$  is a steady state if  $\dot{I}(I^*, \delta^*) = \dot{\delta}(I^*, \delta^*) = 0$ .

We restrict attention to steady states that are *stable* in the sense that arbitrarily small changes in the proportion of the population that cheat do not cause a permanent divergence from the steady state. The following definition formalizes this.

**Definition 2.** A steady state is stable if there exists a  $\bar{\varepsilon} > 0$  such that for all  $\varepsilon \in [0, \bar{\varepsilon}]$ ,  $\dot{\delta}(I^*, \delta^* - \varepsilon) > 0$  and  $\dot{\delta}(I^*, \delta^* + \varepsilon) < 0$ .

In order to find stable steady states, we combine the curves  $\hat{\delta}$  and  $\hat{I}$  from Figures 2 and 3 in Figure 4, along with their associated dynamics. Figure 4 reveals that two types of stable steady state may emerge in the long-run. To see this, we can sketch out how the economy transitions from any given point,  $(I, \delta)$ , to a stable steady state. Some possible transition paths are illustrated in Figure 5, where each dark dot represents an initial condition. Of the transition paths that start at  $I = 0$ , those with relatively high initial values of  $\delta$  end in a region where  $\delta = 1$  whereas those with low values of  $\delta$  end in a region where  $\delta = 0$ .

**Proposition 2.** Let  $(I^*, \delta^*)$  be a stable steady state. Then the steady state either exhibits High Trust/Functional Institutions ( $\delta^* = 1$ ) or Low Trust/Dysfunctional Institutions ( $\delta^* = 0$ ). Furthermore

1.  $I^* \in [0, 1]$  is part of a high trust steady state if and only if  $I^* > I_1$ .
2.  $I^* \in [0, 1]$  is part of a low trust steady state if and only if  $\min\{\hat{I}(0), 1\} \leq I^* < I_2$ .

*Proof.* See appendix. □

Given these possibilities, we can partition the  $(I, \delta)$  space into a set of points that lie on transition paths that lead to a functional institutions/high trust steady state, and the complementary set of points that lead to a dysfunctional institutions/low trust steady state. We focus on a situation

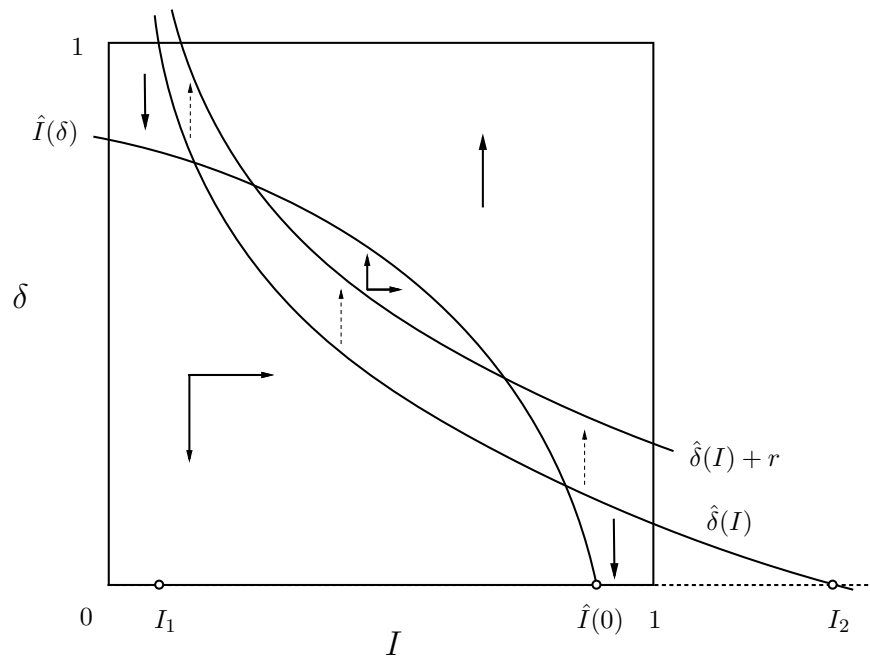


Figure 4: Long-Run Dynamics: Phase Diagram

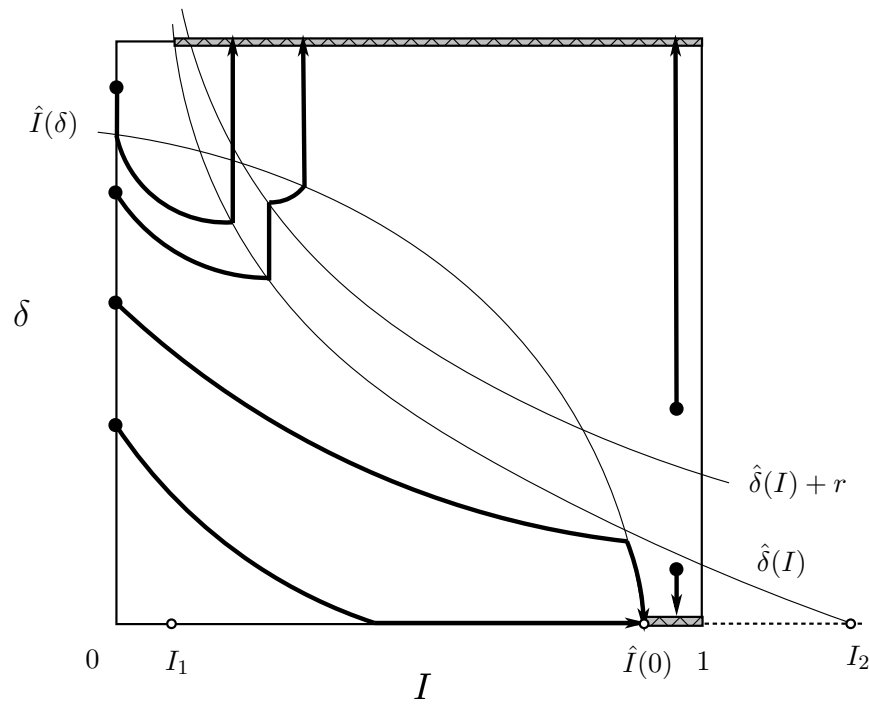


Figure 5: Long-Run Dynamics: Transition Paths



in which both types of steady states exist in order to analyze the ways in which the parameters of the model influence the likelihood that an economy transitions to a high trust steady state.<sup>19</sup>

Note that a high trust steady state will be reached if we start at any point in the region where honesty norms are strictly improving ( $I > \hat{\delta}(I)$ ), such as the shaded region in Figure 6. This is because, as mentioned previously, the payoff to acting honestly is increasing in the proportion of agents that act honestly (since the payoff to cheating, the equilibrium price, falls). But a high trust steady state may also be reached from a point at which honesty norms are deteriorating. This is because institutions may improve sufficiently quickly that the payoff to cheating falls to the point at which honesty norms begin to strictly improve (putting us back into the case just mentioned).

To formalize this, let  $Z = \{(I, \delta) \mid (I, \delta) \in [0, 1]^2, \dot{\delta}(I, \delta) > 0\}$  be the set of points at which we know will transition to the high trust steady state. Any points that lie on transition paths that enter  $Z$  will therefore also transition to the high trust steady state. Consider the points outside of  $Z$  that i) lie on an open boundary of  $Z$ , and ii) experience institutional growth. Such points must also transition to a high trust steady state since they will transition into  $Z$  within an arbitrarily short period of time. Let this set of points be denoted  $\tilde{Z}$ . That is

$$\tilde{Z} \equiv \{(I, \delta) \mid (I, \delta) \in [0, 1]^2, \dot{\delta}(I, \delta) = 0, \dot{I}(I, \delta) > 0\} \quad (20)$$

Associated with each point in  $\tilde{Z}$  is a transition path<sup>20</sup>, and such paths become flat as they approach their associated point in  $\tilde{Z}$  ‘from the left’, since  $\dot{\delta} = 0$  at such points by definition. This, along with fact that transition paths can not cross, implies that all points ‘above’ any such path also transition to some point in  $\tilde{Z}$ , and therefore also transition to a high trust steady state. The *critical transition path* is the highest path for which no lower path encounters a point in  $\tilde{Z}$ . It is therefore the path that passes through the point on the lower boundary of  $\tilde{Z}$ .<sup>21</sup> The critical transition path is depicted as the dark transition path in Figure 6.

By adding all the points that lie above the critical transition path, Figure 7 shows the entire set of points that lead to high trust steady states (the shaded area). The left panel corresponds the case depicted in Figure 6. The middle panel shows how high trust steady states need not be accessible

<sup>19</sup>If  $\min\{I_2, 1\} \leq \min\{\hat{I}(0), 1\}$ , then there are no low trust steady states and *all* points transition to the high ones. Similarly, if  $1 < I_1$  then steady states with high trust do not exist and *all* points transition to the low ones. Thus, we focus on the case where  $I_1 \leq 1$  and  $\min\{\hat{I}(0), 1\} \leq \min\{I_2, 1\}$ .

<sup>20</sup>It is not possible to produce an explicit expression for a transition path, since the  $\phi$  function is left general. Even if  $\phi$  were given a functional form, it is unlikely that an explicit expression for the transition path would exist. Despite this, in the region where institutions are strictly increasing, the transition path that passes through the point  $(I_x, \delta_x)$  can be described by a function in  $(I, \delta)$  space. This function, denoted  $T(I)$ , is the solution to the initial values problem defined by  $T'(I) = \dot{\delta}(I, \delta)/\dot{I}(I, \delta)$  and  $T(I_x) = \delta_x$ . In all other regions the transition path is vertical, and therefore is straightforward to describe. Given that explicit expressions for transition paths are not guaranteed to exist, and along with a desire to present the main ideas as transparently as possible, we choose to focus on the *qualitative* dynamic properties of the system.

<sup>21</sup>Formally, this point is  $(\hat{\delta}^{-1}(\delta_c), \delta_c)$  where  $\delta_c = \inf\{\delta \mid \exists I : (I, \delta) \in \tilde{Z}\}$ .

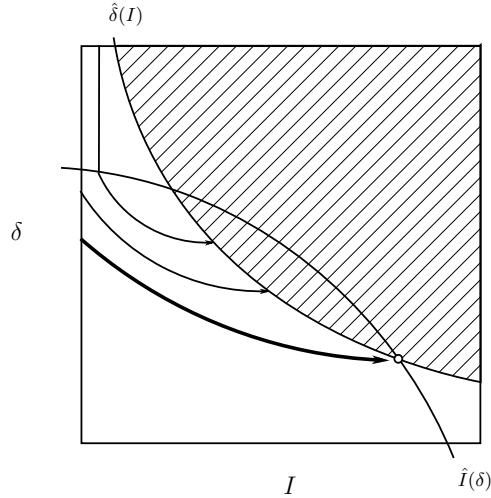


Figure 6: The Critical Transition Path

when institutional quality is very low. The right panel covers the case in which  $\hat{\delta}$  and  $\hat{I}$  intersect only once.<sup>22</sup>

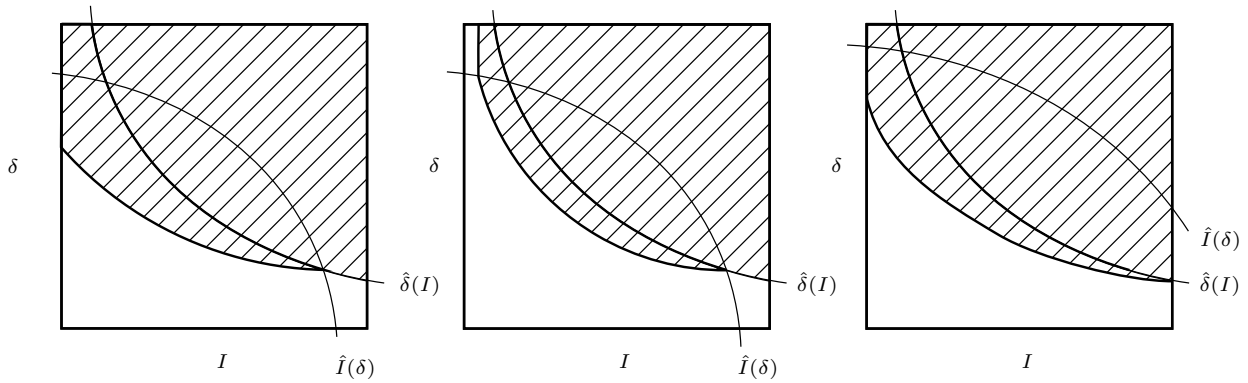


Figure 7: Regions that Lead to High Trust Steady States

### 2.5.1 The Effect of Scale

In this section we explore how a larger population influences the likelihood that an economy ends up with high trust in the long-run, as measured by the size of the set of points that lead to high trust steady states.

Simply, a greater population increases the designer's return to institutional improvement. This has two related consequences; first, when institutional improvements occur they occur to a greater extent, and second, institutional improvements occur over a greater range of states. This can only

<sup>22</sup>The fourth case is one in which  $\hat{\delta}$  and  $\hat{I}$  never intersect. In this case, a high trust steady state is reached from a point if and only if the point lies in Z.

increase the likelihood that the economy will achieve the minimal institutional quality required in order to dissuade rational types from cheating. Geometrically, these two effects respectively correspond to i) a flattening of all transition paths in the region where institutions improve, and ii) an outward shift of the  $\hat{I}$  function.<sup>23</sup> These two effects are displayed in Figure 8 (and are established more formally in the appendix), where the cross-hatched region indicates the additional set of points that lead to a high trust steady state as a result of a greater population.

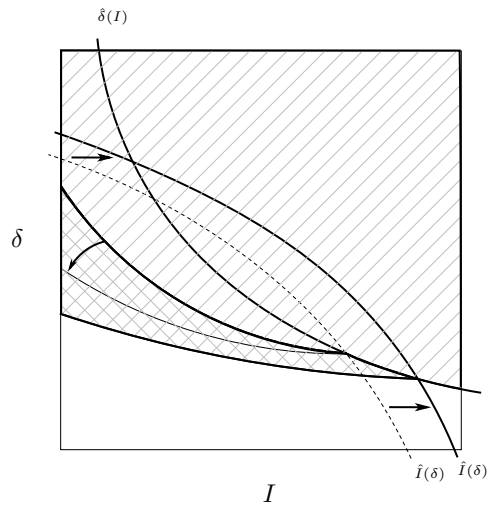


Figure 8: Effect of Scale

While a greater  $N$  increases the likelihood that an economy arrives at a high trust steady state, the effect of  $N$  on long-run institutional quality is ambiguous. This is illustrated in Figure 9 (which assumes  $r = 0$  for clarity), where a greater  $N$  flattens the transition path but also increases  $\hat{\delta}(I)$ . If the slope of the transition path is only slightly flattened relative to the shift in  $\hat{\delta}(I)$ , then institutional quality increases whereas it decreases if the path is greatly flattened. Bigger countries have higher returns to institutional improvement, but because they have greater incentive to improve institutions, the transition to functional institutions is more rapid. On the one hand, this implies institutions should be better in bigger countries. But, on the other, functional institutions will then arise at higher levels of honesty so that institutional quality can be lower at the point when they become functional. The combination of these two effects implies that there is no necessary correlation between country size and institutions, even though institutions are a key part of the reason that trust increases with country size.

These are the predictions that will be taken to the data in the following section. A larger scale helps societies reach the functional institutions/high trust outcome - but there is, in fact, an ambiguous effect of scale on institutional quality.

<sup>23</sup>A greater population has no effect on  $\hat{\delta}$  since higher demand is compensated for with greater entry, thereby keeping prices independent of  $N$ .

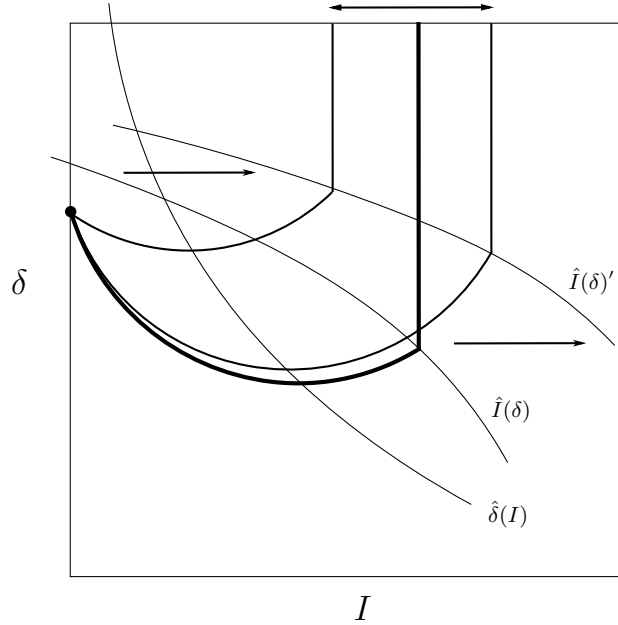


Figure 9: Ambiguous Institutional Quality

**Conjecture 1.** *An economy is more likely to have high trust levels the larger its population size. This should hold even controlling for institutional quality and controlling for other correlates of country-level trust such as income per capita.*

This prediction has not been made by any previous model to our knowledge. There is no prediction made about the relationship between economy size and institutions, nor between trust and institutions. The additional empirical implications generated by the model that are consonant with previous theories, and or for which there is no possibility of establishing causal relations, are not considered.

### 3 Testing the Model

Since our model makes prediction about the steady state relationship between Trust and population size at the country level, we test this by looking at a cross-section of countries. The main limitation to doing this is the availability of information about trust levels for a large number of countries. The World Values' Survey is our source here. We also wish to include institutional quality measures as well, and these derive from various sources detailed below.

### 3.1 Data

The key dependent variable comes from the generalized trust question from the World Values Survey.<sup>24</sup> The World Values Survey has been run for nearly 30 years, and comprises five waves: 1980, 1990, 1995, 2000 and 2005. The trust question has proved reasonably stable in its patterns of response across countries through time. The question is: “Generally speaking, would you say that people can be trusted or that you can’t be too careful in dealing with people?”. For a given sample, our measure of trust is the percentage of individuals answering “Yes” to this question.

Though relatively stable, the correlation of this measure across waves is not perfect.<sup>25</sup> Additionally, the set of countries sampled varies considerably across waves. In order to best account for the sampling variation across waves, maximize our use of the data, and remove any arbitrariness that may arise from focusing on a single year, we construct our “Trust” dependent variable by pooling all survey waves by country and calculating the percentage of individuals that answer “Yes” to the trust question. Our core sample thus consists of 72 countries with a mean of 28.6 and standard deviation 14.6. The country with lowest trust value in our sample is Brazil with 6.2 and the highest is Norway with 66.3. The countries we use, as well as the survey waves for which their trust data are available, are given in Table 4 in the appendix.

To test our theory, the key independent variable is population size. We take the logarithm of population from the Penn World Table 6.3 (Heston *et al.* (2009)) for the 5 years corresponding to the World Values Survey waves and compute a country level average over those years.<sup>26</sup> Our theory predicts that population size should have a positive effect on trust levels, whether or not institutional quality is controlled for. We will test both of these predictions here.

With respect to the other variables to be included in our regressions, we are guided by previous literature. Previous studies on the determinants of trust have repeatedly focused on a core set of key correlates (Berggren and Jordahl (2006) and Bjørnskov (2006)). Some of the baseline regressors we use are historically determined, but others have the potential to be endogenous to country trust levels. Since we are not directly interested in most of these variables, but instead wish to make

---

<sup>24</sup>We obtain the 2005 data from the World Values Survey 1981-2008 official aggregate (v.20090901, 2009. World Values Survey Association ([www.worldvaluessurvey.org](http://www.worldvaluessurvey.org)). Aggregate File Producer: ASEP/JDS, Madrid) and the 1980-2000 data from European and World Values Surveys four-wave integrated data file, 1981-2004 (v.20060423, 2006. Surveys designed and executed by the European Values Study Group and World Values Survey Association. File Producers: ASEP/JDS, Madrid, Spain and Tilburg University, Tilburg, the Netherlands. File Distributors: ASEP/JDS and GESIS, Cologne, Germany.) Further details, including data download, are available at <http://www.worldvaluessurvey.org/>.

<sup>25</sup>For example, for the 33 countries sampled in both 2000 and 2005, the correlation is only 0.7221.

<sup>26</sup>Since relative sizes of countries across our sample years have not varied greatly, it turns out that none of our results are affected by taking the values of population in a single year instead. We can also fully replicate all of the results that will be presented here using population levels instead of the logarithm of population levels. The problem with doing this is that the immense scale differences lead to the very large countries, particularly India and China, exerting undue influence on the regression as outliers. Including population in log scale, as is standard (see Bjørnskov (2006) for example) avoids the outlier influence of larger countries.

sure that we are not mis-specifying our analysis by omitting important trust correlates, we also include these potentially endogenous correlates of trust in our regressions below.

In terms of control variables, GDP per capita is clearly of first order importance as it has almost always been found to be strongly positively correlated with trust.<sup>27</sup> We use data from the Penn World Table 6.3. Income inequality also is among the most consistently significant cross-country determinants of trust, and it exerts a negative effect (Knack and Keefer (1997), Zak and Knack (2001), Uslaner (2002)); see Jordahl (2007) for a survey. We use data on the gini coefficient taken from United Nations Human Development Report.<sup>28</sup> An adverse effect of diversity has been thought to extend across other dimensions as well – Ethnic, Religious and Linguistic. Accordingly we include measures of fractionalization for each one of these three dimensions from Alesina *et al.* (2003). We also include an ex-communist dummy as it has been argued that communist dictatorships encouraged internal spying and mistrust within populations through extensive networks of informants, and that the oppressive communist dictatorships of Central and Eastern Europe eroded trust levels therein; see Bjørnskov (2006). The final variable we add to our core set of controls is a dummy for the Scandanavian countries. These countries score consistently high on trust, irrespective of the set of additional controls used, so through a mechanism distinct from the other trust determinants. Table 1 reports the summary statistics of these variables.

Table 1: Summary statistics

Variable	Mean	Std. Dev.	Min.	Max.	N
Trust	28.545	14.636	6.154	66.342	72
ln(Popn)	16.689	1.664	12.49	20.893	72
GDP/cap. ('000s)	11.812	8.846	0.554	42.359	72
Ethnic Frac.	0.346	0.233	0.012	0.930	72
Religious Frac.	0.403	0.233	0.004	0.86	72
Linguistic Frac.	0.32	0.267	0.018	0.923	72
Scandanavian	0.056	0.231	0	1	72
Ex-Socialist	0.167	0.375	0	1	72
Gini Coefficient	36.88	7.856	24.7	58.5	65
Physical Property (PPR)	6.383	1.294	3.7	8.6	72
Legal/Political (LP)	5.778	2.002	1.8	8.9	72
Intellectual Property (IPR)	5.857	1.775	2.1	8.5	72

<sup>27</sup>For example, Knack and Keefer (1997), Zak and Knack (2001), and Delhey and Newton (2005).

<sup>28</sup>Data available at <http://hdrstats.undp.org/en/buildtables/>. This data represents an average across the 1992-2007 period for years in which data were available.

### 3.1.1 Not Conditioning On Institutions

Table 2 reports our first set of findings where we test the unconditional (upon institutions) predictions of the model. Column 1 reports the basic correlation between log population size and trust in the raw data. These two pieces of information are available for more than our core 72 countries (a total of 93). For completeness, the first column reports the raw correlation between them in this larger sample which we will not use again since the additional 21 countries are ones for which other important trust determinants are unavailable. As can be seen, though the coefficient is positive, it is indistinguishable from zero. Column 2 restricts the sample to countries for which the additional controls that we have discussed are all available. Once again we see a positive raw correlation which is essentially indistinguishable from zero. Column 3 is our first attempt to test the model. In this column we add the most widely reported trust correlate, GDP per capita, to the regression. GDP is itself positive and highly significant in its effect on trust, in line with all previous studies. Additionally, the effect of log population increases and now becomes significant at the 5% level.

One concern about the result presented in Column 3 is that it may be arising because we have omitted the other country level correlates that have been known to be determinants of trust. The key ones are those discussed above. The concern may be that these variables are, for some reason, correlated with population, so that by omitting them, the effects of known determinants are loading onto population. The next column adds the additional trust determinants from previous studies and thus comprises our core set of regressors, Column 4. As consistent with previous studies, income inequality is a negative determinant of trust, though being an ex-communist country has no effect. Scandinavian countries have significantly higher trust levels, but somewhat surprisingly, the social fractionalization measures are not significant. The main result, however, is that the coefficient on log population remains positive (and is even of greater magnitude) and highly significant (now at the 1% level). This is precisely in line with the theory.<sup>29</sup>

Figure 10a displays the partial scatter plot from the regression of log population on trust. As the scatter shows, the correlation does not seem to be strongly influenced by any particular countries. However, it has been argued, see [Bjørnskov \(2006\)](#), that China is an outlier in its unusually high trust levels. It also has the largest population. Consequently, it could be exerting a sizable influence on the results. So Column 5 repeats the specification in the previous column with the omission of China. Though taking China out slightly lowers the magnitude of log population's effect, it still remains significant at the 5% level.<sup>30</sup>

---

<sup>29</sup> Education has been argued to be an important determinant by building trust in the classroom, for example [Knack and Keefer \(1997\)](#) and [Knack and Zak \(2002\)](#). Including years of Education (from Barro and Lee), it is never significant when GDP per capita is also included, and has almost no impact on the effect of population.

<sup>30</sup>The positive and significant effect of population is also robust to the omission of China and India. In fact, the result is robust to the omission of the  $L$  largest countries for all values of  $L$  up to at least 10.

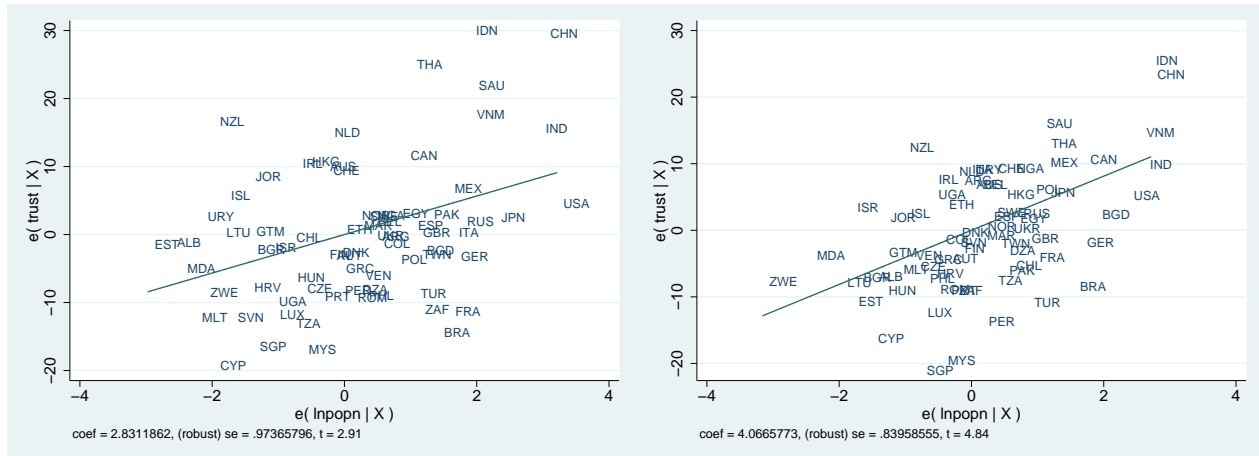
Table 2: Baseline Regressions

VARIABLES	(1) trust	(2) trust	(3) trust	(4) trust	(5) trust
ln(Popn)	1.214 (0.908)	0.139 (1.015)	2.156** (0.933)	2.831*** (0.974)	2.269** (0.909)
GDP/cap. ('000s)			0.964*** (0.198)	0.631*** (0.148)	0.621*** (0.159)
Ethnic Frac.				1.137 (9.139)	3.805 (8.678)
Religious Frac.				5.615 (7.637)	4.132 (7.620)
Linguistic Frac.				-7.362 (8.128)	-7.734 (7.775)
Scandanavian				26.94*** (3.404)	25.65*** (3.345)
Ex-Socialist				0.755 (3.559)	-1.647 (3.140)
Gini Coefficient				-0.438** (0.183)	-0.528*** (0.163)
Constant	6.410 (15.08)	26.22 (17.09)	-18.83 (15.36)	-12.77 (19.71)	-0.0481 (17.44)
Observations	93	72	72	72	71
$R^2$	0.018	0.000	0.287	0.572	0.590

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1





(a) Table 2, Column 4

(b) Table 3, Column 4

Figure 10: Partial Scatter Plots From Tables 2 and 3

### 3.1.2 Conditioning On Institutions

One may conjecture that a much simpler explanation for the observed population/trust relationship is available. Specifically, accepting the [Demsetz \(1967\)](#) argument that there are economies of scale in the creation of institutions, a variant of which is in fact present in our theory, one may conjecture that the previous results are a reflection of the following chain of effect: Scale effects make larger countries have better institutions and better institutions in turn lead to higher levels of trust. Note that this mechanism, though related, is much simpler and distinct from the mechanism in our model. To see the difference, note again that our theory makes no predictions about the effect of population size on institutions. Recall that there is one force leading to a positive effect of size on institutional quality, and another effect leading to a negative impact of size. However, our theory does suggest that functional institutions should arise more readily in larger countries (though perhaps at lower quality levels) so that the positive effect of population size on trust levels is unambiguous. An even sharper contrast to this type of simpler explanation is our theory's prediction that the effect of size on trust should persist over and above any effect that size has on institutions. According to the simpler explanation we have just recounted, including institutions in the regression should negate any positive effect of size on trust. Thus, in Table 3, we test the prediction that population levels continue to exert a positive effect on trust levels even when controlling for any effects of institutions.

Testing this prediction requires coming up with a measure of institutions to use as controls. Though there are a number of institutional variables and indexes that have been compiled that could be used, there is a unique recently constructed set of indexes (compiled by [Stroková \(2010\)](#)) that are ideal for our purposes here. The International Property Rights Index (IPRI) was first pro-

Table 3: Conditioning On Institutions

VARIABLES	(1) trust	(2) trust	(3) trust	(4) trust	(5) trust
ln(Popn)	2.344*** (0.857)	3.424*** (0.868)	3.718*** (0.872)	4.067*** (0.840)	3.643*** (0.845)
IPRI	6.054*** (0.849)				
Physical Property (PPR)		7.379*** (1.738)	7.469*** (1.671)	7.419*** (1.759)	7.002*** (1.722)
Legal/Political (LP)		4.890*** (1.658)	4.309*** (1.613)	2.951* (1.644)	2.740 (1.657)
Intellectual Property (IPR)		-4.651** (1.799)	-6.001*** (1.811)	-6.698*** (1.673)	-6.285*** (1.708)
GDP/cap. ('000s)			0.461 (0.299)	0.603** (0.242)	0.600** (0.261)
Ethnic Frac.				4.769 (6.928)	6.145 (6.698)
Religious Frac.				2.175 (5.598)	1.490 (5.631)
Linguistic Frac.				-7.899 (5.613)	-8.100 (5.467)
Scandanavian				20.45*** (3.358)	20.04*** (3.301)
Ex-Socialist				4.746* (2.436)	3.044 (2.196)
Gini Coefficient				-0.474*** (0.169)	-0.527*** (0.160)
Constant	-46.91*** (15.21)	-76.71*** (14.85)	-76.36*** (15.41)	-56.55*** (17.07)	-46.09*** (16.78)
Observations	72	72	72	72	71
R <sup>2</sup>	0.386	0.510	0.528	0.730	0.731

Robust standard errors in parentheses

\*\*\* p<0.01, \*\* p<0.05, \* p<0.1

duced in 2007. It has three sub-components and aims to capture precisely what we mean by the term “institutions” in this paper, as it is focused on those institutions that are most important in facilitating trade. The first component of the IPRI is the *Legal and Political Environment* (LP), which is itself derived from four sub-components “Judicial Independence” from the World Economic Forum’s Global Competitiveness Index, “Rule of Law”, “Political Stability”, and “Corruption”, all three of which are from the World Bank Institute’s Governance Matters study. These are relevant for our notion of institutions to the extent that they affect agents’ perceptions regarding the enforcement of property rights. The second component is *Physical Property Rights* (PPR), which is derived from the following three sub-components: “Protection of Physical Property Rights”, that is, the strength of a country’s property rights system as it reflects experts’ views on the quality of judicial protection of private property, including financial assets; The ease with which agents can “Register Private Property”, which comes from the World Bank’s Doing Business report, and “Access to Loans” from the World Economic Forum’s 2008-2009 Global Competitiveness Index. The third component is *Intellectual Property Rights* (IPR) which comprises: “Protection of Intellectual Property”, from the World Economic Forum’s 2008-2009 Global Competitiveness Index; “Patent Protection” from the Ginarte-Park Index of Patent Rights; and a measure of “Copyright Piracy” from the International Intellectual Property Alliance.<sup>31</sup> Summary statistics for the IPRI, as well as the three components, are contained in Table 1.

Column 1 of Table 3 adds the IPRI’s single aggregate index (averaging the three sub-indexes) to the regression of trust on log population. Strikingly, the addition of this institutional index, which is itself positive in its effect on trust, makes the coefficient on log population positive and significant at the 1% level, precisely as our theory would predict, even without the addition of traditional trust determinants. Column 2 includes the sub-components of the IPRI index, (PPR, LP and IPR) separately. Despite their strong positive correlation (above 0.8 between any pair) they are distinct in their effects on trust. Once again, population has a strong positive effect on trust.<sup>32</sup> Column 3 adds the first of the previously known trust determinants, GDP per capita, to this specification. Interestingly, GDP per capita is now not significant in this specification, as it is highly correlated with the indexes of institutions (0.87 with the aggregate index). Importantly however, its addition hardly alters the effect of population reported earlier. Column 4 includes the full array of controls, both institutional and previous determinants. The coefficient on log population rises to 4.067 and is highly significant, and the pattern of influence exerted by the additional variables is again in line with how they have been reported in previous studies.

Once again, these findings are precisely in line with the predictions of the model. Population

---

<sup>31</sup>Data, and further details are available at [www.internationalpropertyrightsindex.org](http://www.internationalpropertyrightsindex.org).

<sup>32</sup>Strikingly, Intellectual Property Rights exert a negative effect - though it should be noted that this is only true conditional upon Private Property Rights and Legal and Political ones also being present. When entered on its own, it is positive and significant.

size leads to higher trust, over and above any effect it may have on institutions. The scatter plot, Figure 10b, moreover shows this to be a very robust partial relationship that is not driven by one, or a small set of countries. This is confirmed in column 5 for the case of China, whose omission, though lowering the coefficient on population, affects its significance hardly at all.

One may object to the use of these indices over other measures of institutions that exist. We use them because they are the most recent and comprehensive measures of precisely the types of institutions that we are studying in this paper. However, we have also experimented with other of the commonly used measures that could be thought of as proxies for the model's institutions variables ("government effectiveness" for political components, "rule of law" for legal ones, the International Country Risk Guides "corruption" measure as a measure of bureaucratic probity to name a few). Replacing the comprehensive institutions measures with any of these does not change the basic findings above.

Turning to the magnitude of the effect, in variance normalized terms, a one standard deviation increase in log population, leads to a low of 3.2 (unconditional on institutions) point increase to a high of a 6.4 point (conditional on institutions) increase in trust. Recall that the standard deviation of trust was 14.6.

The positive relationship between population and trust that we document is unlikely to derive from reverse causation, or to have been predicted on a priori grounds. In micro level studies, it has been found that individuals residing in larger cities have lower levels of trust. Since larger countries also have larger cities, this is not directly a problem for the testing of our theory, as it would suggest an a priori reason to find a negative correlation between country size and trust. Also, as Bjørnskov (2006) notes, much network analysis finds that trust is more likely to evolve in small networks of individuals, suggesting again that larger countries may be less able to build it, and, if anything a negative bias.<sup>33</sup>

We have run many different variations of these baseline regressions without significant changes occurring. To mention a few, the positive population trust effect is also there in all waves treated separately, and significant at least at the 10 per cent level with the full set of regressors.<sup>34</sup> We have also experimented with weighting observations according to how many survey waves the trust data are drawn from, including the average survey year as a regressor, and with various different treatments of the time varying regressors, none of which affect the qualitative results.

---

<sup>33</sup>However, if it is the case that countries with larger populations tend to also be agrarian, and hence most of their population does not reside in cities, this would lead to a positive correlation between trust and population for reasons that have nothing to do with our theory. In order to check for this, we have also run a specification including a variable measuring the percentage of the population that lives rurally, taken from the UN population database. <http://esa.un.org/unup> The United nations World Urbanization Prospects. This is never significant and hardly alters the significance or magnitude of the coefficient on population in either Table 2 or 3.

<sup>34</sup>The only wave for which this is not true is the first, 1980. However, for this wave only 18 countries are in the baseline regression.

In conclusion, though we are aware of the difficulty in making strong causal inferences from cross-country regressions, and wary about testing reduced form implications of a dynamic model, the empirical results we have presented here do, at the very least, fail to rule the model out. There are no a priori grounds to expect population size to exert a positive influence on trust, no grounds for expecting trust to positively affect population size, if anything reasons to expect a negative relation between population and trust, and no theories we are aware of other than the one we have presented which would predict a positive relationship, either conditional upon institutions or unconditionally.

## 4 Conclusions

This paper has explored two factors that make individuals trustworthy; internalized honesty norms and incentives that are shaped by institutions. The model we have developed explicitly studies the dynamic evolution of honesty norms, that are affected by institutional quality, and institutions, whose development is simultaneously affected by honesty norms. The dynamic co-dependence between norms and institutions implies two types of social outcomes: (i) functional institutions/high trust where honesty norms are widely followed, rational agents are honest and trade is widespread and (ii) dysfunctional institutions/low trust, where honesty norms are not followed, agents act opportunistically and trade is limited.

The model predicts that countries with larger populations, where incentives for institutional improvement are greater, should be more likely to attain the functional institutions/high trust outcomes. It thus predicts a positive correlation between population size and trust, controlling for extraneous factors, previously studied trust correlates, and institutional quality. The data is consistent with that theoretical prediction.

[Acemoglu et al. \(2005\)](#) identify a number of reasons that institutions are important for economic growth. Primarily, they affect incentives: to invest in physical and human capital, to develop new technologies, to organize production. To these factors we add another benefit; functional institutions, by making promised actions incentive compatible, also build trust. Trust is important, in and of itself, but is increasingly being understood as a key determinant of economic activity.<sup>35</sup> So our analysis indicates an additional indirect channel whereby institutions may matter for economic growth.

The above analysis also suggests further dynamic benefits to the development of functional institutions. By continually rewarding good actions and punishing bad, functional institutions reinforce internalized honesty norms. When technology introduces new trading opportunities, and when the institutions required to regulate these opportunities do not yet exist, these norms are all that stand in the way of trade breaking down. When these norms are widespread, trade can

---

<sup>35</sup>A recent paper that makes this case forcefully is [Algan and Cahuc \(2009\)](#).

thus commence, and the future institutions that will work to regulate and maintain it have the chance to follow.

## APPENDIX

### A Model

#### A.1 Proofs

##### Proof of Proposition 1

*Proof.* We shall begin by looking for values of  $I$  such that rational types cheating (i.e.  $\sigma = 1$ ) is an equilibrium. If rational types cheat, then  $\delta = (1 - r) \cdot \beta$ , which needs to be less than  $\hat{\delta}(I)$  in order for cheating to be optimal. Thus, rational types cheating is an equilibrium if  $(1 - r) \cdot \beta \leq \hat{\delta}(I)$ .

Now we look for values of  $I$  such that rational agents not cheating (i.e.  $\sigma = 0$ ) is an equilibrium. If  $\sigma = 0$ , then  $\delta = (1 - r) \cdot \beta + r$ , which we require to be greater than  $\hat{\delta}(I)$  if rational agents do not find cheating optimal. Thus, rational types not cheating is an equilibrium if  $\hat{\delta}(I) \leq (1 - r) \cdot \beta + r$ .

Since  $(1 - r) \cdot \beta \leq (1 - r) \cdot \beta + r$  it then follows that at least one pure strategy instantaneous equilibrium always exists. Thus, cheating is the *unique* pure-strategy equilibrium if and only if not cheating is not an equilibrium: i.e. when  $\hat{\delta}(I) > (1 - r) \cdot \beta + r$ . Similarly, not cheating is the *unique* pure-strategy equilibrium if and only if cheating is not an equilibrium: i.e. when  $\hat{\delta}(I) < (1 - r) \cdot \beta$ . In all other cases, both equilibria exist.<sup>36</sup> □

##### Proof of Proposition 2

*Proof.* We begin by establishing that all stable steady states involve  $\delta^* \in \{0, 1\}$ .

**Lemma 1.** *Let  $(I^*, \delta^*)$  be a stable steady-state. Then either  $\delta^* = 1$  or  $\delta^* = 0$ .*

*Proof.* By contradiction, suppose  $\delta^* \in (0, 1)$ . From the equations of motion (15), the only way that  $\dot{\delta} = 0$  is if  $\delta^* = \hat{\delta}(I^*)$ . Thus, for any  $\varepsilon > 0$  we have  $\delta^* - \varepsilon < \hat{\delta}(I^*)$ . Furthermore, for  $\varepsilon$  sufficiently small we must have  $\delta^* - \varepsilon > 0$ . But then  $\dot{\delta}(I^*, \delta^* - \varepsilon) < 0$  for sufficiently small  $\varepsilon$ , contradicting the supposed stability. □

From this we know that if  $(I^*, \delta^*)$  is a stable steady state, then  $\delta^* \in \{0, 1\}$ . These two cases are analyzed in turn.

**Case 1:**  $\delta^* = 1$ .

In order for  $(I^*, 1)$  to be a stable steady state,  $(I^*, 1)$  must be a steady state. In order for  $\dot{\delta} = 0$  we must have  $\delta^* \geq \hat{\delta}(I^*)$ . Since  $\delta^* = 1$ , this is equivalent to  $\hat{\delta}(I^*) \leq 1$ . Given that  $\hat{\delta}(\cdot)$  is strictly

---

<sup>36</sup>In fact there is another (non pure strategy) equilibrium in this range where rational types are indifferent to cheating because a proportion  $\sigma \in (0, 1)$  of rational types cheat. This equilibrium is unstable and is ignored.

decreasing, this is equivalent to  $I^* \geq I_1$ . In order for  $\dot{I} = 0$ , the value of  $I^*$  must satisfy either i)  $I^* \geq \hat{I}(1)$ , or ii)  $I^* = 1$ . Since  $\hat{I}$  goes to  $-\infty$  as  $\delta \rightarrow 1$ , it follows that  $I^* \geq \hat{I}(1)$  for all  $I$ . Therefore  $(I^*, 1)$  is a steady state if and only if  $I^* \geq I_1$ .

In order for  $(I^*, 1)$  to be a *stable* steady state,  $(I^*, 1)$  must be stable.  $(I^*, 1)$  is stable if and only if  $1 - \varepsilon > \hat{\delta}(I)$  for sufficiently small  $\varepsilon$ . That is, if and only if  $\hat{\delta}(I) < 1$ . Since  $\hat{\delta}(I)$  is strictly decreasing, this holds if and only if  $I^* > I_1$ . Therefore  $(I^*, 1)$  is stable if and only if  $I^* > I_1$ .

Therefore  $(I^*, 1)$  is a stable steady state if and only if  $I^* > I_1$ .

**Case 2:  $\delta^* = 0$ .**

In order for  $(I^*, 0)$  to be a stable steady state,  $(I^*, 0)$  must be a steady state. In order for  $\dot{\delta} = 0$  we must have  $\delta^* \leq \hat{\delta}(I^*)$ . Since  $\delta^* = 0$ , this is equivalent to  $0 \leq \hat{\delta}(I^*)$ . Given that  $\hat{\delta}(\cdot)$  is strictly decreasing this is equivalent to  $I^* \leq I_2$ . In order for  $\dot{I} = 0$ , the value of  $I^*$  must satisfy either i)  $I^* \geq \hat{I}(0)$ , or ii)  $I^* = 1$ . Together, these are equivalent to  $I^* \geq \min\{\hat{I}(0), 1\}$ . Therefore  $(I^*, 0)$  is a steady state if and only if  $\min\{\hat{I}(0), 1\} \leq I^* \leq I_2$ .

In order for  $(I^*, 0)$  to be a *stable* steady state,  $(I^*, 0)$  must be stable.  $(I^*, 0)$  is stable if and only if  $\varepsilon < \hat{\delta}(I)$  for sufficiently small  $\varepsilon$ . That is, if and only if  $0 < \hat{\delta}(I)$ . Since  $\hat{\delta}(I)$  is strictly decreasing, this holds if and only if  $I^* < I_2$ . Therefore  $(I^*, 0)$  is stable if and only if  $\min\{\hat{I}(0), 1\} \leq I^* < I_2$ .

Therefore  $(I^*, 0)$  is a stable steady state if and only if  $\min\{\hat{I}(0), 1\} \leq I^* < I_2$ . □

## Deriving The Designer's Problem

*Proof.* The change in consumer surplus over time, as a function of  $e$ , is:

$$\dot{\Psi} \equiv \frac{d\Psi}{dt} = \frac{d\Psi}{dp} \cdot \frac{dp}{dI} \cdot \frac{dI}{dt}.$$

Using the fact that  $dI/dt = Z(e)$ , as well as (3) to get the expression for  $p$ , along with

$$\begin{aligned} \frac{d\Psi}{dp} &= -ND(p) \\ \frac{dp}{dI} &= -\frac{C \cdot (1 - \delta) \cdot \alpha}{[\delta + (1 - \delta) \cdot \alpha \cdot I]^2} \end{aligned}$$

produces the result.

Notice that  $M$  can be written as

$$M(I, \delta) = \gamma_1 + \gamma_2 \cdot N \cdot D(p(I, \delta)) \cdot p(I, \delta)^2 \cdot \frac{(1 - \delta) \cdot \alpha}{C}.$$

Since  $I$  affects  $M$  only via  $p$ , and since  $p$  is decreasing in  $I$ ,  $M$  is decreasing in  $I$  if and only if  $M$  is increasing in  $p$ . This holds if and only if

$$\frac{d}{dp} D(p) \cdot p^2 > 0,$$

which holds if and only if:

$$2 \cdot D(p) \cdot p + D'(p) \cdot p^2 > 0.$$

Dividing both sides by  $D(p) \cdot p$  and rearranging gives that we require

$$-\frac{D'(p) \cdot p}{D(p)} < 2.$$

The left side of this is the absolute value of the price elasticity of demand. □

## A.2 Generalization of the Demand Function

Consider  $D(p) = s \cdot p^{-\eta}$ , where  $\eta \in (0, 2)$ . We have

$$M(I, \delta) = \gamma_1 + \gamma_2 \cdot N \cdot s \cdot \left[ \frac{C}{\delta + (1 - \delta) \cdot \alpha \cdot I} \right]^{2-\eta} \cdot (1 - \delta) \cdot \frac{\alpha}{C},$$

which is strictly decreasing in  $I$ . In this case we have

$$\hat{I}(\delta) \equiv B \cdot \left[ \frac{1}{1 - \delta} \right]^{1-\eta} - \frac{\delta}{1 - \delta} \cdot \frac{1}{\alpha},$$

where  $B$  is a constant defined by:

$$B \equiv \left[ \frac{C}{\alpha} \right]^{\frac{1-\eta}{2-\eta}} \cdot \left[ N \cdot s \cdot \frac{z \cdot \gamma_2}{1 - z \cdot \gamma_1} \right]^{\frac{1}{2-\eta}}. \quad (21)$$

Notice that  $\hat{I}$  shares the essential features of the unit elastic case. First, it is straightforward to see that  $\hat{I}(0) > 0$ . Second, since  $\lim_{\delta \rightarrow 1} (1 - \delta) \cdot \hat{I}(\delta) = -\alpha^{-1}$ , we know that  $\lim_{\delta \rightarrow 1} \hat{I}(\delta) = -\infty$ . Furthermore, we have

$$\hat{I}'(\delta) = \frac{\alpha \cdot B \cdot (1 - \eta) \cdot (1 - \delta)^{\eta} - 1}{\alpha \cdot (1 - \delta)^2}, \quad (22)$$

so that it is a decreasing function if the numerator is negative. This is ensured to hold for  $\eta \geq 1$ , and always holds at the point where  $\hat{I}(\delta) = 0$ . This is because  $\delta$  satisfies  $B = \delta \cdot (1 - \delta)^{-\eta} \cdot \alpha^{-1}$  at this point. In addition, the second derivative is

$$\hat{I}''(\delta) = \frac{\alpha \cdot B \cdot (2 - \eta) \cdot (1 - \eta) \cdot (1 - \delta)^{\eta} - 2}{\alpha \cdot (1 - \delta)^3}, \quad (23)$$

so that the function is concave if the numerator is negative. One immediate result is that concavity is ensured for all  $\eta \geq 1$ . Second, regardless of  $\eta$  this function is always concave at the point at which  $\hat{I}(\delta) = 0$  (which can be seen by substituting in  $B = \delta \cdot (1 - \delta)^{-\eta} \cdot \alpha^{-1}$  as before).



### A.3 The Effect of Scale

*Proof.* That a higher  $N$  leads to an outward shift in  $\hat{I}$  is apparent from (18), where  $N$  simply increases the constant (a horizontal shift in  $(I, \delta)$  space).

The effect of  $N$  on the slope of the transition path in the region where institutions are improving is seen by noting that the slope in this region is given by:

$$\frac{d\delta}{dI} = \frac{\dot{\delta}(I, \delta)}{\dot{I}(I, \delta)} = \frac{\delta \cdot (1 - \delta) \cdot \phi(-u^c(I, \delta))}{Z(e^*(I, \delta))}.$$

An increase in  $N$  increases  $M$ , which increases  $e^*$ , which increases the denominator without affecting the numerator. The absolute value of the slope is therefore decreasing in  $N$ .  $\square$

## B Empirical

### B.1 Sample

Table 4 describes the core sample - an “X” indicates that the country was surveyed in that wave.

	country	code	1980	1990	1995	2000	2005		country	code	1980	1990	1995	2000	2005
1	Albania	ALB			X	X		37	Luxembourg	LUX				X	
2	Algeria	DZA				X		38	Malaysia	MYS					X
3	Argentina	ARG	X	X	X	X	X	39	Malta	MLT	X	X		X	
4	Australia	AUS	X		X		X	40	Mexico	MEX		X	X	X	X
5	Austria	AUT		X		X		41	Moldova	MDA			X	X	X
6	Bangladesh	BGD			X	X		42	Morocco	MAR				X	X
7	Belgium	BEL	X	X		X		43	Netherlands	NLD	X	X		X	X
8	Brazil	BRA		X	X		X	44	New Zealand	NZL			X		X
9	Bulgaria	BGR		X	X	X	X	45	Nigeria	NGA		X	X	X	
10	Canada	CAN	X	X		X	X	46	Norway	NOR	X	X	X		X
11	Chile	CHL		X	X	X	X	47	Pakistan	PAK			X	X	
12	China	CHN		X	X	X	X	48	Peru	PER			X	X	X
13	Colombia	COL			X	X	X	49	Philippines	PHL			X	X	
14	Croatia	HRV			X	X		50	Poland	POL		X	X	X	X
15	Cyprus	CYP					X	51	Portugal	PRT		X		X	
16	Czech Republic	CZE		X	X	X		52	Romania	ROM		X	X	X	X
17	Denmark	DNK	X	X		X		53	Russia	RUS		X	X	X	X
18	Egypt	EGY				X	X	54	Saudi Arabia	SAU				X	
19	Estonia	EST		X	X	X		55	Singapore	SGP				X	
20	Ethiopia	ETH					X	56	Slovenia	SVN		X	X	X	X
21	Finland	FIN		X	X	X	X	57	South Africa	ZAF		X	X	X	X
22	France	FRA	X	X		X	X	58	Spain	ESP	X	X	X	X	X
23	Germany	GER		X	X	X	X	59	Sweden	SWE	X	X	X	X	X
24	Greece	GRC				X		60	Switzerland	CHE		X	X		X
25	Guatemala	GTM					X	61	Taiwan	TWN			X		X
26	Hong Kong	HKG					X	62	Tanzania	TZA				X	
27	Hungary	HUN	X	X	X	X		63	Thailand	THA					X
28	Iceland	ISL	X	X		X		64	Turkey	TUR		X	X	X	X
29	India	IND		X	X	X	X	65	Uganda	UGA				X	
30	Indonesia	IDN				X	X	66	Ukraine	UKR			X	X	X
31	Ireland	IRL	X	X		X		67	United Kingdom	GBR	X	X	X	X	X
32	Israel	ISR				X		68	United States	USA	X	X	X	X	X
33	Italy	ITA	X	X		X	X	69	Uruguay	URY			X		X
34	Japan	JPN	X	X	X	X	X	70	Venezuela	VEN			X	X	
35	Jordan	JOR				X	X	71	Vietnam	VNM				X	X
36	Lithuania	LTU		X	X	X		72	Zimbabwe	ZWE				X	

Table 4: Sample Description

## References

- ACEMOGLU, D., JOHNSON, S. and ROBINSON, J. A. (2001). The colonial origins of comparative development: An empirical investigation. *The American Economic Review*, **91** (5), 1369–1401.
- , — and — (2002). Reversal of fortune: Geography and institutions in the making of the modern world income distribution. *The Quarterly Journal of Economics*, **117** (4), 1231–1294.
- , — and — (2005). Institutions as a fundamental cause of long-run growth. In P. Aghion and S. N. Durlauf (eds.), *Handbook of Economic Growth*, Amsterdam: North Holland, pp. 385–472.
- AGHION, P., ALESINA, A. and TREBBI, F. (2004). Endogenous political institutions. *The Quarterly Journal of Economics*, **119** (2), 565–611.
- , ALGAN, Y. and CAHUC, P. (2008). *Can Policy Interact with Culture? Minimum Wage and the Quality of Labor Relations*. NBER Working Papers 14327, National Bureau of Economic Research, Inc.
- , —, — and SHLEIFER, A. (2009). *Regulation and Distrust*. NBER Working Papers 14648, National Bureau of Economic Research, Inc.
- ALESINA, A. and ANGELETOS, G.-M. (2005). Corruption, inequality, and fairness. *Journal of Monetary Economics*, **52** (7), 1227–1244.
- , DEVLEESCHAUWER, A., EASTERLY, W., KURLAT, S. and WACZIARG, R. (2003). Fractionalization. *Journal of Economic Growth*, **8** (2), 155–94.
- ALGAN, Y. and CAHUC, P. (2009). *Cultural Change and Economic Development*. mimeo, University of Paris 1.
- BARR, A. (2003). Trust and expected trustworthiness: experimental evidence from zimbabwean villages. *Economic Journal*, **113** (489), 614–630.
- BERG, J., DICKHAUT, J. and McCABE, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, **10** (1), 122–142.
- BERGGREN, N. and JORDAHL, H. (2006). Free to trust: Economic freedom and social capital. *Kyklos*, **59** (2), 141–169.
- BISIN, A., TOPA, G. and VERDIER, T. (2004). Culture as a transmitted cultural trait. *Rationality and Society*, **16** (4), 477–507.
- and VERDIER, T. (1998). On the cultural transmission of preferences for social status. *Journal of Public Economics*, **70** (1), 75–97.
- and — (2001). The economics of cultural transmission and the dynamics of preferences. *Journal of Economic Theory*, **97** (2), 298–319.
- and — (2005). *Work Ethic and Redistribution: A Cultural Transmission Model of the Welfare State*. mimeo, NYU.
- and — (2008). Cultural transmission. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.
- BJØRNSKOV, C. (2006). *The Determinants of Trust*. Ratio Working Papers 86, The Ratio Institute.
- BÉNABOU, R. and TIROLE, J. (2006). Incentives and prosocial behavior. *American Economic Review*, **96** (5), 1652–1678.

- CARLIN, B. I., DOROBANTU, F. and VISWANATHAN, S. (2007). *Public Trust, the Law, and Financial Investment*. Mimeo.
- COOTER, R. D. (2000). Do good laws make good citizens? *Virginia Law Review*, **86** (8), 1577–1601.
- DELHEY, J. and NEWTON, K. (2005). Predicting cross-national levels of social trust: Global pattern or nordic exceptionalism? *Eur Sociol Rev*, **21** (4), 311–327.
- DEMSETZ, H. (1967). Toward a theory of property rights. *The American Economic Review*, **57** (2), 347–359.
- DHILLON, A. and RIGOLINI, J. (2006). *Development and the Interaction of Enforcement Institutions*. The Warwick Economics Research Paper Series (TWERPS) 748, University of Warwick, Department of Economics.
- DIXIT, A. (2004). *Lawlessness and Economics*. Princeton: Princeton University Press.
- (2008). Economic governance. In S. N. Durlauf and L. E. Blume (eds.), *The New Palgrave Dictionary of Economics*, Basingstoke: Palgrave Macmillan.
- FEHR, E. (2008). *On the Economics and Biology of Trust*. IZA Discussion Papers 3895, Institute for the Study of Labor (IZA).
- , FISCHBACHER, U., SCHUPP, J., VON ROSENBLADT, B. and WAGNER, G. G. (2003). *A Nationwide Laboratory Examining Trust and Trustworthiness by Integrating Behavioural Experiments into Representative Surveys*. CEPR Discussion Papers 3858, C.E.P.R. Discussion Papers.
- FRANCOIS, P. and ZABOJNIK, J. (2005). Trust, social capital, and economic development. *Journal of the European Economic Association*, **3** (1), 51–94.
- GLAESER, E. L., LAIBSON, D. I., SCHEINKMAN, J. A. and SOUTTER, C. L. (2000). Measuring trust. *The Quarterly Journal of Economics*, **115** (3), 811–846.
- GRADSTEIN, M. (2004). Governance and growth. *Journal of Development Economics*, **73** (2), 505 – 518.
- GREIF, A. (1994). Cultural beliefs and the organization of society: A historical and theoretical reflection on collectivist and individualist societies. *Journal of Political Economy*, **102** (5), 912–50.
- (2002). Institutions and impersonal exchange: From communal to individual responsibility. *Journal of Institutional and Theoretical Economics (JITE)*, **158** (1), 168–.
- (2006). *Institutions and the Path to the Modern Economy*. Cambridge: Cambridge University Press.
- GUIO, L., SAPIENZA, P. and ZINGALES, L. (2006). Does culture affect economic outcomes? *Journal of Economic Perspectives*, **20** (2), 23–48.
- HAUK, E. and SAEZ-MARTI, M. (2002). On the cultural transmission of corruption. *Journal of Economic Theory*, **107** (2), 311–335.
- HEIFETZ, A., SHANNON, C. and SPIEGEL, Y. (2007). What to maximize if you must. *Journal of Economic Theory*, **133** (1), 31–57.
- HESTON, A., SUMMERS, R. and ATEN, B. (2009). Penn world table version 6.3. Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania.

- HOFF, K. and STIGLITZ, J. E. (2004). After the big bang? obstacles to the emergence of the rule of law in post-communist societies. *American Economic Review*, **94** (3), 753–763.
- JELLAL, M. and WOLFF, F.-C. (2002). Cultural evolutionary altruism: theory and evidence. *European Journal of Political Economy*, **18** (2), 241–262.
- JORDAHL, H. (2007). *Inequality and Trust*. Working Paper Series 715, Research Institute of Industrial Economics.
- KNACK, S. and KEEFER, P. (1997). Does social capital have an economic payoff? a cross-country investigation. *The Quarterly Journal of Economics*, **112** (4), 1251–88.
- and ZAK, P. J. (2002). Building trust: public policy, interpersonal trust, and economic development. *Supreme Court Economic Review*, **10**, 91–107.
- LA PORTA, R., LOPEZ-DE SILANES, F., SHLEIFER, A. and VISHNY, R. W. (1998). Law and finance. *Journal of Political Economy*, **106** (6), 1113–1155.
- LEUKEART, A. (2005). *The Dynamics of Institutional Change: Formal and Informal Institutions and Economic Performance*. mimeo, munich.
- LI, J. S. (2003). Relation based versus rule based governance: An explanation of the east asian miracle and asian crisis. *Review of International Economics*, **11**, 651–73.
- LINDBECK, A. and NYBERG, S. (2006). Raising children to work hard: Altruism, work norms, and social insurance. *The Quarterly Journal of Economics*, **121** (4), 1473–1503.
- MCADAMS, R. H. and RASMUSEN, E. B. (2007). Norms in law and economics. In A. Polinsky and S. Shavell (eds.), *Handbook of Law and Economics*, vol. 2, Amsterdam: Elsevier.
- MULLIGAN, C. B. and SHLEIFER, A. (2005). The extent of the market and the supply of regulation. *The Quarterly Journal of Economics*, **120** (4), 1445–1473.
- NORTH, D. (1990). *Institutions, Institutional Change, and Economic Performance*. Cambridge: Cambridge University Press.
- (2005). *Understanding the Process of Economic Change*. Princeton: Princeton University Press.
- NUNN, N. and WANTCHEKON, L. (2009). *The Slave Trade and the Origins of Mistrust in Africa*. NBER Working Papers 14783, National Bureau of Economic Research, Inc.
- OLCINA, G. and PENARRUBIA, C. (2004). Hold up and intergenerational transmission of preferences. *Journal of Economic Behavior & Organization*, **54** (1), 111–132.
- OLIVIER, J., THOENIG, M. and VERDIER, T. (2008). Globalization and the dynamics of cultural identity. *Journal of International Economics*, **76** (2), 356 – 370.
- SAEZ-MARTI, M. and ZENOU, Y. (2007). *Cultural Transmission and Discrimination*. IEW - Working Papers iewwp348, Institute for Empirical Research in Economics - IEW.
- SAPIENZA, P., TOLDRA, A. and ZINGALES, L. (2007). *Understanding Trust*. NBER Working Papers 13387, National Bureau of Economic Research, Inc.
- STROKOVA, V. (2010). International property rights index 2010 report. Available at <http://www.internationalpropertyrightsindex.org>.

TABELLINI, G. (2008). The scope of cooperation: Values and incentives. *The Quarterly Journal of Economics*, **123** (3), 905–950.

— (2010a). Culture and institutions: Economic development in the regions of Europe. *Journal of the European Economic Association*, (forthcoming).

— (2010b). Presidential address: Institutions and culture. *Journal of the European Economic Association*, **6** (2-3), 255–294.

USLANER, E. (2002). *The Moral Foundations of Trust*. Cambridge: Cambridge University Press.

ZAK, P. J. and KNACK, S. (2001). Trust and growth. *Economic Journal*, **111** (470), 295–321.