TRUST AND WEALTH

by
METİN NEBİLER

Submitted to the Graduate School of Social Sciences in partial fulfillment of the requirements for the degree of Master of Arts

Sabancı University August 2008

TRUST AND WEALTH

METİN NEBİLER

MA THESIS, 2008

Thesis Supervisor: Prof. MEHMET BAÇ

Abstract

This thesis studies the relation between wealth and trust in a model where the ability to elicit trustworthiness from unrelated people depends on own wealth as well as the wealth of other agents in the economy. Betray of trust leads to an enforcement stage.

In equilibrium, rich people trust and betray when matched with a relatively poor agent, while the poor do not trust, and betray if matched with low wealth agents. A constrained equilibrium is characterized in which agents devote all their wealth in the enforcement stage, therefore, wealth changes directly influences the equilibrium resources.

A betray region (combinations of wealth levels) in which all agents are untrustworthy is identified and characterized. Redistribution of wealth concentrated on the betray region is more likely to induce a zero trust equilibrium.

TRUST AND WEALTH

METİN NEBİLER

YÜKSEK LİSANS TEZİ, 2008

Tez Danışmanı: Prof. Dr. MEHMET BAÇ

Özet

Kişisel servet ve dağılımı ile tanımadığımız kişilere olan güven arasındaki ilişki kuramsal açıdan incelenmektedir. Ampirik bulgular böyle bir ilişkinin olduğuna işaret etmekte, ancak iktisat literatüründe bu olguyu açıklayacak teorik modeller henüz geliştirilmemiştir. Güvenin kötüye kullanıldığı durumlarda, verilen sözün yerine getirilip getirilmeyeceğini taraflar arasında oynanan (yasal veya yasal olmayan) bir güç savaşı belirlemektedir.

Oyunun dengesinde, zengin insanların güvenmeye daha yatkın olduğu, fakat karşılarındaki insanların zenginlik seviyesi kendilerine oranla düşük ise, onların güvenini de kötüye kullanmaya daha yatkın oldukları saptanmıştır. Aynı zamanda fakir insanların güvenmedikleri ve eğer karşılarındaki insanlar daha fakir ise kendilerine olan güveni de kötüye kullandıkları saptanmıştır. Kısıtlı diye tabir ettiğimiz dengede insanlar sınırlı zenginlik seviyelerine sahiplerdir ve güç savaşı evresinde bütün zenginliklerini ayırmaktadırlar.

Çalışma, güvenin kötüye kullanıldığı servet kombinasyonlarını belirlemektedir. Sonuç olarak bu tür kombinasyonların yoğun olduğu toplumlarda güvenin yoğun olarak kötüye kullanılacağı, dolayısıyla güven duygusunun da sarsılacağı, öngörüsü ortaya çıkmaktadır.

Contents

List of Figures

Chapter 1

Introduction

Trust is a very important part of people's lives since many transactions among people involve trust. Arrow (1973) states that "there is an element of trust in every transaction". The nature of trust is strange: it bears some natural risks, which is amplified when we trust individuals whom we don't know. This "general" trust in people draws attentions of many branches of social sciences including political science, sociology, and economics. The definition of trust does not vary much among disciplines, however, the general definition can be summarized following Gambetta (1988), as "There is a degree of convergence in the definition of trust which can summarized as follows: trust... is a particular level of the subjective probability with which an agent assesses that another agent or group of agents will perform a particular action... When we say we trust someone... , we implicitly mean that the probability that he will perform an action that is beneficial or at least not detrimental to us is high enough for us to consider engaging in some form of cooperation with him." This definition may also be called a "rational trust" or a "calculative trust" definition. Not all academicians agree on such a definition. Outside the realm of economics, trust is mainly considered a non-calculative decision.

Recent literature on trust has been rapidly growing with the contributions of researchers from different branches of social science. In most of those analysis, researchers try to identify the impacts of trust by using some measures of the trust from surveys. Generally, much of the trust research relies upon the survey question "Generally speaking, would you say that most people can be trusted or that you can't be too careful in dealing with people?" in the General Social Survey(GSS)/World Values Survey (WVS). The survey is aimed to understand the evidence on trust and social capital and also bears interesting results. It is worthwhile to overview the previous works and results in order to understand the trust issue and its determinants and impacts.

Alesina and La Ferrara (2000) used GSS data for the U.S. for the years 1974-1994 and showed that on average, 40 percent of people respond the question in the survey "most people can be trusted" and the trend is generally decreasing from 1970 to 1980 and 1990. Scandinavian countries has the highest trust level while the lowest trust levels are in Latin American countries (La Porta, Lopez-de-Silanes, Shleifer, Vishny (1997)). [1] Glaeser, Laibson, Scheinkman, Soutter (2000) finds that on average, rich and well educated people are more likely to trust, college graduates more 30 percent respond the question as "most people can be trusted" than the high school dropouts while white people are more than 21 percent more likely to say yes to the trust question than blacks.[2]

As Coleman (1990) argues, familiarity between people promotes trust most of research on trust try to investigate the effects of heterogeneity/homogeneity on trust. La Ferrara and Alesina (2000) report effects of heterogeneity variables such as Gini index, race, and ethnicity on trust. They found that trust is negatively and significantly correlated with all three measures of heterogeneity. They also showed that income inequality has a negative impact on trust: "an increase in Gini by one standard deviation decreases the likelihood of trust by 2.5 percentage points". They also reported that the states with the highest trust levels are homogenous in ethnicity and race, also income inequality is very low in those states, on the other hand, states with the lowest level of trust have a fragmented population with a great variation in ethnicity and race and very high income inequality.

In addition to the descriptive statistics and the determinants of trust, it is worthwhile to pay attention to the impacts of the trust on economic and political activities. Slemrod and Katuscak (2002) states that "In high-trust societies, individuals need to spend less resources to protect themselves from being exploited in economic transactions." They study a transaction game and find a connection between trust and trustworthiness: in societies where trustworthiness is low, trust does not pay off, on the other hand, trust pays off more in societies where trustworthiness is higher. They also use WVS data between years of 1990 and 1993 in 18 countries. Their empirical results show that trust increases individuals' income in most countries.[3] Knack and Keefer

[1] La porta et al. used the WVS data for the years 1980-1990.
[2] Glaeser et al. use the GSS data for U.S.
[3] In fact, they suggest that trust increases income if average trustworthiness is high

(1997) used the WVS data of 1981 and 1990 a cross-section of 29 countries to show that people are more likely to innovate and accumulate physical and human capital in the existence of a high level of trust in their country, therefore, the economy grow more rapidly. An extension of Knack and Keefer (1997) is the work of Zak and Knack (2000). They conclude that trust has a positive effect on growth. Uslaner (2002 and forthcoming) identifies the relationship between economic inequality and trust. Uslaner (forthcoming) states that people "Economic equality promotes both optimism and the belief that we all have a shared fate, across races, ethnic groups, and classes." Economic inequality is an important predictor of trust, moreover, decreases trust (Uslaner (2002)). As Gini index increases, trust declines: trust is 21 percent higher in Switzerland than South Africa (Uslaner, forthcoming). In addition to trust's effects on economic activity, several studies in the literature shows that trust has a positive impact on political activity. In political science, it is showed that trust increases the collective action (e.g., Levi (1998); Uslaner (2002)) and political and civic participation (Knack and Keefer (1997)). La Porta, Lopez-de-Silvanes, Shleifer, and Vishny (1999) find that, across countries, a one-standard deviation increase in the measure of trust increases judicial efficiency by 0.7 of a standard deviation and reduces government corruption by 0.3 of a standard deviation. Putnam (1993) suggests that in the existence of higher civic engagement the local governments performs more efficiently.[4]

Trust in people may also be motivated by the formal institutions and organizations. Formal institutions, including judicial system and some civil organizations like bureaus, are ways of enforcing formal transactions like contracts, in a possible case of cheating. As Zak and Knack (2001) argue, a more efficient formal institution promotes trust in society. In addition to formal institutions, sometimes informal activities may also be used in order to force people to perform the expected action instead of betraying. A betrayed person can apply to civil courts for the enforcement, even hire the best lawyer while betraying person has the same opportunity to countervail the enforcement activity. At the same time, as Bac (forthcoming) argues, agents may use private means of enforcement even illegal activities such as paying gangsters to exercise threats. All these enforcement and counter-enforcement activities are closely related to the income levels of individuals since all of them requires resources. If there is a big gap between agents' wealth levels, relatively richer ones have more resources to devote

enough. In their data the only exception is Mexico where average trustworthiness is too low.

[4]Putnam (1993) uses the Italian data based on cross-regions.

and are more powerful in the enforcement stage. Therefore, relative wealth or resource levels determines the enforcement capabilities of agents.

In this paper, we study a model of an economy where agents are homogenous except their wealth levels. Unlike the previous works on trust, individual characteristics such as ethnicity, race, religion, etc. are not included. Wealth is public information. In the economy, agents are randomly matched to form pairs in which one of the agents take the role of beneficiary and the other one becomes trustee, with equal probability. Beneficiary starts the transaction by deciding to trust or not for a future delivery of surplus. Trustee then decides to perform the expected action or betray to grab a proportion of the surplus. If the trustee betrays, the game extends to enforcement game where beneficiary and trustee devote resources to win the enforcement. Agents do have limited resources to devote in the enforcement stage: we classify the agents as unconstrained if they can respond their partners' enforcement strategies and constrained, otherwise. In the case of beneficiary wins the enforcement game, in addition to the resources devoted, the trustee incurs a loss and delivers the surplus he grabbed. However, if trustee wins the enforcement game, he keeps the surplus while the beneficiary loses the resources he devoted and the promised surplus.

The equilibrium of the trust game characterizes the trust and betray strategies of the agents in the economy. The betray and trust decisions are functions of own resource levels and the matched agents' resource levels. Constrained agents are more likely not to trust and be trustworthy while unconstrained agents generally inclined to both betray and trust. Unconstrained agents are more likely to trust since they have enough resources for a strong enforcement while relatively poor, constrained, people are less likely to betray and trust because of the lack of enough resources. There exists a betray region where trustee in a matched pair in that region always betrays. Betray region includes poor beneficiaries and relatively rich trustees. Like the betray region, there also exists a trust region in which in a matched pair, beneficiaries always trust. All unconstrained beneficiaries are in the trust region since they can stand strongly against their partners. These two regions also have an intersection where beneficiaries trust although trustees betray.

The wealth distribution directly determines the trust and betray strategies of agents. If only one agent's wealth level increases, it may change the action of that agent. A continuous increase of a trustworthy agent's wealth level may change his action from performing to betraying, however, it may not be the case all the time. An agent who does not trust eventually changes his action from not trusting to trusting if his wealth increases continuously. Especially, constrained equilibrium resources are very sensitive to wealth changes since it directly changes the equilibrium resources. On the other hand, redistribution of wealth in the economy may change the strategies completely. Specific wealth distributions give clear results about the probability of betray in the economy: A redistributed wealth concentrated on the betray region increases the probability of betray to occur while a wealth redistribution which does not assign any pair in the betray region eliminates the betray in the economy. In a specific wealth redistribution which assigns equal wealth to all agents in an economy, there exists a critical mean wealth where all agents betray if the mean wealth is higher than that critical wealth, performs otherwise. Similarly, a critical wealth level for the trust decision exists: all agents trust if the mean wealth is higher than the critical wealth, does not trust otherwise.

Recent theoretical background on trust shows variety. Many scholars developed new perspectives in order to understand the dynamics of trust. Kreps (1990) presented a trust game which is based on prisoner's dilemma framework. In his one-shot "trust game", he concludes that agents always have an incentive to betray, as a result, their partners not to trust. However, he showed that if game is repeated sufficient times to develop a repetition among players, trust could occur. Similarly, Harvey S. James Jr. (2002) develops the Kreps' work by developing a new model which is an alternative version of principal-agent problem and principals are able to choose to monitor the agents instead of only trusting. One of the conclusions he made is that the principal trust more reliably when he perceives that the agent' temptation to betray is low. Bohnet, Frey, Huck (2001) models a stage game in which player 1 decides to trust or not while player 2 decides to perform or breach. If player 2 breaches the game extends to litigation stage where agents try to win the trial. In all of these models, payoffs are exogenously determined, however, in our model payoffs are determined by the enforcement strategies of agents. Most similar work to our model are Zak and Knack (2001). They study a model where agents are heterogenous by means of income, religion, race, etc. Agents are assigned a wealth and have an income for working in production. In their model, cheating agents punished by formal institutions if they are detected, in addition to that partners are able to inspect the cheating agents by devoting time, called diligence. However, their partners do not have the opportunity to claim that they did not cheat. In Zak and Knack, trust is inversely proportional to diligence; the more diligent the individuals, the less trusting they are. One of their main findings is that wealthy agents are more likely to investigate their partners and protect their wealth. On the other hand, a higher income decreases the diligence since the opportunity cost of diligence is higher (one has to forego current consumption to increase diligence). In addition to that, high income inequality decreases trust while more egalitarian system promotes it.

In this paper, we extend the model used by Bac (forthcoming). The main difference is wealth is public information in our setup. Agents have beliefs about their partners and trust and betray is decided according to their beliefs, however, we allow agents to know the resources of their partners. Therefore, wealth distribution and mean wealth has a great importance in Bac's work. In equilibrium, agents at the lower tail of the wealth distribution are more trustworthy and less likely to trust while wealthy people both trust and betray since they have enough resources for the effective enforcement. He suggests any mean-preserving wealth distribution decreases the wealth inequality may not increase trust since in the existence of very wealthy agents in an economy with low per-capita trust will be positive because of those few wealthy agents and trustworthiness will also be positive since most of agents have few resources for countervailing the enforcement. Homogenizing wealth around its low mean may cause those few wealthy agents to quit trusting. In an economy with sufficiently low per-capita may result with zero-trust since all agents may be too vulnerable if they trust. However, full wealth equality provides two equilibria: full-trust occurs when mean wealth is sufficiently high. On the other hand, in an economy with sufficiently low per-capita there exists zero-trust equilibria.

The thesis is organized as follows. In section 2, we present a model where agents play a trust game. Section 3 provides the equilibrium of the trust and betray strategies. A special case of wealth equality is presented in the Section 4. Section 5 summarizes the results.

Chapter 2

A Model of Trust and Enforcement

Consider a large economy in which people are classified according to their wealth levels. Agents' wealth levels are common knowledge.

The sequence of events is as follows:

- Agents are randomly matched to form pairs. In each pair, one agent takes the position of trustee and the other takes the position of beneficiary. Positions are also randomly determined: each agent is equally likely to assume the role of trustee.

- Agents play a dynamic "trust game" in which first, the beneficiary decides on whether to trust the trustee. The game ends if the beneficiary chooses not to trust.

- If the beneficiary trusts, he incurs a cost $x > 0$. The trustee decides on whether to betray trust or perform. If the trustee performs, the game ends.

- If the trustee betrays, the parties play an enforcement game in which they determine simultaneously the resources, the beneficiary to secure performance, the trustee to avoid performing.

- The binary outcome of the enforcement game is realized and the parties obtain their final payoffs.

In this economy, the maximal individual wealth is denoted W and we assume that the minimum wealth level that one can have is $x > 0$ to ensure that all beneficiaries can trust if they find it in their own interest to do so. The strategies in the overall trust game are defined as follows:

The beneficiary's trust strategy maps $[x, W]^2$ into $[0, 1]$; that is, assigns to each given (beneficiary-trustee) wealth combination (w_B, w_T) a probability of trusting, denoted $t(w_B, w_T)$. The trustee's strategy can be defined similarly. To each given a pair of wealth levels and a binary realization of trust strategy of the beneficiary, the trustee's betray strategy assigns a probability of betrayal, τ. Thus, $\tau : [x, W]^2 \times \{0, 1\} \rightarrow [0, 1]$.

In the enforcement subgame, which is played only if $t = 1$ and $\tau = 1$ is realized, the beneficiary's strategy (called enforcement strategy) assigns to the wealth pair (w_B, w_T) a feasible resource $R_B \in [0, W - x]$. Similarly the trustee's strategy (called counterenforcement strategy) assigns to (w_B, w_T) a feasible resource $R_T \in [2x, W + x]$. These resources are devoted to exercising threats, on monitoring and investigating the most effective way of putting pressure on the trustee perform, or paying third parties to perform these functions. The resources R_B and R_T jointly determine the binary outcome of the enforcement subgame. The beneficiary succeeds, meaning that he is in a position to secure performance, with probability $\frac{R_B}{R_B+R_T}$. With probability $1 - \frac{R_B}{R_B+R_T}$ the trustee "wins" and neutralizes the beneficiary's enforcement activity; in this case, the trustee reaps an extra payoff which we assume is fraction $\alpha \in (0, 1]$ of the net surplus S she owes the beneficiary.

If the beneficiary "wins," then the trustee is forced to perform and beneficiary gets the promised surplus S while trustee incurs a loss $\kappa > 0$, a possible interpretation of which is punishment. The probability that the trustee wins, $\frac{R_T}{R_B+R_T}$, is increasing in the trustee's expenditures on counterenforcement, R_T given R_B, but it is decreasing in of R_B for given R_T constant. The opposite behavior obtains for the probability that the beneficiary wins. We shall assume that if $R_B = R_T = 0$, then each side wins with probability 0.5. Concerning the terminal payoffs, we assume

$$(1 - \alpha)S > \kappa > \alpha S \quad \text{and} \quad S > x.$$

The assumption $\kappa > \alpha S$ rules out the case of ineffective punishment, whereas $(1 - \alpha)S > \kappa$ implies that the surplus is large enough so that the trustee may sometimes betray. An implication of these two assumptions is $\alpha < 0.5$: the maximum benefit from betraying trust is less than the half of the entire surplus. The previous assumption also restricts the punishment. We are only interested with the impacts of marginal changes

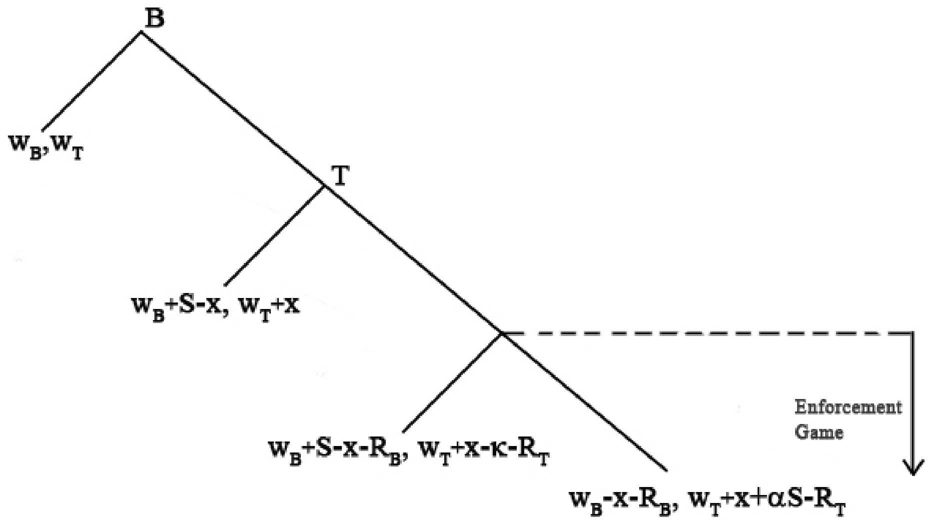

Figure 2.1: Game Tree

in punishment, κ, on results and moreover, a high punishment does not give descriptive results in our framework. On the other hand, $S > x$ is a minimal requirement; it ensures that the beneficiary trusts if the trustee does not betray.

The terminal payoffs are as follows:

If the beneficiary does not trust, $t = 0$, the pair dissolves. Therefore payoffs are equal to wealth.

$$u^N = w_B, \qquad v^N = w_T.$$

When the beneficiary trusts ($t = 1$), the trustee does not betray ($\tau = 0$) and performs, the beneficiary gets the surplus $S > 0$ and the game ends. Thus,

$$u^P = w_B - x + S, \qquad v^P = w_T + x.$$

If, however, the trustee betrays, the parties proceeds to the enforcement subgame, which has two potential outcomes. If the trustee wins terminal payoffs are:

$$u^E = w_B - x - R_B, \qquad v^E = w_T + x + \alpha S - R_T.$$

If the beneficiary wins terminal payoffs are

$$u^E = w_B - x - R_B + S, \qquad v^E = w_T + x - R_T - \kappa.$$

Chapter 3

Equilibrium of the Trust Game

The strategy profile in the trust game $\{(t, R_B), (\tau, R_T)\}$ must form a subgame-perfect Nash equilibrium (SPE). Therefore the enforcement strategies must also form a Nash equilibrium of the enforcement subgame. In this subgame enforcement strategies must be mutually best responses. The best response function of beneficiary is obtained by maximizing the expected payoff,

$$u^M = w_B - x + \frac{R_B}{R_B + R_T} S - R_B,$$

subject to $R_B \leq w_B - x$, given R_T. Ignoring the wealth constraint, the best-response function of the beneficiary, denoted $R_B(R_T)$, must satisfy

$$R_B(R_T) = (R_T S)^{0.5} - R_T, \tag{3.1}$$

Similarly the trustee's best response function, denoted $R_T(R_B)$ maximizes his expected payoff

$$v^M = w_T + x + \frac{R_T}{R_B + R_T} \alpha S - \frac{R_B}{R_B + R_T} \kappa - R_T.$$

subject to $R_T \leq w_T + x$, given R_B. Assuming that the wealth constraint is not binding, the trustee's best-response function is

$$R_T(R_B) = (R_B(\kappa + \alpha S))^{0.5} - R_B. \tag{3.2}$$

These best-response functions are derived under the assumption that the resource constraints, given by $R_B \leq w_B - x$ for the beneficiary and $R_T \leq w_T + x$ for the trustee, are not binding. Obviously, if one of the agents wants to devote more than his net wealth permits, his best choice is to devote entire net wealth on enforcement. We will say that an agent is constrained if he or she cannot afford the best response of a given enforcement, or counter-enforcement level. Incorporating the resource constraints, Then the optimal R_T^* will be given by

$$R_T^* = \begin{cases} R_T(R_B) & \text{if} \quad w_T + x \geq R_T(R_B); \\ w_T + x & \text{if} \quad w_T + x < R_T(R_B). \end{cases} \tag{3.3}$$

The beneficiary's optimal R_B^* is obtained similarly as:

$$R_B^* = \begin{cases} R_B(R_T) & \text{if} \quad w_B - x \geq R_B(R_T); \\ w_B - x & \text{if} \quad w_T - x < R_B(R_T). \end{cases} \tag{3.4}$$

We say that agent i is *constrained* if R_i^* exceeds his available wealth at the enforcement stage. This available wealth is $w_B - x$ for the beneficiary, $w_T + x$ for the trustee. An agent is *unconstrained* otherwise, if R_i^* does not exceed available wealth. Thus, a constrained agent would prefer to devote any additional wealth to enforcement or counterenforcement activities whereas an unconstrained agent's optimal strategy in the enforcement subgame is insentive to small changes in wealth.

Consider a beneficiary in the enforcement game, given a fixed resource R_T which the trustee devotes to counterenforcement. The beneficiary's net expected payoff from marginally increasing his own resource R_B is

$$\frac{du^M}{dR_B} = \frac{R_T S}{(R_B + R_T)^2} - 1.$$

Rearranging this equation, we find that

$$\frac{du^M}{dR_B} \geq 0 \qquad if \qquad R_B \leq (R_T S)^{0.5} - R_T \tag{3.5}$$

The beneficiary is constrained given R_T if $\partial u^M / \partial R_B > 0$ when evaluated at $R_B = w_B - x$; then the optimal enforcement strategy for the beneficiary is $R_B^* = w_B - x$, to devote all of his wealth to enforcement. Similarly, using the net betray payoff of the trustee, v^M for a given R_B, we have

$$\frac{dv^M}{dR_T} = -1 - \frac{(0 - R_B)\kappa}{(R_B + R_T)^2} + \frac{((R_B + R_T) - R_T)\alpha S}{(R_B + R_T)^2},$$

which after arrangements yields the following condition:

$$\frac{dv^M}{dR_T} \geq 0 \qquad if \qquad R_T \leq (R_B(\kappa + \alpha S))^{0.5} - R_B. \tag{3.6}$$

We have a constrained trustee in the enforcement game given R_B, if $\partial v^M / \partial R_T$ evaluated at $R_T = w_T + x$ is strictly positive. Such a trustee will devote all his available wealth to counterenforcement, that is, if $R_T^* = w_T + x$.

The best response functions in (3.1) and (3.2) describe the behavior of the unconstrained levels of R_i^* as a function of R_j. These critical resource levels are not monotonic in R_j; For instance, an increase in R_B increases $R_T(R_B)$ for small amounts of R_B, decreases otherwise. More precisely, from the best response function, we see that

$$\frac{dR_T(R_B)}{dR_B} = \frac{\kappa + \alpha S}{2(R_B(\kappa + \alpha S))^{0.5}} - 1 \geq 0 \quad \text{if and only if} \quad R_B \leq \frac{(\kappa + \alpha S)}{4}. \tag{3.7}$$

On the other hand, we get

$$\frac{dR_B(R_T)}{dR_T} = \frac{S}{2(R_B S)^{0.5}} - 1 \geq 0 \quad \text{if and only if} \quad R_T \leq \frac{S}{4}. \tag{3.8}$$

These properties are displayed in Figure 3.1, where the wealth configurations in which the trustee is constrained are shown as the area lying below the $R_T(R_B)$ schedule. Wealth combinations under which the beneficiary is constrained is located to the left of the $R_B(R_T)$ schedule.

An interesting case is the amount of the exogenous variable x, the trust cost of the beneficiary paid to the trustee. An increase in x, decreases the resources of the beneficiary and increases the resources of the trustee in the enforcement game. Therefore, for higher values of x, beneficiary is more likely to be constrained and trustee is more likely to be unconstrained. However, x does not change the optimal strategies when both agents are unconstrained, unless making the beneficiary constrained.

Before proceeding with the analysis, it is useful to recall the main assumption made so far:

$$(1 - \alpha)S > \kappa > \alpha S \quad and \quad S > x > 0$$

.

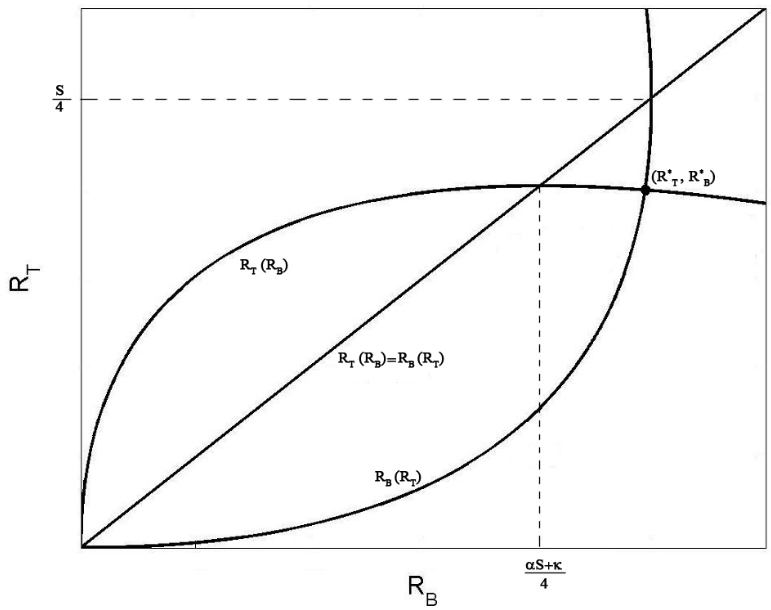

Figure 3.1: Best Response Functions and the Unconstrained Equilibrium

3.1 Analysis of the Enforcement Subgame

Given a matched pair with resources $(w_B^E, w_T^E) \in [x, W]^2$, an equilibrium of the enforcement subgame is a pair of resources $(R_B^*, R_T^*) \in [0, W - x] \times [2x, W + x]$ satisfying (3.3) and (3.4). It is useful to study separately the equilibria according to the status of the wealth constraints, that is, for the case in which the resource constraints are not binding, and the cases in which they are binding only for one or both of the parties. An enforcement subgame equilibrium exists and is unique.

Let us start with the game in which both agents are unconstrained. The Nash Equilibrium pair of resources are obtained as solutions to (3.1) and (3.2)[1]:

$$R_B^* = \frac{(\kappa + \alpha S)S^2}{(\kappa + \alpha S + S)^2} \qquad R_T^* = \frac{(\kappa + \alpha S)^2 S}{(\kappa + \alpha S + S)^2} \tag{3.9}$$

The following result locates the unconstrained equilibrium, displayed in Figure 3.1:

Lemma 1 *(i) The maximal best response $R_T(R_B)$ of the trustee is greater than the Nash equilibrium resource R_T^*. Moreover, $R_T^* < (\kappa + \alpha S)/4$.*
(ii) The maximal best response $R_B(R_T)$ of the beneficiary is greater than Nash equilib-

[1]See the Appendix

rium resource R_B^**. Moreover,* $R_B^* < S/4$*.*

Proof.

(i) The maximum value of $R_T(R_B)$ can be found by solving the equation

$$\frac{dR_T(R_B)}{R_B} = \frac{(\kappa + \alpha S)^{0.5}}{2R_B^{0.5}} - 1 = 0,$$

$$R_B^{0.5} = \frac{(\kappa + \alpha S)^{0.5}}{2} \quad \text{that is, } R_B = (\kappa + \alpha S)/4.$$

Substituting this R_B into the best response function $R_T(R_B)$, we get

$$R_T(R_B) = \frac{(\kappa + \alpha S)^{0.5}(\kappa + \alpha S)^{0.5}}{2} - \frac{(\kappa + \alpha S)}{4} = \frac{(\kappa + \alpha S)}{4}.$$

Using the expression of the Nash equilibrium in $R_T^* < (\kappa + \alpha S)/4$ yields:

$$R_T^* < (\kappa + \alpha S)/4 \Rightarrow \frac{(\kappa + \alpha S)^2 S}{(\kappa + \alpha S + S)^2} < (\kappa + \alpha S)/4,$$

which implies

$$4S(\kappa + \alpha S) < (\kappa + \alpha S + S)^2, \quad \text{or,} \quad 0 < (\kappa + \alpha S - S)^2,$$

which follows from the assumption $S > (\kappa + \alpha S)$.

(ii) As in part (i), the maximum value of $R_B(R_T)$ is given by the following equation

$$\frac{dR_B(R_T)}{R_T} = \frac{S^{0.5}}{2R_T^{0.5}} - 1 = 0$$

$$R_T^{0.5} = \frac{S^{0.5}}{2}$$

$$R_T = S/4$$

The same procedure as in (i) yields

$$R_B(R_T) = \frac{S^{0.5}S^{0.5}}{2} - \frac{S}{4} = \frac{S}{4}.$$

Using the expression of the Nash equilibrium in $R_B^* < S/4$ we get:

$$\frac{(\kappa + \alpha S)S^2}{(\kappa + \alpha S + S)^2} < S/4 \Rightarrow 4S(\kappa + \alpha S) < (\kappa + \alpha S)^2 + 2S(\kappa + \alpha S) + S^2,$$

$$\Rightarrow 0 < (\kappa + \alpha S - S)^2,$$

which holds because $S > (\kappa + \alpha S)$.

Q.E.D.

When one of the two parties is constrained, that is, when R_i^* in (3.9) exceeds agent i's available resources, in the Nash equilibrium this agent will devote all his available resources, to which the unconstrained party will respond optimally according to his best response function. So, when the beneficiary is constrained but the trustee is not, the Nash Equilibrium is given by

$$R_B^* = w_B - x \qquad R_T^* = [(w_B - x)(\kappa + \alpha S)]^{0.5} - w_B + x, \tag{3.10}$$

such that

$$\frac{\partial u^M(R_B^*, R_T^*)}{\partial R_B} > 0 \quad \text{and} \quad \frac{\partial v^M(R_B^*, R_T^*)}{\partial R_T} = 0.$$

Similarly, when the trustee is constrained while the beneficiary is not, the Nash Equilibrium is

$$R_T^* = w_T + x \qquad R_B^* = ((w_T + x)S)^{0.5} - w_T - x \tag{3.11}$$

such that

$$\frac{\partial u^M(R_B^*, R_T^*)}{\partial R_B} = 0 \quad \text{and} \quad \frac{\partial v^M(R_B^*, R_T^*)}{\partial R_T} > 0.$$

Both of these partial derivatives will be strictly positive when both agents are constrained; then the Nash equilibrium is

$$R_T^* = w_T + x \qquad R_B^* = w_B - x. \tag{3.12}$$

The Nash equilibrium of the enforcement subgame (R_B^*, R_T^*) will change as the surplus S or the punishment κ changes, provided the resource constraint is not binding. In the Appendix we show that when both sides are unconstrained, an increase in S leads the beneficiary and the trustee to increase the resources they devote to enforcement and counter-enforcement, that is,

$$\frac{dR_T^*}{dS} > 0, \qquad \frac{dR_B^*}{dS} > 0.$$

An increase in the surplus means a larger resource to compete for. As a result, both the beneficiary and the trustee will increase their Nash equilibrium resources (R_B^*, R_T^*), to raise their chances from the larger surplus.[2] If one of the parties is constrained in the enforcement subgame, a small change in S will have no effect on his enforcement strategy; but the unconstrained party will, as mentioned above, increase R_i^*.

The impact of an increase in the punishment for the betray is different. If κ increases, the trustee will respond by increasing the resource he devotes to counter-enforcement because, if he betrays he faces a larger cost in case the beneficiary wins the enforcement game. Given this change in the trustee's behavior, the beneficiary will also modify his enforcement strategy. He, too, will increase the amount he spends for the enforcement, if the surplus is sufficiently large such that $(1 - \alpha)S > \kappa$.

If the trustee is constrained, a small increase in κ will have no effect on the equilibrium outcome, because the change in the equilibrium resources comes through its effect on the trustee's equilibrium strategy. On the other hand, if it the beneficiary who is constrained while the trustee is not, R_B^* is constant but R_T^* will increase.

3.2 The Equilibrium of the Overall Trust Game

We now probe into the overall trust game, which includes trust and betray decisions plus the potential enforcement subgame. One step before the enforcement subgame is the trustee's betray decision if the beneficiary trusts. The trustee betrays if his corresponding payoff is larger than the sure payoff he gets from performing, $w_T + x$. At this stage the optimal strategy of the trustee is to set $\tau = 1$ and betray if, given the Nash equilibrium $R^* = (R_B^*, R_T^*)$ of the enforcement subgame to which the betray decision leads, the following *net* betray payoff is positive:

$$v(R^*) = -R_T^* + \frac{R_T^*}{R_B^* + R_T^*}\alpha S - \frac{R_B^*}{R_B^* + R_T^*}\kappa \geq 0. \tag{3.13}$$

$v(R^*)$ is decreasing in R_B and, as expected, a larger surplus share αS or a smaller

[2] R_T^* exceeds only if α is sufficiently large, for small values of α, R_T^* may decrease. Also see Lemma 5

punishment κ enhances incentives to betray.

Consider now the beneficiary's trust strategy. The beneficiary's *net* expected payoff from trust is stated below, given the trustee's betray strategy τ and the Nash equilibrium $R^* = (R_B^*, R_T^*)$ of the enforcement subgame to which the betray decision leads.

$$u(R^*, \tau) = -x + \tau(\frac{R_B^*}{R_B^* + R_T^*}S - R_B^*) + (1 - \tau)S. \tag{3.14}$$

The beneficiary's optimal strategy is to set $t = 1$ and trust if the expected payoff in (3.14) is strictly positive. Note that both $v(R^*)$ and $u(R^*, \tau)$ depend on the wealth combination because R^* is a function of $w = (w_B, w_T)$.

A SPE of the overall trust game, denoted $\{t^*, \tau^*, R_B^*, R_T^*\}$, is a collection of trust and betray/perform strategies for each agent, as well as enforcement and counter-enforcement strategies in the enforcement subgame. Lemma 2 describes the trustee's optimal betray-perform response when the beneficiary trusts, according to status of wealth constraints. Note that part (i) of the lemma holds irrespective of whether the trustee is unconstrained or constrained.

Lemma 2 *Suppose that the beneficiary trusts.*
(i) If the beneficiary is unconstrained, the trustee does not betray.
(ii) If the beneficiary is constrained but the trustee is not, the trustee betrays if and only if $w_B - x < 2\kappa + \alpha S - 2(\kappa^2 + \kappa\alpha S)^{0.5} \equiv \bar{R}_B$.

Proof. (i) The proof is divided in two parts, according to whether the trustee is constrained or not. Assume, first, that both parties are unconstrained. Using (R_B^*, R_T^*) defined by (3.9) in (3.13) we get

$$v(R_B^*, R_T^*) = -\frac{S}{S + \kappa + \alpha S}\kappa + \frac{\kappa + \alpha S}{\kappa + \alpha S + S}\alpha S - \frac{S(\kappa + \alpha S)^2}{(\kappa + \alpha S + S)^2}$$

The net betray payoff can be arranged and simplified as follows:

$$v(R_B^*, R_T^*) = \frac{-S(\kappa + \alpha S + S)\kappa + (\kappa + \alpha S)\alpha S(\kappa + \alpha S + S) - S(\kappa + \alpha S)^2}{(\kappa + \alpha S + S)^2} < 0$$

$$-S(\kappa + \alpha S + S)\kappa + (\kappa + \alpha S)\alpha S(\kappa + \alpha S + S) - S(\kappa + \alpha S)^2 < 0$$

$$(\kappa + \alpha S)\alpha S(\kappa + \alpha S + S) < S(\kappa + \alpha S + S)\kappa + S(\kappa + \alpha S)^2$$

$$(\kappa + \alpha S)(\alpha S(\kappa + \alpha S + S) - S\kappa - S(\kappa + \alpha S)) < S^2\kappa.$$

Note that $(\alpha S(\kappa + \alpha S + S) - S\kappa - S(\kappa + \alpha S)) < 0$ because

$$\alpha S\kappa + (\alpha S)^2 + \alpha SS - S\kappa - S\kappa - \alpha SS < 0 \text{ implies } (\alpha - 1)S\kappa + ((\alpha S)^2 - S\kappa) < 0,$$

where the first term is negative by $\alpha < 1$ and the second term is negative by $\kappa > \alpha S$. Thus, $v(R_B^*, R_T^*) < 0$.

Assume, now, that the trustee is constrained by his resource R_T. Since the beneficiary is unconstrained, the Nash equilibrium will lie on beneficiary's best-response function $R_B(R_T)$. Using this fact, the net betray payoff is

$$v(R_B(R_T), R_T) = R_T^{0.5}\frac{(\kappa + \alpha S)}{S^{0.5}} - \kappa - R_T.$$

We show that this net betray payoff is negative even if the trustee could devote his best response to R_B. The value of R_T that maximizes the net betray payoff above satisfies the following first-order condition:

$$\frac{d(v(R_B(R_T), R_T))}{dR_T} = 0.5\frac{\kappa + \alpha S}{R_T^{0.5}S^{0.5}} - 1$$

which yields

$$R_T = \frac{(\kappa + \alpha S)^2}{4S}.$$

Substituting this R_T into the net betray payoff yields:

$$v(R_B(R_T), R_T) = \frac{(\kappa + \alpha S)}{2S^{0.5}}\frac{(\kappa + \alpha S)}{S^{0.5}} - \kappa - \frac{(\kappa + \alpha S)^2}{4S}$$

which is negative if

$$\frac{(\kappa + \alpha S)^2}{2S} - \frac{(\kappa + \alpha S)^2}{4S} - \kappa < 0 \text{ or if } \frac{(\kappa + \alpha S)^2}{4S} - \kappa < 0,$$

which reduces to $(\kappa + \alpha S)^2 < 4S\kappa$. This inequality holds by the assumption $(1 - \alpha)S > \kappa > \alpha S$, which implies $S > \kappa + \alpha S$ and $4\kappa > \kappa + \alpha S$. The proof of part (i) is complete.

(ii) When the beneficiary is constrained but the trustee is not, the former's best counterenforcement response R_B^* is smaller than his available wealth. Using the trustee's unconstrained enforcement best response $R_T^* = (R_B(\kappa + \alpha S))^{0.5} - R_B$, his net betray payoff in (3.13) can be written as:

$$v(R_B, R_T(R_B)) = -2(R_B(\kappa + \alpha S))^{0.5} + R_B + \alpha S \quad \text{for a given } R_B. \tag{3.15}$$

The betray payoff function in (3.15) is the trustee's payoff along his best-response curve. It is decreasing in the beneficiary's resource R_B if

$$1 - 2\frac{(\kappa + \alpha S)^{0.5}}{2R_B^{0.5}} < 0, \quad \text{or if} \quad R_B < \kappa + \alpha S.$$

There exists a critical resource level $\tilde{R}_B = \kappa + \alpha S$ which minimizes the betray payoff of the trustee in (3.15). We note that $\tilde{R}_B$ is larger than the resource R_B^* which the unconstrained beneficiary puts in the unconstrained Nash equilibrium. In fact,

$$\kappa + \alpha S > R_B^* = \frac{(\kappa + \alpha S)S^2}{(\kappa + \alpha S + S)^2} \quad \text{because} \quad S^2 < (\kappa + \alpha S + S)^2.$$

Because the beneficiary is constrained, $R_B = w_B - x < R_B^* < \tilde{R}_B$. It follows that the trustee's betray payoff is declining along his best-response curve as R_B is increased towards the unconstrained Nash equilibrium point.

The betray payoff in (3.15) is strictly positive if $4R_B\kappa < (\alpha S - R_B)^2$, which clearly holds as $R_B \to 0$ and therefore the trustee betrays if the constrained beneficiary's resources are small enough. The betray payoff has two roots, of which the smaller one is relevant:[3]

$$\bar{R}_B = 2\kappa + \alpha S - 2(\kappa^2 + \alpha S\kappa)^{0.5}. \tag{3.16}$$

The betray payoff is positive for $R_B < \bar{R}_B$, negative for larger values of R_B in the relevant range. Thus, the trustee betrays if $R_B = w_B - x < \bar{R}_B$, which yields the condition stated in the lemma.

Q.E.D.

[3]Setting $v(R_B) = 0$, we get $2(R_B(\kappa + \alpha S))^{0.5} = R_B + \alpha S$. Taking the squares of both sides and arranging the terms yields $R_B^2 - R_B(4\kappa + 2\alpha S) + (\alpha S)^2 = 0$. Solving this equation for R_B yields the two roots,

$$R_{B_1} = 2\kappa + \alpha S - 2(\kappa^2 + \kappa\alpha S)^{0.5} \quad \text{and} \quad R_{B_2} = 2\kappa + \alpha S + 2(\kappa^2 + \kappa\alpha S)^{0.5}.$$

See the Appendix

Lemma 2 part (i) requires the assumption $(1-\alpha)S > \kappa > \alpha S$, which is unlikely to hold if α is 0.5 or larger. If this assumption does not hold, then a trustee playing against an unconstrained beneficiary can choose to betray trust; the intuition is that the share he gets from the surplus is sufficiently large, so, even if the trustee is not as resourceful as the beneficiary, he may find it worthwhile to risk betraying trust.

Lemma 2 does not cover the trustee's betray response in the case in which both sides are constrained−shown by the lens-shaped area between the best-response functions in Figure 3.2. We focus below on this case, where $R_T = w_T + x$ and $R_B = w_B - x$ such that the parties' expected marginal payoffs in the enforcement subgame are strictly increasing in w_T and w_B respectively. That is (3.5) and (3.6) evaluated at $R_T = w_T + x$ and $R_B = w_B - x$ hold with strict inequality:

$$w_T + w_B < [(w_T + x)S]^{0.5} \quad \text{and} \quad w_T + w_B < [(w_B - x)(\kappa + \alpha S)]^{0.5} \tag{3.17}$$

where $w_B \geq x$ and $w_T \geq x$. Alternatively, (3.17) can be expressed in terms of the resources available in the enforcement subgame:

$$R_T + R_B < [R_T S]^{0.5} \quad \text{and} \quad R_T + R_B < [R_B(\kappa + \alpha S)]^{0.5}, \tag{3.18}$$

where $R_B \geq 0$ and $R_T \geq 2x$.

When the wealth combination satisfies (3.17)−equivalently, when the resource combination satisfies (3.18)−we have a constrained equilibrium and, if trusted, the trustee's net betray payoff is given by:

$$v(w) = \frac{(w_T + x)\alpha S - (w_T + x)(w_B + w_T) - (w_B - x)\kappa}{w_B + w_T}, \text{ or}$$

$$v(R) = \frac{R_T \alpha S - R_B \kappa}{R_B + R_T} - R_T$$

in terms of the resources. It can be shown that $v(w)$ and $v(R)$ are decreasing in the beneficiary's wealth w_B and resources R_B, respectively.

We define the locus $\rho_B(R_T)$ through the zero-betray-payoff condition $v(R) = 0$, which delivers a resource for the beneficiary such that the trustee is indifferent between betraying and performing given his resources R_T. Thus,

$$\rho_B(R_T) = \begin{cases} \frac{R_T \alpha S - (R_T)^2}{\kappa + R_T} & \text{if} \quad R_T(\bar{R}_B) \geq 2x; \\ \bar{R}_B & \text{if} \quad R_T(\bar{R}_B) < 2x. \end{cases} \tag{3.19}$$

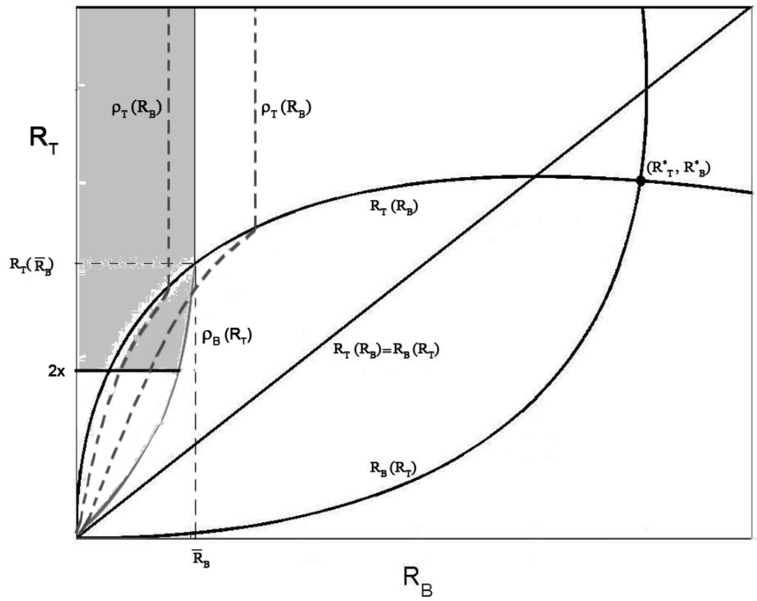

Figure 3.2: Betray and Trust Locus: the betray region is the shaded area which lies left of the betray locus,$\rho_B(R_T)$.

Given his resources R_T, the trustee betrays if the beneficiary's resource R_B is smaller than $\rho_B(R_T)$ and performs otherwise. The locus $\rho_B(R_T)$ is shown in the resource space by the bold convex curve until the trustee's best response curve and the region in which the trustee betrays is given by the shaded area in Figure 3.2 (where the left hand side of (3.19) exceeds the right hand side.) As the resources of the beneficiary increases, the betray payoff decreases and eventually leads to a switch in the trustee's decision from betray to perform.

As we move along the $\rho_B(R_T)$ locus, that is, as R_T approaches $(R_B(\alpha S+\kappa))^{0.5} - R_B$, eventually we hit the intersection of $\rho_B(R_T)$ and the trustee's best response function $R_T(R_B)$.[4] At the intersection, the beneficiary is constrained and is indifferent between

[4]To see this, we can write the condition (3.19) where we let $R_T \to (R_B(\alpha S+\kappa))^{0.5} - R_B$ as

$$R_B = \frac{((R_B(\alpha S+\kappa))^{0.5} - R_B)(\alpha S) - ((R_B(\alpha S+\kappa))^{0.5} - R_B)^2}{\kappa + (R_B(\alpha S+\kappa))^{0.5} - R_B}.$$

Rearranging this condition leads at once to $4(R_B(\kappa+\alpha S)) = (\alpha S + R_B)^2$, which has as solution $R_B = \bar{R}_B$, defined in Lemma 2 as a critical resource of the beneficiary such that the unconstrained trustee is indifferent between betraying and performing.

betraying and performing, while the trustee gives the best unconstrained enforcement response. As shown in Figure 3.2, the resource combinations under which both sides are constrained and the trustee betrays if trusted is a lens-shaped subset. Below we derive a necessary condition for existence of a "betray region" within the set of constrained pairs of resources.

As mentioned, $(R_B, R_T) = (0, 2x)$ is not a resource combination under which both sides are constrained; the beneficiary is constrained but the trustee is not. If R_B is increased while R_T is fixed at $2x$, the net betray payoff falls. The trustee will also become constrained as his best response curve is reached and continues to betray if trusted at low levels of R_B. When a critical beneficiary resource R'_B is reached such that $v(R'_B, 2x) = 0$, the trustee, constrained, becomes indifferent between betraying and performing. A resource $R'_B \in (0, \bar{R}_B)$ satisfying the trustee's indifference condition $R'_B + 2x = [R'_B(\kappa + \alpha S)]^{0.5}$ exists if

$$2x < R_T(\bar{R}_B), \tag{3.20}$$

where $\bar{R}_B$ is defined in the proof of Lemma 2 as the critical beneficiary resource which makes the trustee indifferent between betraying and performing when the trustee is not constrained. Obviously Condition (3.20) holds if x is relatively small.

We now move to the initial stage of the overall game and study the beneficiary's trust decision. The trust game ends if the beneficiary does not trust; in that case his expected net payoff in (3.14) is negative. This constitutes a Nash Equilibrium if in addition to the beneficiary does not trust, trustee decides to betray if trusted. However, this Nash equilibrium is Pareto dominated when beneficiary trusts, therefore, we do not pay attention to this equilibrium. Moreover, the game proceeds to the trustee‘s betray-perform decision. Lemma 3 describes the beneficiary‘s optimal trust decision in all wealth combinations.

Lemma 3 *(i) The unconstrained beneficiary always trusts.*
(ii) If the trustee is unconstrained and betrays whereas the beneficiary is constrained, the beneficiary trusts if and only if $w_B - x > \hat{R}_B$*, where* $\hat{R}_B$ *is the root of the net zero-trust payoff* $u(R^*, \tau) = 0$ *in (3.14), that is,*

$$\hat{R}_B = \frac{S^2 - 2x(\alpha S + \kappa) - (S^2 - 4x(\alpha S + \kappa))^{0.5}}{2(\alpha S + \kappa)}.$$

Proof. (i) It is obvious that in any SPE, the beneficiary trusts if the trustee does not betray. In (3.14), if $\tau = 0$, the expected payoff from trust is positive, so the beneficiary sets $t = 1$. In that case, the beneficiary‘s net trust payoff is $u(R) = S - x$ and since $S > x$, he gets a net positive payoff.

(ii) Assume that the beneficiary is constrained but the trustee is not. Using the trustee's unconstrained enforcement best response $R_T^* = (R_B(\kappa + \alpha S))^{0.5} - R_B$, the beneficiary's net trust payoff in (3.14) can be written as:

$$u(R_B, R_T(R_B)) = R_B^{0.5} \frac{S}{(\alpha S + \kappa)^{0.5}} - R_B - x. \tag{3.21}$$

The payoff in (3.21) is the beneficiary's trust payoff along the trustee's best-response curve. We have

$$\frac{du_M}{dR_B} \geq 0, \quad \text{if} \quad R_B \leq \frac{S^2}{4(\kappa + \alpha S)}.$$

The maximand of the trust payoff (3.21) is unique and denoted $\breve{R}_B = \frac{S^2}{4(\kappa+\alpha S)}$. We note that $\breve{R}_B$ is larger than the resource R_B^* which the unconstrained beneficiary puts in the unconstrained Nash equilibrium. In fact,

$$\frac{S^2}{4(\kappa + \alpha S)} > R_B^* = \frac{(\kappa + \alpha S)S^2}{(\kappa + \alpha S + S)^2} \quad \text{because} \quad S < (\kappa + \alpha S).$$

Because the beneficiary is constrained, $R_B = w_B - x < R_B^* < \breve{R}_B$. It follows that the beneficiary's trust payoff is increasing along trustee‘s best-response curve as R_B is increased towards the unconstrained Nash equilibrium point. On the other hand, the trust payoff in (3.21) is strictly negative if $R_B S^2 \leq (R_B + x)^2(\kappa + \alpha S)$. This condition clearly holds as $R_B \to 0$, therefore, the beneficiary does not trust if his resources are small enough. The trust payoff never turns positive and the beneficiary never trusts if $S^2 < 4x(\kappa + \alpha S)$, given that trust will be betrayed. However, if $S^2 \geq 4x(\kappa + \alpha S)$ the zero-trust payoff has two real roots, of which the smaller one is relevant:[5]

[5]Setting $u(R_B) = 0$, we get $R_B^{0.5} S = (\kappa + \alpha S)^{0.5}(R_B + x)$. Taking the squares of both sides and arranging the terms yields $(\alpha S + \kappa)R_B^2 - (S^2 - 2x(\alpha S + \kappa))R_B + x^2(\alpha S + \kappa) = 0$. Solving this equation for R_B yields the two roots (See the Appendix),

$$R_{B_1} = \frac{S^2 - 2x(\alpha S + \kappa) - ((S^2 - 2x(\alpha S + \kappa))^2 - 4x^2(\alpha S + \kappa)^2)^{0.5}}{2(\alpha S + \kappa)}$$

$$\hat{R}_B = \frac{S^2 - 2x(\alpha S + \kappa) - ((S^4 - 4xS^2(\alpha S + \kappa))^{0.5}}{2(\alpha S + \kappa)}.$$

the trust payoff is negative, hence the beneficiary does not trust, if and only if $R_B = w_B - x < \hat{R}_B$, which yields the condition stated in the lemma.

Q.E.D.

Lemma 3 covers the trust decision of the beneficiary in all wealth combinations except the case in which both agents are constrained. In that region, both agents wealths and, equivalently, their resources, must satisfy (3.17) and (3.18), respectively, where the beneficiary's net trust payoff is (for $\tau = 1$ and assuming the trustee betrays)

$$u(w, 1) = \frac{-x(w_B + w_T) + (w_B - x)S - (w_B - x)(w_B + w_T)}{w_B + w_T}, \text{ or}$$

$$u(R, 1) = \frac{R_B(S - x) - xR_T}{R_B + R_T} - R_B$$

It can be shown that $u(w, 1)$ and $u(R, 1)$ are decreasing in the trustee's wealth w_T and resources R_T, respectively. We define the locus $\rho_T(R_B)$ through $u(R, 1) = u(R, 0) = 0$, to determine the critical trustee resource that makes the beneficiary indifferent between trusting and not trusting for each given resource R_B of the beneficiary:

$$\rho_T(R_B) = \frac{R_B(S - x) - R_B^2}{R_B + x} \tag{3.22}$$

The constrained beneficiary trusts if and only if $R_T < \rho_T(R_B)$. The locus $\rho_T(R_B)$ is shown in Figure 3.2 by the bold concave dashed curve; to its left lies the region in which the beneficiary does not trust. Pick a resource combination in the no trust region and increase the beneficiary's resources, keeping those of the trustee constant. The net trust payoff will increase and eventually turn positive, in which case the beneficiary‘s action switches from not trusting to trust. As we move along the $\rho_T(R_B)$ locus, that is, as R_T approaches $(R_B(\alpha S + \kappa))^{0.5} - R_B$, eventually we hit the intersection of $\rho_T(R_B)$

$$R_{B_2} = \frac{S^2 - 2x(\alpha S + \kappa) + ((S^2 - 2x(\alpha S + \kappa))^2 - 4x^2(\alpha S + \kappa)^2)^{0.5}}{2(\alpha S + \kappa)}$$

and the trustee's best response function $R_T(R_B)$.[6]

The comparison between $\bar{R}_B$ and $\hat{R}_B$ is an important issue but unfortunately analytically difficult. Magnitudes of parameters such as κ, α, x, and S determines the relative positions of these two critical resource levels. In Figure 3.2, there are two dashed lines representing the trust locus. If $\bar{R}_B \geq \hat{R}_B$, then the steeper dashed line represents the trust locus and two loci intersect each other so that there exists a region in which the beneficiary trusts although the trustee betrays. In this case, the intersection of two regions is larger and number of agents who trust if their partners betray is greater. On the other hand, if $\bar{R}_B \leq \hat{R}_B$, flatter dashed line represents the trust locus and we still have such a region; however, the intersection is smaller now.

Until now, we analyzed the overall "Trust Game". First of all, we specified the best response functions $R_B(R_T)$ and $R_T(R_B)$ given by the equations (3.1) and (3.2), respectively. We characterized the equilibrium of the enforcement game $\{R_B^*, R_T^*\}$. Then, we characterized the betray decision of trustee in detail and summarized it in the lemma 2. In a case where the beneficiary is constrained but the trustee is not, if $w_B - x < \bar{R}_B$, the trustee betrays, where $\bar{R}_B$ is given by the equation (3.16). In the case where both agents are constrained, there exists a betray locus $\rho_B(R_T)$ (equation (3.19)) under some conditions. Finally, we characterized the trust decision of beneficiary which is given by lemma 3. Similar to betray locus, we find a trust locus, $\rho_T(R_B)$, (equation (3.22)) when both agents are constrained. Now, we are ready to define the betray and trust decisions of the trustee and the beneficiary, respectively. The first part of the proposition below states the conditions under which the beneficiary trusts; it determines a trust region in which, for all matched pairs in that region, beneficiaries trust. The second part of the proposition deals with the betray decision of the trustee and draws the boundaries of the betray region.

Proposition 1 *(i) The beneficiary trusts if* $R_B \geq R_B(R_T)$ *or* $R_B \geq \rho_B(R_T)$ *or* $R_T \leq$

---

[6]To verify that $\rho_T(R_B) = R_T(R_B)$ as R_T converges to $(R_B(\alpha S + \kappa))^{0.5} - R_B$, write the condition (3.22) where $R_T \to (R_B(\alpha S + \kappa))^{0.5} - R_B$ as

$$(R_B(\alpha S + \kappa))^{0.5} - R_B = \frac{R_B(S - x) - R_B^2}{R_B + x}$$

Rearranging this condition leads at once to $(\alpha S + \kappa)(R_B + x)^2 = R_B S$, which has as solution $R_B = \hat{R}_B$, defined in Lemma 3 as a critical resource of a constrained beneficiary when he is indifferent between trusting and not trusting, matched with a unconstrained trustee.

$\rho_T(R_B)$.
(ii) The trustee betrays if $R_B \leq \rho_B(R_T)$.

Proof.

(i) The beneficiary trusts if the expected net trust payoff is non-negative. The condition $R_B \geq R_B(R_T)$ corresponds to the case where the beneficiary is unconstrained. Lemma 2 states that in such a condition trustee never betrays and first part in lemma 3 suggests that the beneficiary trusts. $R_B \geq \rho_B(R_T) or R_B \geq \rho_T(R_B)$ specifies the conditions that a constrained the beneficiary trusts. The ambiguity results from the fact that the relative positions of $\bar{R}_B$ and $\hat{R}_B$ to each other is not known. If $\bar{R}_B > \hat{R}_B$, $R_B \geq \rho_B(R_T)$ is irrelevant since the beneficiary's expected net trust payoff is non-negative even though the trustee betrays and in that case lemma 3 suggests that the beneficiary trusts. However, $\bar{R}_B < \hat{R}_B$, then the betray locus and trust locus intersect at one point where both agents are constrained. In this case, both equations are relevant. Then, lemmas 2 and 3 suggests that the beneficiary trusts.

(ii) This part is straightforward. The trustee decides to betray if the expected net betray payoff is non-negative. Lemma 2 suggests that the trustee never betrays to an unconstrained beneficiary. The condition $R_B \leq \rho_B(R_T)$ specifies the betray region since in that region the beneficiary is constrained and the expected net betray payoff is non-negative.

Q.E.D.

The analysis above showed that there exists a betray region in which in a matched trustee-beneficiary pair, betray occurs and the game proceeds to the enforcement subgame. If betray occurs in a transaction, it results with loss of resources which reduces the welfare. So, somehow, the size of the betray region displays the welfare loss in a society: if the betray region expands, then the welfare loss is growing. Therefore, the analysis of the size of the betray region gains importance. The size of the betray region can be reduced or enlarged by using the parameters. Below, we study the size of the betray region by using the parameters such as α, S, and κ.

Let us start with the behavior of the best response functions, $(R_B(R_T), R_T(R_B))$, with possible changes in the parameters. The best responses of the beneficiary and the trustee are functions of each other's resources and model parameters. A change in the resource level of one party changes the matched party‘s best response. Also, a change in parameters can create different results in the best response functions of both parties. The following lemma clearly reflects the effects of possible changes of parameters on best response functions.

Lemma 4 *(i)*$R_B(R_T)$ *is increasing in* S *and independent from* κ *and* α.
(ii) $R_T(R_B)$ *is increasing in* S, κ, *and* α.

Proof.

(i) The best response of the beneficiary is given by equation (3.1). By taking the first derivatives of the best response function, we can easily reach some conclusions about its behavior. First of all take the derivative with respect to S.

$$\frac{\partial R_B(R_T)}{\partial S} = \frac{\partial((R_T S)^{0.5} - R_T)}{\partial S}$$

$$\frac{\partial R_B(R_T)}{\partial S} = \frac{R_T^{0.5}}{2S^{0.5}} \geq 0.$$

Now, we take the derivative of $R_B(R_T)$ with respect to α and κ, however, since $R_B(R_T)$ is not functions of α and κ we can immediately see that they are equal to zero.

$$\frac{\partial R_B(R_T)}{\partial S} = \frac{\partial R_B(R_T)}{\partial S} = 0.$$

(ii) The trustees' best response function is given in equation (3.2). The derivatives of the best response function of the trustee can give us some conclusions about the parameters effects. Let us start with taking the derivative with respect to S:

$$\frac{\partial R_T(R_B)}{\partial S} = \frac{\partial((R_B(\kappa + \alpha S))^{0.5} - R_B)}{\partial S}$$

$$\frac{\partial R_T(R_B)}{\partial S} = \frac{R_B^{0.5}\alpha}{2(\kappa + \alpha S)^{0.5}} \geq 0.$$

Now, take the derivative with respect to κ:

$$\frac{\partial R_T(R_B)}{\partial \kappa} = \frac{\partial((R_B(\kappa + \alpha S))^{0.5} - R_B)}{\partial \kappa}$$

$$\frac{\partial R_T(R_B)}{\partial \kappa} = \frac{R_B^{0.5}}{2(\kappa + \alpha S)^{0.5}} \geq 0.$$

Finally, take the derivative with respect to α:

$$\frac{\partial R_T(R_B)}{\partial \alpha} = \frac{\partial((R_B(\kappa + \alpha S))^{0.5} - R_B)}{\partial \alpha}$$

$$\frac{\partial R_T(R_B)}{\partial \alpha} = \frac{R_B^{0.5} S}{2(\kappa + \alpha S)^{0.5}} \geq 0.$$

Q.E.D.

In this lemma, the change in the best response functions corresponds to shifts in best response functions. An increase in the best response functions corresponds to a upward shift for the trustee's best response and a rightward shift for the beneficiary's best response: in a sense that the resources devoted for the enforcement game is increasing in all levels of other agent's resource level.

Now, we are ready to analyze the behavior of the equilibrium $\{R_B^*, R_T^*\}$ when one of the parameters changes. It is expected that an increase in the surplus has a positive effect on both R_B^* and R_T^*. When the surplus increases, both expected net trust payoff and expected net betray payoff are increasing, therefore, the trustee and the beneficiary have stronger incentives to devote more resources in order to win the enforcement game and grab the higher surplus. The following lemma describes the effects of changes in parameters on the unconstrained equilibrium of the enforcement game.

Lemma 5 *(i) R_B^* is increasing in α, κ, and S.*
(ii) R_T^* is increasing in α, κ; however, the effect of surplus is ambiguous.

Proof.

(i) The amount of resources that an unconstrained beneficiary devotes in equilibrium is given by the equation (3.9). The effects of the parameters can be characterized by taking the derivatives of R_B^*. First, we start with the effect of a possible change in S on R_B^*.

$$\frac{\partial R_B^*}{\partial S} = \frac{(2S(\kappa + \alpha S) + S^2\alpha)(\kappa + \alpha S + S)^2 - 2S^2(\kappa + \alpha S)(\kappa + \alpha S + S)(\alpha + 1)}{(\kappa + \alpha S + S)^4} \geq 0$$

$$(\kappa + \alpha S + S)((2S(\kappa + \alpha S) + S^2\alpha)(\kappa + \alpha S + S) - 2S^2(\kappa + \alpha S)(\alpha + 1)) \geq 0$$

$$(2S\kappa + 3\alpha S^2)(\kappa + \alpha S + S) \geq (2S^2\kappa + 2\alpha S^3)(\alpha + 1)$$

$$2S\kappa^2 + 2\kappa\alpha S^2 + 2\kappa S^2 + 3\alpha\kappa S^2 + 3\alpha^2 S^3 + 3\alpha S^3 \geq 2\alpha\kappa S^2 + 2\alpha^2 S^3 + 2\kappa S^2 + 2\alpha S^3$$

$$3\kappa\alpha S^2 + 2S\kappa^2 + \alpha^2 S^3 + \alpha S^3 \geq 0.$$

Now, take the derivative with respect to κ

$$\frac{\partial R_B^*}{\partial \kappa} = \frac{S^2(\kappa + \alpha S + S)^2 - 2S^2(\kappa + \alpha S)(\kappa + \alpha S + S)}{(\kappa + \alpha S + S)^4} \geq 0$$

$$S^2(\kappa + \alpha S + S)((\kappa + \alpha S + S) - 2(\kappa + \alpha S)) \geq 0$$

$$(\kappa + \alpha S + S) - 2(\kappa + \alpha S) = S - (\kappa + \alpha S) \geq 0 \Rightarrow S \geq \kappa + \alpha S.$$

Last, take the derivative with respect to α

$$\frac{\partial R_B^*}{\partial \alpha} = \frac{S^3(\kappa + \alpha S + S)^2 - 2S^3(\kappa + \alpha S)(\kappa + \alpha S + S)}{(\kappa + \alpha S + S)^4}$$

$$S^3(\kappa + \alpha S + S)((\kappa + \alpha S + S) - 2(\kappa + \alpha S)) \geq 0$$

$$(\kappa + \alpha S + S) - 2(\kappa + \alpha S) = S - (\kappa + \alpha S) \geq 0 \Rightarrow S \geq \kappa + \alpha S.$$

(ii) Similarly, an unconstrained trustee devotes R_T^*, given by equation (3.9). Take the derivative with respect to α in order find its effect on R_T^*.

$$\frac{\partial R_T^*}{\partial \alpha} = \frac{2S^2(\kappa + \alpha S)(\kappa + \alpha S + S)^2 - 2S^2(\kappa + \alpha S)^2(\kappa + \alpha S + S)}{(\kappa + \alpha S + S)^4} \geq 0$$

$$2S^2(\kappa + \alpha S)(\kappa + \alpha S + S)((\kappa + \alpha S + S) - (\kappa + \alpha S)) \geq 0$$

$$(\kappa + \alpha S + S) - (\kappa + \alpha S) \geq 0 \Rightarrow S > 0.$$

Now, take the derivative with respect to κ

$$\frac{\partial R_T^*}{\partial \kappa} = \frac{2S(\kappa + \alpha S)(\kappa + \alpha S + S)^2 - 2S(\kappa + \alpha S)^2(\kappa + \alpha S + S)}{(\kappa + \alpha S + S)^4} \geq 0$$

$$2S(\kappa + \alpha S)(\kappa + \alpha S + S)((\kappa + \alpha S + S) - (\kappa + \alpha S)) \geq 0$$

$$(\kappa + \alpha S + S) - (\kappa + \alpha S) \geq 0 \Rightarrow S > 0.$$

Finally, we look at the effect of surplus on R_T^*. Take the derivative with respect to S

$$\frac{\partial R_T^*}{\partial S} = \frac{(2\alpha S(\kappa + \alpha S) + (\kappa + \alpha S)^2)(\kappa + \alpha S + S)^2 - 2S(\alpha + 1)(\kappa + \alpha S)^2(\kappa + \alpha S + S)}{(\kappa + \alpha S + S)^4} \geq 0$$

$$(\kappa + \alpha S)(\kappa + \alpha S + S)((\kappa + 3\alpha S)(\kappa + \alpha S + S) - 2S(\alpha + 1)(\kappa + \alpha S)) \geq 0$$

$$(\kappa + 3\alpha S)(\kappa + \alpha S + S) \geq 2S(\alpha + 1)(\kappa + \alpha S))$$

$$\kappa^2 + \kappa S + 4\alpha\kappa S + 3(\alpha S)^2 + 3\alpha S^2 \geq 2\alpha\kappa S + 2(\alpha S)^2 + 2\kappa S + 2\alpha S^2$$

$$\kappa^2 + 2\alpha\kappa S + (\alpha S)^2 + \alpha S^2 \geq S\kappa. \tag{3.23}$$

Equation (3.23) does not hold in every case, therefore, the result of S on R_T^* is ambiguous. However, it is obvious that (3.23) does not hold as $\alpha \to 0$. The reason of a change in the unconstrained equilibrium of the enforcement game is the shifts of the best response functions. When S increases both of the best response functions shifts as lemma 4 states. Therefore, the magnitude of the shifts of best response functions determine the levels of R_B^* and R_T^*. If α is small enough, the shift of the beneficiary's best response function exceeds the trustee's best response function, which results in an increase in R_B^* but a decrease in R_T^*. In higher α values, this effect disappears and an increase in S increases both R_B^* and R_T^*.

Q.E.D.

Let us state the intuition behind the previous lemma. As we mentioned, one can expect that an increase in surplus, S, results in agents devoting more resources in the enforcement phase since both agents expected net payoffs increase. However, while this is the case for R_B^*, R_T^* decreases with an increase in S at small values of α. Similarly, an increase in α directly increases the expected net betray payoff of the trustee because the betraying trustee can grab a larger portion from the surplus. Lemma 4 states that if α increases, the best response function of the trustee shifts upward while the beneficiary's best response function is constant. Therefore, the trustee devotes more resources in every level, and eventually we have a new unconstrained equilibrium where both agents devote higher resources. Perhaps, the most surprising result of the previous lemma is the effect of κ on the unconstrained equilibrium. When κ increases, the punishment imposed on the betraying trustee if he loses the enforcement game increases. As a result, betraying trustees should devote higher resources to win the enforcement game and avoid paying the higher punishment. The trustee's best response function shifts upward if κ increases, that is, the trustee devotes more resources in every level. Even though the beneficiary's best response stays constant, the shift in trustee's best response results with a new unconstrained equilibrium where both R_B^* and R_T^* are higher.

The other important variable in determining the size of the betray region is $\bar{R}_B$ since it determines the critical R_B level: a constrained beneficiary whose resources are smaller than $\bar{R}_B$ is betrayed by an unconstrained trustee. Also, as we mentioned, if we move along the $\rho_B(R_T)$ locus, that is, as R_T approaches $(R_B(\alpha S + \kappa))^{0.5} - R_B$, eventually we hit the intersection of $\rho_B(R_T)$ and the trustee's best response function $R_T(R_B)$: this is the resource level $\bar{R}_B$. Therefore, for an accurate analysis of the betray region, $\bar{R}_B$ plays an important role: keeping everything constant a right shift of $\bar{R}_B$, shifts the betray locus and increases the betray region. The following lemma studies the behavior of $\bar{R}_B$ for possible changes in the model parameters.

Lemma 6 *$\bar{R}_B$ increases in α and S and decreases in κ.*

Proof.

Let us start with the effect of α on $\bar{R}_B$. Taking the derivative with respect to α yields:

$$\frac{\partial \bar{R}_B}{\partial \alpha} = \frac{\partial(2\kappa + \alpha S - 2(\kappa^2 + \kappa\alpha S)^{0.5})}{\partial \alpha}$$

$$\frac{\partial \bar{R}_B}{\partial \alpha} = S - \frac{2\kappa S}{2(\kappa^2 + \kappa\alpha S)^{0.5}} \geq 0$$

$$2S(\kappa^2 + \kappa\alpha S)^{0.5} \geq 2\kappa S \Rightarrow 4S^2\kappa^2 + 4S^2\kappa\alpha S \geq 4S^2\kappa^2 \Rightarrow 4S^2\kappa\alpha S \geq 0$$

Now, take the derivative with respect to κ

$$\frac{\partial \bar{R}_B}{\partial \kappa} = \frac{\partial(2\kappa + \alpha S - 2(\kappa^2 + \kappa\alpha S)^{0.5})}{\partial \kappa}$$

$$\frac{\partial \bar{R}_B}{\partial \kappa} = 2 - \frac{2(2\kappa + \alpha S)}{2(\kappa^2 + \kappa\alpha S)^{0.5}} \leq 0$$

$$4(\kappa^2 + \kappa\alpha S)^{0.5} \leq 2(2\kappa + \alpha S) \Rightarrow 16\kappa^2 + 16\kappa\alpha S \leq 16\kappa^2 + 16\kappa\alpha S + 4(\alpha S)^2$$

$$\Rightarrow 0 \leq 4(\alpha S)^2$$

Finally, take the derivative with respect to S

$$\frac{\partial \bar{R}_B}{\partial S} = \frac{\partial(2\kappa + \alpha S - 2(\kappa^2 + \kappa\alpha S)^{0.5})}{\partial S}$$

$$\frac{\partial \bar{R}_B}{\partial S} = \alpha - \frac{2\alpha\kappa}{2(\kappa^2 + \kappa\alpha S)^{0.5}} \geq 0$$

$$2\alpha(\kappa^2 + \kappa\alpha S)^{0.5} \geq 2\alpha\kappa \Rightarrow 4\alpha^2\kappa^2 + 4\alpha^3\kappa S \geq 4\alpha^2\kappa^2 \Rightarrow 4\alpha^3\kappa S \geq 0$$

Q.E.D.

The previous lemma summarizes the effects of the changes in parameters when the beneficiary is constrained. The lemma 2 states that the trustee never betrays to an unconstrained beneficiary. However, when the beneficiary is constrained, there exists cases where the trustee betrays and $\bar{R}_B$ is the critical point where an unconstrained trustee is indifferent between betray and perform. Also, as stated above, the betray locus covers the case where both agents are constrained; this locus hits the critical level $\bar{R}_B$ as the trustee becomes unconstrained. Therefore, a right shift of $\bar{R}_B$ enlarges the betray locus and increases the number of constrained[7] and unconstrained trustees that betray to constrained beneficiaries.

Previous lemma states that an increase in S and α shifts $\bar{R}_B$ to the right. An increase in both parameters directly increases the expected net betray payoff of the trustee. Therefore, the trustee may change his action depending on his resources, from perform to betray. However, if κ increases, then the trustee's expected net betray payoff declines and the trustee may change his action from betray to perform.

Until now, we analyzed the effects of parameters on some relevant variables and best responses for an accurate characterization of the betray region. The size of the betray region can be reduced or enlarged by using the parameters. The previous three lemmas express the effects of changes of parameters. We can use these lemmas in order to adjust the size of the betray region. The following proposition summarizes the effects of parameters on the betray region.

Proposition 2 *The betray region enlarges with S and α and shrinks with κ.*

Proof.

The betray region is shown in Figure 3.2 by the shaded area. When α or S increases, lemma 4 states that the trustee's best response function shifts upward and lemma 6 states that $\bar{R}_B$ shifts rightward. A rightward shift of the $\bar{R}_B$ also shifts the betray locus rightward. Therefore, betray region enlarges.

An increase in κ shifts the trustee's best response function upward while $\bar{R}_B$ shifts leftward. The leftward shift of the $\bar{R}_B$ shifts the betray locus leftward. Eventually, the betray region shrinks.

Q.E.D.

In the betray region, the trustee has a non-negative expected net betray payoff, hence, in every matched pairs whose resources lie in that region, the trustee betrays if the beneficiary trusts. However, some wealth combinations in that region are out of consideration because the trustee is never constrained for resource levels below $2x$. An increase in both α or S increases the expected net betray payoff of trustees and more of the trustees want to betray and get hold of the higher proportion of the surplus

[7]Increases only if "betray region" within the set of constrained pairs of resources exists, in other words, Equation (3.20) should hold.

or higher surplus. This enlarges the betray region. On the other hand, an increase in punishment, κ, decreases the number of betraying trustees since their expected net betray payoff decreases, therefore, the betray region shrinks.

In addition to the betray region, there exist a trust region in our framework. The trust region includes the complement of the betray region because beneficiary certainly trusts if trustee does not betray as stated in lemma 2. However, sometimes it may be possible that a beneficiary trusts even trustee betrays. Hence, there may exist a intersection of both betray and trust regions where the beneficiary trusts even if the trustee betrays. Therefore, the trust region can be defined as the sum of complement of the betray region and the intersection of both regions. However, we have some difficulties in achieving a clear characterization of the trust region and obtaining comparative statics results regarding the effects of the parameters (κ, α, and S) on the trust region.

First of all, unlike the betray region, the size of the trust region is ambiguous since the relative positions of betray and trust locus is unknown. Figure 3.2 shows betray and trust loci together, but two dashed lines represent the trust locus since we do not determine the relative positions of both locus. The ambiguity results from failure in reaching a clear comparison between $\bar{R}$ and $\hat{R}$. Different magnitudes of the parameters change the relative positions of both loci, which makes the analysis very difficult. Also, we can not determine the effects of parameters on the trust locus, hence, the trust region. The characterization of the changes in the trust locus due to changes in parameters is very difficult since calculations do not give clear results. In addition to that, $\hat{R}_B$, which is the critical wealth level determining the trust decision of the agents, does not exist in every case. The condition $S^2 > 4x(\kappa + \alpha S)$ must hold for existence of a trust locus. These difficulties cause an ambiguity in the behavior of the trust region in model parameters, so, we cannot explain the changes of the trust region when any of the parameters change.

Chapter 4

Effects of Wealth Distribution

The wealth distribution in the economy is an important determinant of betray and trust regions. Specification of the wealth distribution is not necessary but it gives clear results about the betray and trust regions. On the other hand, individual wealth changes may cause agents' actions to change. A potential increase in an agents' wealth makes the agent more powerful in the enforcement game with higher resources available. Even a small increase of wealth of an agent at the margin, may change the action profile of that agent such as from not trusting to trusting or from performing to betraying. Similarly, a reduction in wealth may have the opposite impact on agents' actions.

As Bac (forthcoming) proposed, the trust and betray decisions are strongly related with the relative enforcement power and wealth. Therefore, wealth distribution and relative wealth levels of matched agents are the determinant of trust and betray in the economy. A specific wealth redistribution may have drastic changes in the trust and betray decisions such that we may have a full-trust equilibrium or even zero-trust equilibrium. In the constrained equilibrium, agents devote all their resources in order to win the enforcement stage. Hence, if one of the agents' wealth changes slightly, the equilibrium resources change which will affect the probabilities of being successful in the enforcement stage. The betray region can be seen in the Figure 3.2 with the shaded area. A specific wealth distribution that is concentrated on the betray region increases the likeliness to betray in that economy while similarly, wealth distributions concentrated on the trust region increases the likeliness to trust. In the extreme case, suppose that the wealth distribution is completely concentrated in the trust region. In this case all agents who take the role of beneficiary trusts and we have full-trust in the economy. Similarly, a specific wealth distribution may assign all pairs to the region where no agent trusts. Hence, we have a zero-trust equilibrium.

Uniform wealth distribution may be very expressive and give clear results. In this case, the betray region is the area of that shaded region in Figure 3.2, while the trust region is the area of the complement of that region plus the intersection of both regions. ¿From a different perspective, we can also define betray probability as the likeliness of betray to occur in all transactions in the society. Hence, the betray probability is the shaded area over the all area. The betray probability may have welfare implications: in an economy high betray probability corresponds to higher numbers of pairs to move to the enforcement stage and higher amounts of resources devoted in the enforcement stage which will result with a decrease in welfare. Clearly, a specific wealth distribution concentrated on the betray region increases the betray probability in transactions, therefore, results with welfare losses. On the contrary, a specific wealth distribution concentrated on the trust region increases the welfare by decreasing the losses in the enforcement stage.

A special case of wealth equality is studied since it may have nice implications to analyze an economy where wealth is equally distributed. The next section studies the trust game and implications in the case of equally distributed wealth.

4.1 The Case of Wealth Equality

We now consider an economy in which wealth is distributed perfectly equal: all agents in the economy have same wealth levels. Similarly, this corresponds to the case that two agents with same wealth levels match in order to complete a transaction. In this perfectly egalitarian economy, per-capita wealth is also the mean wealth in the economy. Therefore, per-capita wealth is very important since it completely determines the trust and trustworthiness in the economy. A wealth distribution depending on the level of per-capita wealth may give very different results.

In this economy, the only difference among agents is disappeared, and agents are perfectly identical. However, the enforcement game is still asymmetric which is resulted from the initial amount x that the beneficiary delivers to the trustee as the cost of trust. In a low per-capita economy both agents have small amounts of resources to devote in the enforcement game. At low wealth levels, when beneficiary trusts, he has to incur x to the trustee, moreover, if trustee betrays, the enforcement resources of the beneficiary decreases while the counterenforcement resources of the trustee increases.

Since the trustee has an extra amount of x, the trustee is more advantageous in the enforcement game. In other words, the probability that the trustee to be successful is higher in low wealth levels. Therefore, in a low per-capita economy, the trustee is more likely to be untrustworthy and attempt to grab a fraction of the surplus while the beneficiary is less likely to trust. However, as the mean wealth in the economy increases, the asymmetry that results from x is diminishing and the probability of trustee being successful in the enforcement stage is decreasing while the probability of beneficiary to win the enforcement increases. Hence, in higher wealths, the trustee is more likely to be trustworthy and the beneficiary is more likely to trust.

In a perfectly egalitarian economy, when two agents with equal wealth levels match and play the trust game, the resources they devote in the enforcement game is given by the line in Figure 4.1: $R_T = R_B + 2x$. The case in which both agents are constrained with equal wealth levels occurs if following condition holds: [1]

$$\kappa + \alpha S > 16x \tag{4.1}$$

If the above condition does not hold, the available wealth combinations reveals an equilibrium where beneficiary is always constrained and trustee is unconstrained. Therefore, the analysis of both cases in condition 4.1 is necessary in order to obtain clear results.

Let's start with the first case in which condition 4.1 holds. In this case, we have a constrained equilibrium such that both agents are constrained and have same wealth levels. As it is seen in the Figure 4.1, at very low per-capita wealth, because of the asymmetry results from the delivery of x makes the trustee unconstrained and the beneficiary constrained. Therefore, for that levels of mean wealth the second part of the lemma 2 applies which states that the trustee betrays. For higher levels of mean wealth, beneficiary stays as unconstrained, however, after a critical resource level,

[1]The intersection of $R_T = R_B + 2x$ and $R_T(R_B) = (R_B(\kappa + \alpha S))^{0.5} - R_B$ is found by simply putting R_T into the best response function. By arranging the equation we get $2(R_B + x) = (R_B(\kappa + \alpha S))^{0.5}$. Taking the squares of the both sides and set it equal to zero we have $4R_B^2 + R_B(8x - (\kappa + \alpha S)) + 4x^2 = 0$. Solving this equation with respect to R_B, the roots of the equation are

$$R_{B_1} = \frac{(\kappa + \alpha S) - 8x - ((8x - (\kappa + \alpha S))^2 - 64x^2)^{0.5}}{8}$$

. Inside of the squareroot is positive only when $\kappa + \alpha S > 16x$

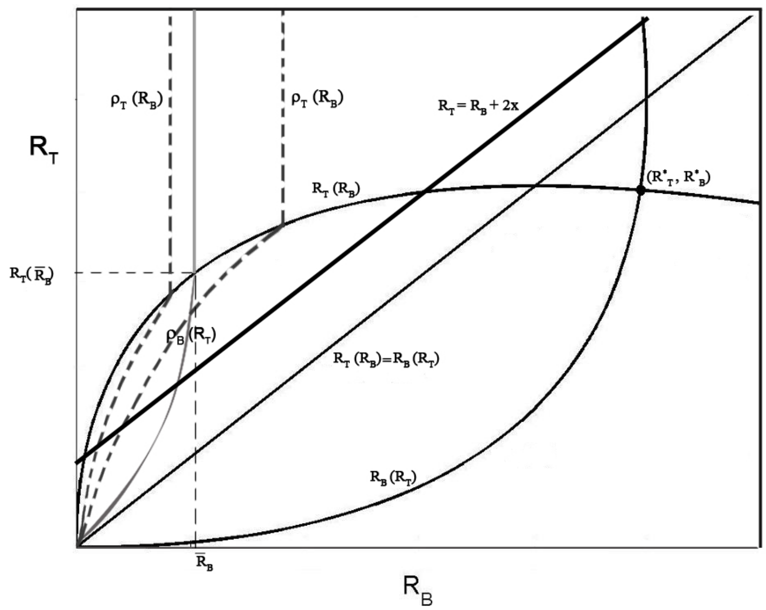

Figure 4.1: If condition 4.1 holds $R_T = R_B + 2x$ intersects with the trustee's best response function, hence, we have pairs in which agents have same wealth levels and are constrained

trustee becomes constrained too. In this case, both agents are constrained and devote all their resources in a potential enforcement game. Therefore, there exists a critical per-capita wealth level,$\hat{w}$, which gives trustee zero *net* expected betray payoff. It is also shown in Figure 4.1, at the intersection of the betray locus, $\rho_B(R_T)$, and the $R_T = R_B + 2x$ line in lens-shaped area, trustee has zero *net* expected betray payoff. Therefore, there exist a mean wealth level, $\hat{w}$ in a perfectly egalitarian economy where wealth is equally distributed: for lower per-capita wealth levels than critical wealth $\hat{w}$, all trustees are untrustworthy, on the other hand, if the per-capita wealth is higher than $\hat{w}$, all trustees are trustworthy and we have a full trust equilibrium.

Similar to betray decision, the trust decision of beneficiaries in a perfectly egalitarian economy has important implications. Beneficiaries do not trust in a low per-capita economy since they do not have enough resources to be successful in the enforcement game, however, we have a full trust equilibrium if per-capita wealth is sufficiently high. There exists a critical mean wealth, $\bar{w}$, which makes the beneficiaries' indifferent between trusting and not trusting. Hence, in a perfectly egalitarian economy where per-capita wealth is lower than the $\bar{w}$, beneficiaries do not trust, however, if per-capita wealth is higher than the $\bar{w}$, we have a full trust equilibrium. As Figure 4.1 shows, the critical resource level which makes the beneficiary indifferent in the trust decision is the intersection of $R_T = R_B + 2x$ line and the trust locus. The following lemma covers the trust and betray decisions of beneficiary and trustee, respectively, when condition (4.1) holds.

Lemma 7 *(i) In an economy where wealth is equally distributed, suppose beneficiary trusts, trustee betrays if* $\hat{w} \leq \frac{1}{4}((\alpha S - \kappa - 2x) + \sqrt{(\alpha S - \kappa - 2x)^2 + 8x(\alpha S + \kappa)})$.

(ii)In a perfectly egalitarian economy, the beneficiary trusts if $\bar{w} \geq (S - (S^2 - 8Sx)^{0.5})/4$.

Proof.

(i)Note that as $w \longrightarrow 0$ $v(w) > 0$

My claim is $\frac{d}{dw}v(w) \leq 0$, $v(w)$ is decreasing with w, eventually becomes negative,

$$\frac{d}{dw}v(w) = -4w + \alpha S - \kappa - 2x \leq 0.$$

Now, find the roots

$$v(R_B, R_T) = -R_T + \frac{R_T}{R_B + R_T}\alpha S - \frac{R_B}{R_B + R_T}\kappa$$

$$v(w) = \frac{w + x}{2w}\alpha S - w - x - \frac{w - x}{2w}\kappa$$

$$v(w) = w\alpha S + x\alpha S - 2w^2 - 2wx - w\kappa + x\kappa$$

$$v(w) = -2w^2 + w(\alpha S - \kappa - 2x) + x(\alpha S + \kappa)$$

The greater root is relevant[2]

$$\hat{w} = \frac{1}{4}((\alpha S - \kappa - 2x) + \sqrt{(\alpha S - \kappa - 2x)^2 + 8x(\alpha S + \kappa)}).$$

I need to show $\hat{w} \geq 0$,

$$\frac{1}{4}((\alpha S - \kappa - 2x) + \sqrt{(\alpha S - \kappa - 2x)^2 + 8x(\alpha S + \kappa)}) \geq 0$$

$$(\alpha S - \kappa - 2x)^2 \leq 0(\alpha S - \kappa - 2x)^2 + 8x(\alpha S + \kappa)$$

$$8x(\alpha S + \kappa) \geq 0.$$

[2]See the Appendix

(ii)Note that as $w \to 0$ $v(w) < 0$

I need to show that $\frac{du(w)}{dw} \geq 0$ in order to show $u(w)$ is increasing with w and becomes positive eventually,

$$\frac{du(w)}{dw} = Sx/2w^2 - 1 \geq 0 \quad if \quad Sx > 2w^2$$

$$u(R_B, R_T) = -x + \frac{R_B S}{R_B + R_T} - R_B$$

replacing $R_B = w - x$ and $R_T = w + x$,

$$u(w) = -x + \frac{(w-x)S}{w - x + w + x} - (w - x)$$

Setting $u(w) = 0$ and arranging the equation, we get

$$2w^2 - Sw + Sx = 0$$

Solving the equation, the smaller one is relevant[3].

$$\bar{w} = \frac{S - (S^2 - 8Sx)^{0.5}}{4}$$

I need to show $\bar{w} \geq 0$,

$$(S - (S^2 - 8Sx)^{0.5})/4 \geq 0 \Rightarrow S^2 \geq S^2 - 8Sx \Rightarrow 8Sx > 0.$$

Now, Sx should be greater than $\bar{w}$, in order to show that $u(w)$ is increasing with x and become positive,

$$Sx \geq 2((S - (S^2 - 8xS)^{0.5})/4)^2 \Rightarrow 8Sx \geq S^2 - 2S(S^2 - 8xS)^{0.5} + S^2 - 8xS$$

$$2S(8x + (S^2 - 8xS)^{0.5}) \geq 2S^2 \Rightarrow 8x + (S^2 - 8xS)^{0.5} \geq S$$

[3]See the Appendix

Take the square of both sides

$$64x^2 + S^2 - 8xS + 16x(S^2 - 8xS)^{0.5} \geq S^2 \Rightarrow 8x + 2(S^2 - 8xS)^{0.5} \geq S$$

Again take the square of both sides we get

$$64x^2 + 32x(S^2 - 8xS)^{0.5} + 4S^2 - 32Sx \geq S^2 \Rightarrow Sx > \bar{w}$$

Since condition (4.1) holds $S > \kappa + \alpha S > 16x$

Q.E.D.

The previous lemma covers the betray and trust decisions of constrained agents when there exists a region in which all agents have same wealth levels. In an economy where wealth is equally distributed among agents, per-capita wealth level is very deterministic. The relative positions of $\hat{w}$ and $\bar{w}$ is ambiguous since magnitudes of α, κ, x, and S determines their relative positions. In a sufficiently low per-capita economy, both trust and trustworthiness is zero: trustees betray while beneficiaries may never trust since trustees are more advantageous in the enforcement stage while beneficiaries do not have enough resources to win the enforcement.[4] On the other hand, in an economy with a sufficiently high per-capita wealth level we may have a full trust equilibrium where beneficiaries trust and trustees never betray. [5]

The second case is that condition (4.1) does not hold: the equal wealth level line, $R_T = R_B + 2x$, never intersects with the trustee‘s best response function $R_T(R_B)$, as in Figure 4.2. In this case, matched agents‘ resource combination (R_B, R_T) never exists in the constrained region. Therefore, there does not exist an equilibrium in which both agents are constrained, instead, trust game has an equilibrium in which trustee is unconstrained but beneficiary is constrained. The following lemma summarizes the result in the case that condition (4.1) does not hold: pairs formed by agents with equal wealth levels does not exist in the lens-shaped area.

Lemma 8 *(i) The trustee betrays if $R_B \leq \bar{R}_B$, performs otherwise.*

(ii)The beneficiary trusts if $R_B \geq \hat{R}_B$, does not trust otherwise.

[4]Let denote the per-capita with w_p. If w_p is smaller than both $\hat{w}$ and $\bar{w}$, a zero trust equilibrium exists where beneficiaries do not trust while all trustees betray. However, if $\hat{w} > w_p > \bar{w}$, although all trustees betray, all beneficiaries trust

[5]A full trust equilibrium exists only when $w_p > \bar{w}$ regardless of the magnitude of the $\hat{w}$. If $\hat{w} > w_p > \bar{w}$ in a full trust equilibrium all trustees betray, however if $w_p > \bar{w} > \hat{w}$ in the full trust equilibrium all trustees are trustworthy.

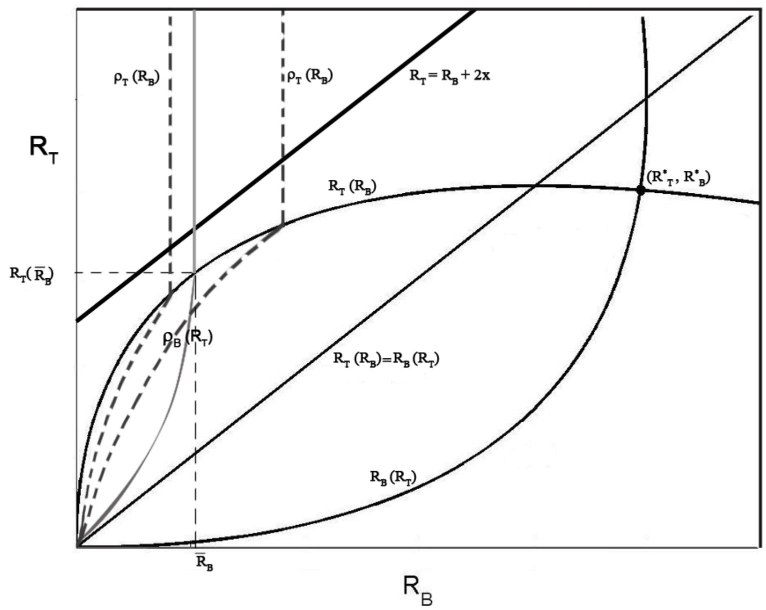

Figure 4.2: If condition (4.1) does not hold, the $R_T = R_B + 2x$ never intersects with the trustee's best response.

Proof.

(i)If the condition (4.1) does not hold, then we have a case in which the beneficiary is constrained and the trustee is unconstrained. In Lemma 2 it is stated that the trustee betrays if $R_B < \bar{R}_B$ and performs otherwise. Lemma 2 perfectly applies here.

(ii)If the condition (4.1) does not hold, then we have a case in which the beneficiary is constrained and the trustee is unconstrained. In Lemma 3 it is stated that the beneficiary does not trust if $R_B < \hat{R}_B$ and does not trust, otherwise. Hence, Lemma 3 perfectly applies here.

Q.E.D.

The previous lemma covers the betray and trust decisions in an economy where wealth is equally distributed but at least one of the agents is unconstrained. In this case, pairs are formed by constrained beneficiaries and unconstrained trustees. Therefore, the previous results explain the behaviors of both agents. A sufficiently low per-capita wealth in the economy results with zero trust equilibrium where all trustees betray.[6]

[6]If w_p is smaller than both $\hat{R}_B$ and $\bar{R}_B$, a zero trust equilibrium exists where beneficiaries do not trust while all trustees betray. However, if $\bar{R}_B > w_p > \hat{R}_B$, although all trustees betray, all beneficiaries trust

However, in an economy with sufficiently high per-capita wealth full trust equilibrium exists in the economy where trustees never betray.[7]

[7]A full trust equilibrium exists only when $w_p > \hat{R}_B$ regardless of the magnitude of the $\bar{R}_B$. If $\bar{R}_B > w_p > \hat{R}_B$ in a full trust equilibrium all trustees betray, however if $w_p > \hat{R}_B > \bar{R}_B$ in the full trust equilibrium all trustees are trustworthy.

Chapter 5

Conclusion

The relationship between wealth and trust is analyzed in our theoretical model where enforcement game is introduced if betray occurs. We demonstrate that agents' wealth and relative wealth of their partners are the main determinants of trust and trustworthiness since agents trust or are trustworthy when they believe they have enough resources to enforce broken promises. We showed that in the constrained equilibrium, agents devote all their resources in order to be successful in the enforcement stage. In equilibrium, wealthy agents trust and are trustworthy if their partners are wealthy, on the other hand, they are untrustworthy if their partners are sufficiently poor. Incentives to be untrustworthy is positively related with the surplus and own wealth, however, negatively related with the punishment and partners' wealth. Similarly, trusting incentives increases with an increase in own wealth and decreases with partners' wealth. A betray region region (combinations of wealth) is identified in which agents are always untrustworthy. The number of the untrustworthy agents increases with surplus, on the other hand, decreases if punishment is higher.

The trust and trustworthiness decisions are strongly related with the wealth distribution. Individual wealth increases may change the trust and trustworthiness decisions. Interestingly, all positive wealth changes may not change the actions of agents even being the wealthiest in the economy. Wealth redistributions may induce drastic changes in the overall trust and trustworthiness. A sufficiently concentrated wealth distribution in the betray region results with no trustworthiness and less trust. However, if wealth is concentrated in the trust region, as expected, full-trust equilibrium is reached in the economy. A wealth redistribution assigns the same wealth to all agents in the economy may result with two equilibria; in a sufficiently low per-capita economy there is zero trust while a sufficiently high mean wealth results with full-trust.

Although this paper brings a new perspective to the trust issue, more advanced models should be developed for examining the determinants and the impacts of trust on economic issues. First of all, our model suggest income as the only heterogeneity factor in the society, however, it is much more complex in reality. In addition to wealth differentiation among people, race, ethnicity, education level, etc. can be embedded into the model as another heterogeneity factors in the society. A measure in the light of these variables between $(0, 1)$ could be used to classify the people and agents with measure slightly higher than 0 can be classified as the poor agents such as black, uneducated agents, and agents with higher measure could be assigned as the relatively rich agents. Hence, effects of wealth and other variables on trust and trustworthiness could be analyzed in this framework.

Another extension could be the addition of judicial efficiency into the model. Judicial efficiency is very important in the enforcement stage, since we expect a higher probability of success if judicial system works well in an economy. The probability of being successful in the enforcement stage can be redefined as the probability we used multiplied with a variable δ which represents the judicial efficiency and takes values in the interval $(0, 1)$. If judicial system works perfectly, this corresponds to $\delta = 1$ and judicial system does not work well may correspond to $0.5 > \delta > 0$.

Bibliography

[1] Alesina, A. and E. La Ferrara., The Determinants of Trust., NBER Working Paper No. 7621, March 2000.

[2] Arrow, Kenneth J. Information and Economic Behavior. Stockholm: Federation of Swedish Industries. 1973.

[3] Bac, M. "Generalized Trust and Wealth." International Review of Law and Economics forthcoming.

[4] Bohnet, I., Frey, B. S. and Huck, S. "More Order with Less Law: On Contract Enforcement, Trust, and Crowding." American Political Science Review 95 (2001): 131-144.

[5] Coleman, James S. Foundations of Social Theory. Cambridge, MA: Harvard University Press, 1990.

[6] Gambetta D. Trust: Making and Breaking Cooperative Relations. Oxford: Basil Blakwell, 1988.

[7] Glaeser, Edward, David I. Laibson, Jose A. Scheinkman, and Christine L. Soutter. "Measuring Trust." Quarterly Journal of Economics 115(3) (2000): 811-846.

[8] James, Harvey S. "On the Reliability of Trusting". Rationality and Soceity 14(2) (2002): 229256.

[9] Katuscak, P. and E. Slemrod. "Do Trust And Trustworthiness Pay Off?" NBER Working Paper No. 9200, September 2002.

[10] Knack, Stephen and Paul Zak. "Trust and Growth." Economic Journal 111(470) (2001): 295-321.

[11] Knack, Stephen and Philip Keefer. "Does social capital have an economic payoff?: A cross-country investigation." The Quarterly Journal of Economics, 112(4) (1997): 1251-1288.

[12] Kreps, David M., "Corporate Culture and Economic Theory." Perspectives in Positive Political Economy. Ed. J. Alt and K. Shepsle. New York: Cambridge University Press, 1990.

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

[14] La Porta R., F. Lopez de Silanes, A. Shleifer and R. Vishny. "The Quality of Government" Journal of Law and Economics 15(1) 1999, 222-279

[15] Levi, M. "A state of trust". Trust and Governance. Ed. Levi, M. and Braithwaite. New York: Russel Sage Foundation, 1998.

[16] Putnam, R. Making Democracy Work: Civic Traditions in Modern Italy. Princeton: Princeton University Press, 1993.

[17] Uslaner, E.M. The Moral Foundations of Trust. Cambridge. UK: Cambridge University Press, 2002.

[18] Uslaner, E.M. Corruption, Inequality, and the Rule of Law: The Bulging Pocket Makes the Easy Life. Cambridge. UK: Cambridge University Press, 2008.

Appendix

Derivation of equation (1)

$$\frac{d}{dR_B}u^M = -1 + \frac{((R_B + R_T) - R_B)S}{(R_B + R_T)^2}$$

$$((R_B + R_T) - R_B)S = (R_B + R_T)^2$$

$$R_T S = (R_B + R_T)^2$$

$$R_B = (R_T S)^{0.5} - R_T.$$

Derivation of equation (2)

$$\frac{d}{dR_B}v^M = -1 - \frac{(0 - R_B)\kappa}{(R_B + R_T)^2} + \frac{((R_B + R_T) - R_T)\alpha S}{(R_B + R_T)^2} = 0$$

$$R_B(\kappa + \alpha S) = (R_B + R_T)^2$$

$$R_T = (R_B(\kappa + \alpha S))^{o.5} - R_B$$

Derivation of the Equilibrium: Equation (7)

Best responses imply that $R_T S = (R_B + R_T)^2$ and $R_B(\kappa + \alpha S) = (R_B + R_T)^2$ so by combining these two equations, we get

$$\frac{R_B}{R_T} = \frac{S}{\kappa + \alpha S}$$

then we can write R_B in terms of R_T as follows

$$R_B = \frac{S R_T}{\kappa + \alpha S}$$

then by putting this into the equation (1) we get

$$\frac{R_T S}{\alpha S + \kappa} = (R_T S)^{0.5} - R_T$$

$$(R_T S) = (\alpha S + \kappa)(R_T S)^{0.5} - (\alpha S + \kappa)R_T$$

$$R_T(\alpha S + \kappa + S) = (\alpha S + \kappa)(S)^{0.5} R_T^{0.5}$$

$$R_T^{0.5} = \frac{(\alpha S + \kappa)(S)^{0.5}}{\alpha S + \kappa + S}$$

$$R_T^* = \frac{(\alpha S + \kappa)^2 (S)}{(\alpha S + \kappa + S)^2}$$

Then putting this into the equation we can find the

$$R_B^* = \frac{S}{\alpha S + \kappa} R_T^*$$

$$R_B^* = \frac{(\alpha S + \kappa)(S)^2}{(\alpha S + \kappa + S)^2}$$

Relevance of $\bar{R}_B$

The R_{B_1} is relevant since the other root is greater than the unconstrained equilibrium.

First, we show that $R_{B_1} > 0$

$$2\kappa + \alpha S - 2(\kappa^2 + \kappa \alpha S)^{0.5} > 0 \Rightarrow 2\kappa + \alpha S > 2(\kappa^2 + \kappa \alpha S)^{0.5}$$

Take the square of both sides then

$$4\kappa^2 + 4\kappa \alpha S + (\alpha S)^2 > 4\kappa^2 + 4\kappa \alpha S$$

Since $(\alpha S)^2 > 0$.

Now, I show that

$R_{B_2} > R_B^* \Rightarrow 2\kappa + \alpha S + 2(\kappa^2 + \kappa \alpha S)^{0.5} > \frac{(\alpha S + \kappa)(S)^2}{(\alpha S + \kappa + S)^2}$

$$(\alpha S + \kappa + S)^2 (2\kappa + \alpha S + 2(\kappa^2 + \kappa \alpha S)^{0.5}) > (\alpha S + \kappa)(S)^2$$

This holds since $(\alpha S + \kappa + S)^2 (2\kappa + \alpha S) > (\alpha S + \kappa)(S)^2$

Relevance of $\hat{R}_B$

The R_{B_1} is relevant since the other root is greater than the unconstrained equilibrium.

First, we show that $R_{B_1} > 0$

$$R_{B_1} = \frac{S^2 - 2x(\alpha S + \kappa) - ((S^2 - 2x(\alpha S + \kappa))^2 - 4x^2(\alpha S + \kappa)^2)^{0.5}}{2(\alpha S + \kappa)} > 0$$

$$S^2 - 2x(\alpha S + \kappa) > ((S^2 - 2x(\alpha S + \kappa))^2 - 4x^2(\alpha S + \kappa)^2)^{0.5}$$

Take the square of both sides we obtain

$$S^4 - 4xS^2(\alpha S+\kappa) + 4x^2(\alpha S+\kappa)^2 > S^4 - 4xS^2(\alpha S+\kappa)$$

Since $4x^2(\alpha S+\kappa)^2 > 0$

Now, I show that

$R_{B_2} > R_B^* \Rightarrow \frac{S^2-2x(\alpha S+\kappa)+((S^2-2x(\alpha S+\kappa))^2-4x^2(\alpha S+\kappa)^2)^{0.5}}{2(\alpha S+\kappa)} > \frac{(\alpha S+\kappa)(S)^2}{(\alpha S+\kappa+S)^2}$

$$(\alpha S+\kappa+S)^2(S^2-2x(\alpha S+\kappa)+(S^4-4xS^2(\alpha S+\kappa))^{0.5}) > (\alpha S+\kappa)S^2 2(\alpha S+\kappa)$$

Eliminate the $(\alpha S+\kappa+S)^2(S^4-4xS^2(\alpha S+\kappa))^{0.5}$ term. Then, we have

$$((\kappa+\alpha S)^2+S^2+2S(\kappa+\alpha S))(S^2) > (\kappa+\alpha S)(S^2(\kappa+\alpha S)+2x(\kappa+\alpha S+S)^2)$$

Then eliminate the $S^2(\kappa+\alpha S)$ term. We obtain

$$S^3(2(\kappa+\alpha S)+S) > (\kappa+\alpha S)(S^2(\kappa+\alpha S)+2x(\kappa+\alpha S+S)^2)$$

$$S^3(\kappa+\alpha S)+S^3(\kappa+\alpha S+S)) > S^2(\kappa+\alpha S)^2+2x(\alpha S+\kappa)(\kappa+\alpha S+S)^2)$$

$S^3(\kappa+\alpha S) > S^2(\kappa+\alpha S)^2$ since $S > \kappa+\alpha S$

And $S^3(\kappa+\alpha S+S)) > 2x(\alpha S+\kappa)(\kappa+\alpha S+S)^2)$ since $S^2 > 4x(\kappa+\alpha S)$

Relevance of $\hat{w}$

The smaller root is irrelevant since it is less than zero:

$$\hat{w} = \frac{1}{4}((\alpha S-\kappa-2x) - \sqrt{(\alpha S-\kappa-2x)^2+8x(\alpha S+\kappa)}) < 0$$

$$(\alpha S-\kappa-2x) < \sqrt{(\alpha S-\kappa-2x)^2+8x(\alpha S+\kappa)}$$

This obviously holds, since the left side of the equation is negative while the right side is positive.

Relevance of $\bar{w}$

The net trust payoff is increases if the following condition hold

$$Sx > 2w^2$$

The above equation holds for the smaller root. However, it does not hold for the greater root which reveals that the net trust payoff is decreasing if $Sx < 2w^2$, and eventually, reaches zero at the higher root $\bar{w}_2 = \frac{S+(S^2-8Sx)^{0.5}}{4}$. Now, I show that $Sx < 2\bar{w}^2$:

$$2\bar{w}_2^2 = 2(\frac{S + (S^2 - 8Sx)^{0.5}}{4})^2 > Sx$$

$$S^2 + S^2 - 8Sx + 2S(s^2 - 8Sx)^{0.5} > 8Sx$$

$$2S^2 + 2S(s^2 - 8Sx)^{0.5} > 16Sx$$

$2S^2 > 16Sx$, since condition 4.1 holds $S > \kappa + \alpha S > 16x$