NBER WORKING PAPER SERIES

THE POLITICAL COASE THEOREM:
EXPERIMENTAL EVIDENCE

Sebastian Galiani
Gustavo Torrens
Maria Lucia Yanguas

Working Paper 19943
http://www.nber.org/papers/w19943

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2014

We thank the Weidenbaum Center at Washington University in St. Louis for its financial support, Universidad de San Andrés in Argentina for providing a laboratory to conduct the experiment, and the Department of Economics at the University of Zurich for allowing us to use z-tree. We would specially like to thank Daron Acemoglu, James Walker, Arlington Williams, and Juan Camilo Cardenas for very insightful comments. Brian Feld provided excellent research assistance. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Political Coase Theorem: Experimental Evidence
Sebastian Galiani, Gustavo Torrens, and Maria Lucia Yanguas
NBER Working Paper No. 19943
February 2014
JEL No. C92,D72

ABSTRACT

The Political Coase Theorem (PCT) states that, in the absence of transaction costs, agents should agree to implement efficient policies regardless of the distribution of bargaining power among them. This paper uses a laboratory experiment to explore how commitment problems undermine the validity of the PCT. Overall, the results support theoretical predictions. In particular, commitment issues matter, and the existence of more commitment possibilities leads to better social outcomes. Moreover, we find that the link is valid when commitment possibilities are asymmetrically distributed between players and even when a redistribution of political power is required to take advantage of those possibilities. However, we also find that at low levels of commitment there is more cooperation than strictly predicted by our parameterized model while the opposite is true at high levels of commitment, and only large improvements in commitment opportunities have a significant effect on the social surplus, while small changes do not.

Sebastian Galiani
Department of Economics
University of Maryland
3105 Tydings Hall
College Park, MD 20742
and NBER
galiani@econ.umd.edu

Gustavo Torrens
Department of Economics
Indiana University
Wylie Hall, 100 S Woodland Ave
Bloomington, IN 47405-7104
gtorrens@indiana.edu

Maria Lucia Yanguas
Department of Economics
UCLA
Los Angeles
myanguas@ucla.edu

The Political Coase Theorem: Experimental Evidence*

Sebastian Galiani† University of Maryland
Gustavo Torrens‡ Indiana University
Maria Lucia Yanguas§ UCLA

First Version: February, 2013 This Version: February, 2014

Abstract

The Political Coase Theorem (PCT) states that, in the absence of transaction costs, agents should agree to implement efficient policies regardless of the distribution of bargaining power among them. This paper uses a laboratory experiment to explore how commitment problems undermine the validity of the PCT. Overall, the results support theoretical predictions. In particular, commitment issues matter, and the existence of more commitment possibilities leads to better social outcomes. Moreover, we find that the link is valid when commitment possibilities are asymmetrically distributed between players and even when a redistribution of political power is required to take advantage of those possibilities. However, we also find that at low levels of commitment there is more cooperation than strictly predicted by our parameterized model while the opposite is true at high levels of commitment, and only large improvements in commitment opportunities have a significant effect on the social surplus, while small changes do not.

JEL Classification Codes: D72 C92

Key Words: Political Coase Theorem, Limited Commitment, Laboratory Experiment.

1 Introduction

In the context of a legal dispute associated with an externality, the Coase Theorem (Coase 1960) states that if there are no transaction costs and legal rights are well-specified, the allocation of resources will be efficient, whatever the allocation of legal rights and the bargaining power of the parties involved, which only affect the division of the surplus. In principle, an analogous result also applies to politics. The Political Coase Theorem (hereinafter referred to as the PCT), states that, in the absence of transaction costs, political players should agree to implement efficient policies regardless of the distribution of political power among them, which should only affect how gains are distributed. However, inefficient policies are pervasive, which suggests that the extension of the Coase Theorem to politics might not be so simple. In fact, one key implicit assumption of the Coase Theorem does not probably hold in politics. The Coase

*We thank the Weidenbaum Center at Washington University in St. Louis for its financial support, Universidad de San Andrés in Argentina for providing a laboratory to conduct the experiment, and the Department of Economics at the University of Zurich for allowing us to use z-tree. We would specially like to thank Daron Acemoglu, James Walker, Arlington Williams, and Juan Camilo Cardenas for very insightful comments. Brian Feld provided excellent research assistance.

†E-mail: galiani@econ.umd.edu
‡E-Mail: gtorrens@indiana.edu
§E-Mail: myanguas@ucla.edu

Theorem assumes that agents can commit to their agreements and, without this assumption, the theorem breaks down as a result of time inconsistency and only inefficient outcomes emerge. For a standard legal dispute, this may not be a serious problem because agents can always rely on binding agreements enforced by a third party (e.g. a contract enforced by the courts). Politics, however, involves bargaining among powerful players, who, by definition, can renege agreements. As a consequence, powerful players face a commitment problem which may restrict the scope of agreements they can reach with other players. In this paper, we explore how commitment problems affect the PCT employing a randomized laboratory experiment.

A historical example could help to clarify the crucial role of commitment problems in politics. North and Weingast (1989) study the British Glorious Revolution as an institutional change that solves a crucial commitment problem. Before the revolution the Parliament was not willing to allow new taxes because the king had considerable discretionary power over expenditures. Nor the king could borrow money because he could not credibly commit to repay his debts. As a consequence the Crown used very inefficient sources of revenue such as the sales of monopoly licenses. The revolution opened the era of parliamentary supremacy. The king could not dissolve the Parliament anymore, the Parliament significantly increased its power to control and monitor expenditures, royal prerogative powers were reduced and subordinated to common law, and judges stopped serving at the king's pleasure. Finally, the successful dethroning of Charles I and James II established a credible threat against future monarchs, but the Crown continued playing an important role, working as a balance of power to the Parliament.

The importance of the PCT cannot be overemphasized. On the substantive front, there is hardly a more relevant issue in the social sciences than the identification of the sources of inefficient policies. From a theoretical perspective, most formal political economy models now simply assume that the PCT does not apply. This is the end result of a shift in the literature on institutions and institutional change away from a tacit acceptance of the PCT. Early works in institutional economics suggested that institutional changes were efficient adjustments in response to innovations, implicitly accepting the PCT. Conversely, the identification of the specific transaction costs that block efficient outcomes has been a paramount issue in new institutional economics (North, 1981, 1990). However, only Acemoglu (2003) presents a formal political economy model in which social conflict and limited commitment are the key factors that undermine the validity of the PCT. His model is an infinitely repeated taxation game between a ruler and a citizen. Under no commitment the unique sub-game perfect Nash equilibrium of the stage game leads to a very inefficient outcome. The citizen does not work hard because he knows that the ruler will tax all his income, and the ruler does not have any way to credibly commit herself to refraining from taxing the citizen's income. When the ruler can credibly commit to limit taxation or the citizen can credibly commit to pay a transfer if the ruler resigns, the efficient outcome is restored. Note, however, that the later case, requires an institutional change (the ruler must resign). Even when neither the ruler nor the citizen are able to commit repeated interactions open the door to limited commitment and, hence, to better social outcomes. In other words, repeated interactions make some promises credible.

To test the validity of the PCT, we begin by adapting the model developed by Acemoglu (2003) to a simplified version suitable to be tested in a laboratory environment.[1] Our adaptation keeps three main

[1]Even though field experiments are becoming more common, both in economics and political science (see, among others, Druckman, Green, Kuklinski and Lupia, 2011; and Gerber and Green 2012), the range of causes that researchers can manipulate experimentally is still limited. Thus, laboratory experiments remain the gold standard for research into the causal relationships existing among a broad set of issues (Friedman and Sunder, 1994). This is particularly the case when

features in Acemoglu (2003). First, agents' ability to commit is limited. Second, commitment opportunities could be asymmetrically distributed among players. Third, exploiting commitment opportunities could require an institutional change. Our adaptation is simpler in the sense that we consider a two-period game in which promises are only partially binding. Specifically, there is some probability that a player must keep her promises and some probability that promises are not binding at all. Clearly, by changing these probabilities we are inducing different levels of commitment. This approach to induce limited commitment is now standard in political economy (see for example, Acemoglu and Robinson 2005). The intuition that justifies it is that there are two sources of political power: de-jure political power, which emanates from legal institutions and it is more sustainable, and de-facto political power, which emanates from the ability to change legal institutions and it is only temporary. For example, consider an autocrat (de-jure power) that faces a popular revolt (de-facto power) with some probability. To placate a revolt he can commit to some temporary reforms, but in the future if the revolt calms down (de-facto power is only temporary), he will be tempted to dismantle the reforms. As a consequence, the autocrat's promises during a revolt are only partially credible (he will renege his promises as soon as the revolt has dissipated). More importantly for this paper, our simple two-period game with partially binding promises is able to capture the key insights on commitment problems and the PCT. In particular, we can handle a situation in which nobody can commit; only one player has the ability to partially commit; and only one player has the ability to commit, but an institutional change is required.

We conducted the experiment between August and November 2012 in a computer laboratory at Universidad de San Andrés, Argentina. Participants were graduate and undergraduate students who had differing fields of study and differing degrees of familiarity with game theory. In all, we conducted 10 sessions with 16 subjects each, for a total of 160 participants, who were first randomly assigned to one of two roles - player 1 (the ruler) or player 2 (the citizen)- and then to four different treatments. Each treatment differs only in terms of commitment opportunities: in Treatment 1, neither player 1 nor player 2 can credibly commit (promises are not binding at all); in Treatment 2, player 1 has a slight commitment opportunity (with a probability equal to 0.25 that her promises are binding), while player 2 has none; in Treatment 3, player 1 has significant commitment opportunities (with a probability equal to 0.75 that her promises are binding), while player 2 has none; and, last, in Treatment 4, player 1 has no commitment opportunities while player 2 has full credibility (her promises are binding). The first three treatments increasingly offer more commitment opportunities to player 1, who initially detents the power to tax. The last treatment offers more commitment opportunities to player 2, who initially does not have the power to tax. Thus, in Treatment 4 players must engineer a redistribution in their relative power (an institutional change) in order to exploit commitment opportunities.

Overall, the results of this experiment support the hypothesis that commitment issues matter and that the existence of more commitment opportunities leads, on average, to better social outcomes. Indeed, we find that this link is valid even when a reallocation of political power is required to take advantage of new commitment opportunities. However, we also find that at low levels of commitment there is more cooperation than would, strictly speaking, be predicted by our parameterized model while the opposite is true at high levels of commitment. Furthermore, only large improvements in commitment opportunities have a significant effect on the social surplus, while small changes do not.

There are several experimental papers related to our work. McAdams (1999) reviews experimental

---

dealing with political interactions that are unlikely to be easily manipulated by researchers (see, among others, Camerer 2003 and Palfrey 2009).

works in law and economics, including Hoffman and Spitzer (1982) and Harrison and McKee (1985), who conducted a series of experiments to test the predictions of the Coase Theorem. In these experiments two subjects must select one alternative from a list of payments for each of them. One of the players is randomly selected to be the controller, who has the right to select the alternative from the list. In some treatments players are allowed to bargain and make side payments. In these treatments the experimenters provided an agreement form and indicated that if subjects decided to use it, they will pay them accordingly. Overall, these experiments show that bargaining allows players to reach the efficient outcome. A key difference with our experiment is that we only give partial and different commitment opportunities to each subject, while in these experiments the agreement form provides full commitment opportunities to both subjects. Indeed, our setting might not be relevant for pre-trial bargaining and settlement, the leading application of Coasian ideas in law and economics. However, the existence of partial and asymmetrically distributed commitment opportunities is pervasive in politics.

There have also been many experiments with simple investment games, some of which allow for commitment opportunities. Van Huyck, Battalio, and Walters (1994) implement an experimental design closed to ours. A peasant has to invest part or all of his endowment, for which he receives an interest minus a tax imposed by a dictator. In the control setup, the tax rate is set after the investment (discretion case), while in the treatment the dictator promises a tax rate before the investment is made (commitment case). Bracht and Feltovich (2008) conducted a laboratory experiment employing a simple game in which one player (the investor) can either invest all or nothing, and the other player (the allocator) can keep everything or split the proceeds. In some treatments allocators were allowed to first put some money into escrow (null, low or high), which will be lost if the allocator keeps all the proceeds. Only a high level of escrow works as a commitment device for allocators, inducing the investor to invest and the allocator to split the proceeds. One difference between these papers and our work is that in our experiment players can partially commit in the sense that only with some probability promises are binding. On the contrary, in Van Huyck, Battalio, and Walters (1994), the dictator can either fully commit or does not commit at all. Similarly, in Bracht and Feltovich (2008) the allocator either fully commits (when the escrow is high) or he does not commit at all (when the escrow is null or low). Another difference is that in our setting we can alter commitment opportunities for each player individually. Finally, while in Bracht and Feltovich (2008) the investor can only choose between investing all or nothing, in our experiment the citizen is allowed to choose five different levels of effort.

Ben-Ner and Putterman (2009) conducted a laboratory experiment using a one-shot investment game. In some treatments parties were allowed to interact before playing the investment game. In particular, parties had the chance to "sign" either a non-binding or a binding contract. Unlike our setting, in which the possibility of any of the parties to partially or fully "tie their hands" is imposed to only one of them before the game begins, in their experiment both players must agree to sign a binding contract. Thus, in their setting commitment opportunities equally affects both parties, while our experiment enables to test how political power is reallocated when parties have different commitment opportunities.

Besides these important differences, our results are in line with the findings in previous works. Hoffman and Spitzer (1982) and Harrison and McKee (1985) report that when subjects employ the agreement form (indeed a binding agreement), outcomes tend to be efficient. Van Huyck, Battalio, and Walters (1994) find that total surplus is higher under commitment, but larger (smaller) in the Discretion (Commitment) setup than predicted by the theory. Bracht and Feltovich (2008) find that when escrow is high (high commitment opportunities for the allocator), the outcome is close to be efficient, but for other treatments the surplus is higher than predicted by the theory. Likewise, Ben-Ner and Putterman (2009) find that the total surplus is higher than predicted by the theory when no commitment device is in place, and that players tend to value fairness in the division of the surplus. Our work also supports the idea that more commitment possibilities lead to more efficient outcomes. The novelty of our paper is that we consider a setting in which agreements are only partially credible, commitment opportunities are asymmetrically distributed between the players and, hence, a reallocation of power might be necessary to take advantage of commitment opportunities. We believe this is a relevant situation for many political interactions. As in previous works, we also find that in our setting cooperation is higher than predicted by the theory when there are no commitment opportunities. After we present our main results we briefly explore potential explanations for this departure.

The rest of this paper is organized as follows: Section 2 summarizes the theoretical framework, beginning with a general but informal presentation of the PCT. The section continues with a brief review of a repeated game that formalizes when the PCT applies and finishes with a description of a simplified version of the model that has been specially adapted for use in our laboratory experiment. Section 3 describes the laboratory experiment. Section 4 shows that subjects understood the game and the randomization was balanced. Section 5 presents the main results of the experiment. Section 6 discusses the departures from theoretical predictions. Finally, Section 7 concludes.

2 Theoretical Framework

In this section we review our theoretical framework. First, we begin with a general statement about the main link between commitment opportunities and social outcomes. Second, we briefly summarize an infinite-horizon repeated game developed by Acemoglu (2003) that captures this link. Third, we develop a simpler two-period game that adapts Acemoglu's model to a laboratory environment and then fully characterize its equilibrium. Finally, we briefly discuss how altruism and honesty affect the equilibrium of our model.

2.1 The PCT and Commitment

As we mentioned in the introduction, if we apply Coase's ideas to politics, we must conclude that, in the absence of transaction costs in political bargaining, society should reach an agreement that yields an efficient social outcome. Given that inefficient policies and institutions are pervasive, a fundamental issue in political economy and institutional economics has been the identification of crucial transaction costs in the political system.

Since the seminal work of North and Weingast (1989), commitment problems have been considered to be one of the primary transaction costs in politics. There are two reasons for this. First, political transactions are usually inter-temporal in the sense that one party offers something today in exchange for a promise of something in the future. Second and more fundamentally, in the political arena, parties cannot rely on contracts being enforced by a third party because powerful agents are, by definition, the ones who wield the power needed to enforce agreements. Thus, powerful agents face a commitment problem which may restrict the nature of the agreements they can reach with other agents.

In order to demonstrate the importance of commitment problems and the link between the ability to commit and social outcomes, let us suppose that powerful agents can somehow commit to a course of action: say, to repay loans or not expropriate others' investments. Then, less powerful agents will be willing to lend to powerful agents and start investment projects because they know that their property will not be expropriated. When only agents that actually do not have political power can commit, the situation may become more complicated, and reaching an efficient outcome will probably require a change in the distribution of political power. Conversely, when agents cannot commit, it will be very difficult to reach efficient outcomes that involve inter-temporal transactions. In general, we can summarize the link between commitment opportunities and social outcomes as follows.

The PCT and Commitment: *When parties* ***can make binding promises****, i.e., when a commitment technology is available, social outcomes will be* ***efficient*** *regardless of how much bargaining power the parties have or what the original distribution of political rights was like, since this only affects the distribution of the social surplus. We say that the* ***PCT applies****. When parties* ***cannot make binding promises*** *social outcomes will be* ***inefficient*** *and the bargaining power of the parties and the original distribution of political rights will influence the distribution as well as the size of the social surplus. In this case, we say that the* ***PCT does not apply****.*

2.2 A Formal Model of the PCT

Acemoglu (2003) develops a political economy model that formally illustrates why we should not expect the Coase Theorem to apply to politics. The model is a game between a ruler and a citizen. The citizen can work in the formal or the informal sector. In the formal sector the production function is $Y_F = e^{1-\alpha} + R$, where $e \geq 0$ is the effort level, $0 < \alpha < 1$, and $R > 0$ represents an exogenous source of income (e.g., rents from natural resources). In the informal sector the production function is $Y_I = b^{\alpha} e^{1-\alpha}$, where $0 < b < 1$. There are two differences between the formal and informal sectors. On the one hand, productivity is lower in the informal sector ($b < 1$). On the other hand, only income coming from the formal sector can be taxed. The citizen's disposable income is given by $Y_d = m\,(Y_F - TS) + (1 - m)Y_I$, where $m = 1$ indicates that the citizen works in the formal sector, $m = 0$ that he works in the the informal sector, and TS is any tax that ruler charges or any transfer that the citizen pays to the ruler. The utility function of the citizen is $Y_d - (1 - \alpha)\, e$. The utility function of the ruler is just the resources she gets from the citizen, i.e., TS. The timing of events is as follows. First, the ruler decides to relinquish power ($r = 1$) or not ($r = 0$). If the ruler relinquishes, then the citizen selects a transfer to the ruler $S\,(Y_F)$, and he decides in which sector he is going to work and an effort level. If the ruler does not relinquish, then the citizen selects the sector that he is going to work in and an effort level. Then, the ruler selects a tax schedule $T\,(Y_F)$. Note that taxes and transfers can never exceed the income in the formal sector, i.e., $S\,(Y_F) \leq Y_F$ and $T\,(Y_F) \leq Y_F$.

It is not difficult to prove that the unique sub-game perfect Nash equilibrium for this game is $r = 0$, $m = 0$, $e = b$, and $T\,(Y_F) = Y_F$, which is an inefficient outcome. But suppose for a moment that, before any player had made a decision, parties were able to sign an enforceable agreement. In that case, it is not difficult to see that they would reach an efficient outcome. In other words, the parties would first agree to maximize the total surplus: the citizen would work in the formal sector ($m = 1$) and he would devote the most efficient level of effort ($e = 1$), generating a surplus of $\alpha + R$. Then, taxes $T\,(Y_F)$ or transfers $S\,(Y_F)$ would be set in order to distribute this surplus between the ruler and the citizen.

The problem with this solution is that the agreement will not be a credible. The citizen knows that the ruler will not have any incentive to keep her original promise once the citizen has set $m = e = 1$

and the ruler has the chance to set $T(Y_F)$. In fact, the only reasonable expectation is that the ruler will appropriate all the income, i.e., $T(Y_F) = Y_F$. Hence, the citizen will prefer to work in the informal sector ($m = 0$ and $e = b$). Alternatively, the ruler knows that the citizen will not have any incentive to keep his original promise once she relinquished. In fact, the only reasonable thing for her to expect is that, after she relinquishes, the citizen will set $S(Y_F) = 0$. Thus, the only possible equilibrium when the parties cannot commit to their promises is the inefficient outcome $r = 0$, $m = 0$, $e = b$, which yields a total surplus of $\alpha b + R$.

In order to escape from this outcome, Acemoglu (2003) considers an infinite-horizon repeated game whose stage game is the one just described. He finds that repeated interactions open the door to credible agreements and better, although not necessarily efficient, outcomes. Indeed, for intermediate levels of the common discount factor, he shows that the size of the equilibrium surplus depends on the bargaining power of the parties. Thus, in general, the Coase Theorem does not apply to politics because powerful parties can only partially commit to respect agreements.

Although simple repeated games have been implemented in the laboratory, we believe that it is better to begin testing the PCT with a simpler model. In the next section, we consider a two-period model that is designed to capture most of the key insights in Acemoglu's infinite-horizon model. The crucial simplification is that promises are only partially enforceable, in the sense that with some probability agents can renege them.

2.3 A Simple Model with Limited Commitment

Consider a simple game with only two players: a citizen and a ruler, denoted by C and R, respectively. The citizen has an endowment of one unit of effort that he can use to produce a private good. The production function is given by:

$$Y = \frac{1}{2} + 2e,$$

where $e \in [0,1]$ is the effort level. The citizen values the private good and leisure. The payoff function for the citizen is given by:

$$v_C = Y_d + \frac{1 - e^2}{2}, \tag{1}$$

where Y_d is his disposable income, i.e., his income after taxes and/or transfers.

The ruler has the power to tax the citizen but she can relinquish this power in exchange for a transfer. Thus, the payoff function for the ruler is given by:

$$v_R = rS + (1 - r)T, \tag{2}$$

where T indicates taxes, S indicates the transfer that she gets if she relinquishes and $r \in \{0,1\}$ is her relinquish decision. Note that disposable income is gross income minus taxes or transfers, i.e., $Y_d = Y - rS - (1 - r)T$. Moreover, since only the private good can be taxed and used for transfers, it must be the case that $0 \leq T \leq Y$ and $0 \leq S \leq Y$.

The timing of events is as follows. (*1*) C selects the transfer S. (*2*) R decides to relinquish ($r = 1$) or not ($r = 0$). (*3*) If R relinquishes, then nature decides if S is enforceable (with a probability ρ that it is enforceable). If S is enforceable, then C selects e. If S is not enforceable, then C selects e and has the chance to reset S. If R does not relinquish, then R selects T. C observes T and decides e. Nature decides if T is enforceable or not (with a probability π that it is enforceable). If T is not enforceable, then R decides on a new T. Otherwise, the promised T applies. *(4)* Payoffs are collected.

This model can be represented as an extensive game with perfect information and random moves (see Osborne and Rubinstein, 1994). The appropriate notion of equilibrium for such games is the sub-game perfect Nash equilibrium. The following proposition formally characterizes the equilibrium of the game.

Proposition 1 *The simple model with limited commitment has a unique sub-game perfect Nash equilibrium which is given by:*

1. *Suppose that $\rho = 0$. Then the ruler does not relinquish ($r = 0$). Moreover:*
   (a) *If $\pi < \frac{1}{2}$, then the ruler promises to levy a tax equal to $T = \frac{1}{2} + 2\pi$ and the citizen works $e = 2\pi$;*
   (b) *If $\pi \geq \frac{1}{2}$, then the ruler promises to levy a tax equal to $T = 2.5 - \frac{1}{2\pi}$ and the citizen works $e = 1$.*
2. *Suppose that $\rho = 1$. Then the ruler relinquishes ($r = 1$) and the citizen works $e = 1$. Moreover:*
   (a) *If $\pi < \frac{1}{2}$, then the citizen promises $S = \frac{1}{2} + 4\pi - 2\pi^2$;*
   (b) *If $\pi \geq \frac{1}{2}$, then the citizen promises $S = 2$.*

Proof: See Appendix 1. ■

It is easy to see that the first best allocation is $e = 1$ and $Y = 1.5$, which yields a total social surplus equal to $TSur = 2.5$. The following corollary summarizes the total social surplus and the division of it for each equilibrium in Proposition 1.

Corollary 1 *Under the assumptions of Proposition 1.*

1. *Suppose that $\rho = 0$. Then:*
   (a) *If $\pi < \frac{1}{2}$, then the equilibrium outcome is Pareto inefficient. Moreover, the expected payoff for the ruler is $\mathbf{E}[v_R] = \frac{1}{2} + 4\pi - 2\pi^2$; the expected payoff for the citizen is $\mathbf{E}[v_C] = \frac{1}{2}$ and the total social surplus is $TSur = 1 + 4\pi - 2\pi^2 < 2.5$.*
   (b) *If $\pi \geq \frac{1}{2}$, then the equilibrium outcome is Pareto efficient. Moreover, the expected payoffs for the ruler and the citizen are $\mathbf{E}[v_R] = 2$ and $\mathbf{E}[v_C] = \frac{1}{2}$, respectively.*
2. *Suppose that $\rho = 1$. Then, the equilibrium outcome is Pareto efficient. Moreover:*
   (a) *If $\pi < \frac{1}{2}$, then the expected payoffs for the ruler and the citizen are $\mathbf{E}[v_R] = \frac{1}{2} + 4\pi - 2\pi^2$ and $\mathbf{E}[v_C] = 2 - 4\pi + 2\pi^2$, respectively;*
   (b) *If $\pi \geq \frac{1}{2}$, then the expected payoffs for the ruler and the citizen are $\mathbf{E}[v_R] = 2$ and $\mathbf{E}[v_C] = \frac{1}{2}$, respectively.*

Proof: Straightforward deduction from Proposition 1. ■

The key issue in this game is commitment. The citizen would like to put in more effort, but he knows that the ruler will tax his income. The ruler knows this and she will be better off if she signs a credible agreement to restrict taxation. But, the problem is that she is the ruler and, hence, only she has the power to enforce agreements. Thus, her ability to make a credible promise to limit taxation depends on her ability to tie her hands and commit herself to that. The probability $\pi \in [0, 1]$ is a measure of the strength of this commitment. Another alternative is that she relinquishes her position in exchange for a payment. The problem is again commitment. Once the citizen has the power, he will not be willing to keep his promise and the ruler will get nothing. The citizen would like to commit himself to pay the ruler if she relinquishes, but, here again, his ability to tie his hands is limited. The probability $\rho \in \{0, 1\}$ is a measure of this ability. When neither the ruler nor the citizen can find a way to commit the outcome of the game is very inefficient. However, when either the ruler or the citizen can partially commit, it is possible to support more cooperative outcomes.

2.4 Pro-Social Preferences: Altruism and Honesty

The results in Proposition 1 and Corollary 1 assume that agents are completely selfish and do not face any phycological cost when they do not keep their promises. Now, we explore how these results would be affected if agents had pro-social preferences, i.e., they care about the welfare of others and/or they are honest and do not like to lie.

When $\rho = 0$ and $\pi \geq 1/2$ or $\rho = 1$, the equilibrium when agents are purely selfish and dishonest is Pareto efficient. Thus, there is no way that altruism and/or honesty induce an increase in total surplus. When $\rho = 0$ and $\pi < 1/2$, the equilibrium when agents are assumed purely selfish and dishonest, is not Pareto efficient. It is easy to see how an altruistic ruler can improve this outcome. The citizen knows that an altruistic ruler will not fully tax him and, hence, he will be willing to put more effort. An honest ruler can also improve things because honesty works as an informal commitment device. If the ruler promises a low tax, the citizen knows that the honest ruler will keep her promise. Analogously, a ruler that faces an altruistic citizen knows that he will be more willing to pay her a transfer if she relinquishes. An honest citizen can also make a credible promise that he will pay a transfer if the ruler relinquishes. Note, however, that if only a small proportion of the agents are altruistic and/or honest, the equilibrium outcome will not be Pareto efficient.

In general, although altruism and honesty could easily affect the results in Proposition 1 and Corollary 1, we should not expect that they are strong enough to change the main implication that more cooperative outcomes can be supported when either the ruler or the citizen can partially commit. The main focus of our laboratory experiment will be to test this critical implication.

3 The Laboratory Experiment

In this section we describe our laboratory experiment. First, we provide a general description of the experiment, including its monetary payoffs, number of sessions and rounds, matching procedure, and the instructions received by the subjects. Second, we give a detailed description of the game subjects played. Finally, we summarize the treatments and compute the corresponding predicted outcomes using Proposition 1 and Corollary 1.

3.1 General Description of the Experiment

The experiment was conducted between August and November 2012 at Universidad de San Andrés, Argentina. We recruited undergraduate and graduate students from any field of study and regardless of how familiar they were with game theory. We conducted 10 sessions with 16 subjects each, totalling 160 participants. Subjects were allowed to participate in only one session. Every session included the four treatments, which avoids any selection problem among treatments. In each treatment, subjects were asked to play a simple game involving limited commitment. The experiment was programmed and conducted using z-Tree software (Fischbacher, 2007). Each session lasted approximately 90 minutes.

The experiment proceeded as follows:

1. **Allocation to Computer Terminals**: Before each session began, subjects were randomly assigned to computer terminals.

2. **Instructions**: After the 16 subjects were at their terminals, they received general instructions and, then, the rules of the game were explained using a PowerPoint presentation. Instructions and explanations were always presented to the subjects using neutral words. In particular, subjects were never told that they would be playing a political game and ruler and citizen were always denoted as player 1 and 2, respectively. Appendix 2.1 and 2.2 contain English translations from Spanish of the script we employed for general instructions and the PowerPoint presentation, respectively.

3. **Quiz**: In order to check whether participants understood the rules of the game, we asked them to take a five-question quiz. The quiz was administered after we had given the instructions, but before the rounds began. Subjects were paid approximately US$ 0.81 per correct answer. The quiz questions can be found in Appendix 2.3.

4. **Matching**: In each session, subjects were first randomly assigned to one of two different roles - player 1 (the ruler) or player 2 (the citizen)- and then to four distinct treatments. All pairings were done through the computer. After each round, subjects were re-matched with another partner for the next round. No player knew the identity of the player with whom she was currently paired or the history of decisions made by any of the other players. Due to the fact that subjects maintained their treatment and role, there were only two possible partners for each subject. Note, however, that no subject played with the same partner in two consecutive rounds.

5. **Rounds**: After subjects finished the quiz, they began playing rounds, during which they interacted solely through a computer network using z-Tree software. Subjects played six rounds of the same game, with the caveat that they would never play two consecutive rounds with the same subject. The first two rounds were for practice, and the last four rounds were for pay. The rules of the game were always available through a 'help box' and each screen had access to a calculator. At the end of each round, subjects received a summary of the decisions taken by both themselves and their partners, including payoffs per round, their own accumulated payoffs for paid rounds, effort level, promises, total income before taxes, and nature's decision (when applicable); they were also reminded of the payoff functions of the game. See below for a detailed description of the game that subjects played.

6. **Questionnaire**: Finally, just before leaving the laboratory, all the subjects were asked to complete a questionnaire, which was designed to enable us to test the balance across experimental groups and to control for their characteristics in the econometric analysis presented below. Appendix 2.5 contains the questionnaire.

7. **Payments**: All subjects were paid privately, in cash. After the experiment was completed, a password appeared on each subject's screen. The subjects then had to present this password to the person who was running the experiment in order to receive their payoffs. Subjects earned, on average, US$ 22.50, which included a US$ 2.07 show-up fee, US$ 0.81 per correct answer on the quiz, and US$ 5.18 for each point they received during the paid rounds of the experiment. All payments were made in Argentine currency; at the time, US$22.50 was equivalent to AR$ 108.

3.2 The Game in Each Round

Once they finished the quiz, subjects directed their attention to their computers and proceeded to play the first round of the session. Appendix 2.4 contains a sample of the first screens that subjects visualized, which present the rules of the game. Similar prompts appear on-screen before every round.

In each round players 1 and 2 played the following limited commitment game. Player 1 has the power to tax player 2, who must decide how much effort he puts. Player 1 can also relinquish to her power to tax in exchange for a transfer by player 2. In addition players can exchange promises about taxes and transfers. These promises are enforceable with varying degrees in different treatments. The sequence of play has four stages.

- **Stage 1**: Promises about S. Player 2 promises a transfer $S = min\{S_0, Y\}$ with $S_0 \geq 0$ to player 1 if player 1 relinquishes to her power to tax. This promise is enforceable with probability ρ.

- **Stage 2**: Relinquish Decision and Promises about T: Player 1 decides whether to relinquish or not. If she does not relinquish she also selects and promises a tax $T = min\{T_0, Y\}$ with $T_0 \geq 0$. This promise is enforceable with probability π.

- **Stage 3**: Working Decision:

  – If player 1 relinquishes, then nature determines if S is enforceable or not. If S is enforceable, then player 2 selects the effort level e. If S is not enforceable, then player 2 selects the effort level e and has the chance of resetting the transfer S.

  – If player 1 does not relinquish, then player 2 observes the promised tax $T = min\{T_0, Y\}$ with $T_0 \geq 0$ and decides the effort level e. Then, nature determines if T is enforceable.

- **Stage 4**: Payoffs: Player 1 gets $rS + (1-r)T$ and player 2 gets $Y - rS - (1-r)T + \frac{1-e^2}{2}$, where $r = 0, 1$ is the relinquish decision, T and S are the final effective tax and transfer, $Y = 0.50 + 2e$ and $e = 0, 0.2, 0.5, 0.8, 1$ is the effort level.

Some remarks apply. First, note that the maximum possible tax is Y. The way we implement this restriction in the laboratory is by asking subjects to select $T_0 \geq 0$ and establishing that this will induce a tax equal to $T = min\{T_0, Y\}$. Analogously, for the transfer S. In the PowerPoint presentation we employed to explain the game we clearly describe this point (see Appendix 2.2); the quiz contains two questions about this issue (see Appendix 2.3); and the rules of the game on screen also explain that $T = min\{T_0, Y\}$ and $S = min\{S_0, Y\}$. Second, in order to make computations easier we only allow the subjects to select five levels of effort: very low ($e = 0$), low ($e = 0.2$), medium ($e = 0.5$), high ($e = 0.8$) and very high ($e = 1$).

3.3 Treatments and Predicted Outcomes

The experiment consisted of four different treatments. The first treatment represents a scenario of no commitment opportunities ($\pi = \rho = 0$); the second, a scenario of low commitment opportunities for the ruler ($\pi = 0.25$ and $\rho = 0$); the third, a scenario of high (but not full) commitment opportunities for the ruler ($\pi = 0.75$ and $\rho = 0$); and the fourth, a scenario of full commitment opportunities for the citizen ($\pi = 0$ and $\rho = 1$). Employing Proposition 1 and Corollary 1, we can compute the outcome predicted by our parameterized model for all these treatments. Table 1 depicts our four treatments and the corresponding predicted outcomes. (Recall that $TSur$ is the total social surplus, e is effort, r is the relinquish decision, $\mathbf{E}[v_R]$ is the expected payoff for the ruler, and $\mathbf{E}[v_C]$ is the expected payoff for the citizen).

See Table 1: Treatments and Predicted Outcomes

As shown in Table 1, we should expect that the social surplus, defined as the sum of the individual payoffs for the players matched in each session and period, should be 1 for T1, 1.875 for T2, and 2.5 for T3 and T4. Rulers are expected to relinquish in T4 and do not to relinquish in the other treatments. Effort should be 0 in T1, 0.5 in T2, and 1 in T3 and T4. As for the distribution of the surplus, in T1, T2, and T3 the citizen should obtain an expected payoff of 0.5 and the ruler should receive 0.5, 1.375 and 2, respectively. Thus, in T1, T2 and T3, the ruler should keep all the increase in the social surplus due to the rise in commitment opportunities. Conversely, in T4, the ruler should receive an expected payoff of 0.5 and the citizen should obtain a payoff of 2, collecting all the increase in surplus due to the rise of commitment opportunities.

Our simple model with limited commitment is much simpler than the infinitely repeated game developed by Acemoglu (2003). Still, it is a multistage game in which any departure from the equilibrium path in one node of the game can easily induce changes in predicted outcomes, even if subjects play rationally in successive nodes. Moreover, our matching procedure implies that subjects will meet the same partner every other round. Technically speaking, this makes the game subjects played in the laboratory a two rounds finitely repeated version of our game.[2] Note, however, that the subgame perfect Nash equilibrium of the two rounds finitely repeated version is simply the subgame perfect Nash equilibrium of our game in each round. Finally, as we have seen in section 2.4 altruism and honesty can affect the equilibrium outcomes in Table 1, which are deduced based on the assumption that agents are purely selfish and willing to lie to pursue their goals. Therefore, we should not expect that subjects in the laboratory can perfectly replicate the parameterized model predictions summarized in Table 1.[3] A less demanding test for the model would be to check if laboratory results are consistent with the comparative statics predicted by it.

[2] If we include the two practice rounds, the game is repeated three times.

[3] For a more detailed methodological discussion of what we can learn from experiments about economic primitives and theoretical models, see Smith (2010) and Friedman (2010).

For example, we should expect that the total surplus in T3 and T4 will be higher than in T1 for a wide range of levels of altruism. (Only extremely altruistic subjects could induce a sharp enough rise in the total surplus in T1 to make the difference between T3 and T1 negligible.)

From Table 1 we can easily deduce the following three key comparative static predictions:

1. **Total Social Surplus**: The social surplus should increase when commitment opportunities rise. More precisely $TSur(j)$ indicates the total surplus in treatment Tj:

$$TSur(4) = TSur(3) > TSur(2) > TSur(1).$$

2. **Reallocation of Political Power**: As the citizen gains access to more commitment opportunities, the ruler should relinquish more. More precisely, ($\mathbf{Pr}(r = 1)(j)$ denotes the proportion of rulers that relinquish):

$$\mathbf{Pr}(r = 1)(4) > \mathbf{Pr}(r = 1)(3) = \mathbf{Pr}(r = 1)(2) = \mathbf{Pr}(r = 1)(1).$$

3. **Distribution of the Surplus**: The payoff for a player should increase with his/her own commitment opportunities and does not vary with the other player's commitment opportunities. More precisely, $\mathbf{E}[v_i](j)$ denotes the expected payoff for player $i = C, R$ in treatment Tj):

$$\mathbf{E}[v_C](4) > \mathbf{E}[v_C](3) = \mathbf{E}[v_C](2) = \mathbf{E}[v_C](1),$$

and

$$\mathbf{E}[v_R](3) > \mathbf{E}[v_R](2) > \mathbf{E}[v_R](1) = \mathbf{E}[v_R](4).$$

4 Understanding of the Game and Randomization Balance

In this section we show that subjects understood the game and the randomization was balanced.

Table 2 shows that on average subjects understood the rules of the game. Indeed, from a maximum in the quiz grade of 12 points, subjects scored on average 10.91 points. 89% of the subjects correctly answered question 1, 81% question 2, 89% question 3, and 78% question 4. It seems that subjects found that question 5 was more complicated and only 28% of them correctly answered it.

Table 2 also shows the randomization balance across player roles (citizen vs. ruler). Note that all characteristics and understanding of the rules of the game are perfectly balanced across roles, as the mean difference between citizens and rulers is not significantly different from zero either for subject characteristics or for their understanding of the game.

See Table 2: Balance Across Players

Tables 3 and 4 , show that in the comparisons among the four treatments, all characteristics and levels of understanding of the game were perfectly balanced between T1 and T2 and between T3 and T4. In some of the other cases, there is a slight imbalance in gender and nationality, mostly at a 10% significance level. Nevertheless, it was only in less than 10% of the tests that we rejected the null hypothesis at the 10% level of statistical significance. Moreover, the imbalance in nationality is probably due to the fact that there were very few foreigners in the sample (92.5% of the subjects were Argentines).

See Tables 3 and 4: Balance Across Treatments I and II

5 Main Results: Comparative Statics

In this section we present the main results of the experiment. Overall, we find support for the hypothesis that greater commitment opportunities lead, on average, to better social outcomes. We first provide a descriptive analysis of the decisions taken by the subjects, and then present the econometric results obtained when testing the comparative statics.

5.1 Descriptive Analysis

Table 5 shows descriptive statistics for the main decisions taken by the subjects. For each treatment Table 5 indicates the outcome predicted by the model, the total number of observations, and the sample mean and standard deviation for the corresponding variable in each column. Column (1) reports the total surplus; column (2) effort; column (3) the relinquish decision; column (4) the payoff for the ruler; column (5) the payoff for the citizen; column (6) the tax that the ruler promises to charge; column (7) the transfer that the citizen promises to pay if the ruler relinquishes; column (8) the tax that the ruler sets if his/her promise ends up being non enforceable; column (9) the transfer that the citizen sets if his/her promise ends up being non enforceable; column (10) the tax that the ruler actually charges; and column (11) the transfer that the citizen actually pays.

See Table 5: Decisions Across Treatments (Descriptive Statistics)

Total Social Surplus: Table (5) column (1) and Figure 1.a show the mean total social surplus across treatments. As predicted by the model, the mean surpluses for T3 and T4 are greater than for T1 and T2. Though the mean surplus in T4 is slightly larger than in T3, the difference is very small. Even though the exact magnitudes of the surpluses are not replicated empirically, the relative magnitudes tend to support the predictions of the model. Since the social surplus is an increasing function of effort, essentially the same pattern can be seen in the effort decisions. Indeed, Table (5) column (2) and Figure 1.b show that the average effort levels are 0.32 in T1, 0.33 in T2, 0.62 in T3 and 0.66 in T4. Thus, as predicted by the model, effort is higher in T3 and T4 than in T1 and T2. However, average effort levels are lower than theoretical predictions for all treatments except T1. Note that the mean surplus is concentrated at two points (around 1.5 and around 2). Subjects appear to act similarly when the probability of enforcement is zero or very close to zero (namely, 25%). In these situations, the social surplus is rather small, as citizens decide to reduce their effort in the hope of preventing the ruler from confiscating their income. As the probability of enforcement rises to 75%, the results begin to cluster around a total surplus of about 2 points, as commitment opportunities now allow for more socially efficient results. Finally, Figures 1.c and 1.d show the total surplus and effort level (means and standard deviations) by round. Note that the mean social surplus fluctuates across rounds of the experiment, especially for T1 and T3, but there is no clear pattern. Similarly, the mean effort also fluctuates without any pattern.

See Figure 1: Total Surplus and Effort Level by Treatment and Round

Reallocation of Political Power: Regarding the decision to relinquish power or not, approximately 10% of rulers relinquished power in T1, 3.7% in T2, 2.5% in T3, and 51% in T4 (Table 5 column 3 and Figure 2.a). As predicted by the model, the percentage increases considerably from T1, T2, and T3 to T4; nonetheless, the difference is not as large as theoretical predictions suggested (rulers should never relinquish in T1, T2 or T3 and always relinquish in T4). Figure 2.b shows the relinquish decision (mean and standard deviation) by round. Again we do not note any pattern across rounds.

See Figure 2: Relinquish Decision by Treatment and Round

Distribution of the Surplus: As predicted by the model, the payoff for citizens is, on average, higher in T4 (0.903) than in T1 (0.573), T2 (0.494), or T3 (0.722), while the payoff for rulers is, on average, higher in T3 (1.236) than in T2 (1.037) and higher in T2 than in T1 (0.94) (Table 5 columns 4 and 5 and Figure 3). However, the distribution of the social surplus between rulers and citizens does not exactly coincide with theoretical predictions. In T1, citizens earned an average of 0.573 points and rulers an average of 0.903. Thus, approximately 80% of the 0.513 extra points with respect to theoretical predictions went to rulers. In T2, rulers earned an average of 1.037 points and citizens an average of 0.494. Thus, on average, citizens obtained the payoff predicted by the model, while rulers obtained 0.3 points less. In T3, rulers earned an average of 1.236 points and citizens an average of 0.722, rather than the 2 and 0.5 points predicted by the model, respectively. Last, in T4, rulers earned an average of 1.109 points rather than the 0.5 points predicted by the model, while citizens received only 0.903 points rather than the 2 points predicted by the model. Figures 3.c and 3.d show the payoff for rulers and citizens (mean and standard deviation) by round. Once again we do not note any pattern across rounds.

See Figure 3: Payoffs by Treatment and Round

The last columns in Table 5 also provided useful insights on why expanding commitment opportunities improve social outcomes. Specifically, note that the mean T_{Prom} was always lower than the mean T_{Decid} (columns (6) and (8) in Table 5), while the mean S_{Prom} was always higher than the mean S_{Decid}, except for T4, a treatment in which S_{Prom} must be implemented with probability 1 (columns (7) and (9) in Table 5). This suggests that on average rulers and citizens did not fully keep their promises. In other words, without a formal commitment device subjects had troubles to credibly commit to fulfill their promises.

5.2 Regression Analysis

We now formally test our three comparative static results using regression analysis. Note that in the context of perfect experimental data, where no controls are needed for identification of the causal effects of interest, the analysis is completely non-parametric as it only entails to compare the mean outcome differences across treatment groups and inference also could be made non-parametric. Clustered standard errors are computed both at the pair of subjects matched in a given session, and also, somewhat more conservative, at the group of 4 subjects matched to play among them in a given session.

Total Social Surplus: In order to formally test the hypotheses that more commitment opportunities lead to better social outcomes (a larger social surplus) we use the following regression model:

$$TSur_{gps} = \alpha + \beta_1 D\mathrm{T} + \beta_2 X_{gps} + \sum_{s=1}^{9} \beta_{3s} D\theta_s + \beta_{4s} Q_{gps} + \epsilon_{gps}, \quad (3)$$

where g indexes a particular pairing of subject partners, $p = 1, 2, ..., 4$ indexes experimental rounds, and $s = 1, 2, ..., 10$ indexes experimental sessions. There are 32 pairings of subject partners in the data set, but only 16 of them will enter in each estimated regression.[4]

Total surplus per group, session and round ($TSur$) is the dependent variable. $TSur$ is computed as the sum of the payoffs for the citizen and the ruler in each group, session and paid round. Therefore, each observation corresponds to a given pair of subjects in a particular session and round. The explanatory variable of interest is DT, a dummy variable indicating treatment status (T2, T3 or T4). In some specifications we also include control variables. We control for individual characteristics X_{gps} (for both subjects in group g we control for gender, age, nationality, for racial group, for whether s/he has ever taken a course in game theory, for whether s/he is a graduate or a junior or senior undergraduate student), for the subjects' level of understanding of the game as measured by their answers to the quiz questions Q_{gps} (mean quiz-mark per group and quiz-mark of the citizen in each group), and for fixed effects by session ($D\theta_s$).

According to our theoretical predictions, we should expect β_1 to be positive when comparing T2 to T1, T3 to T2, T3 to T1 and T4 to T1, while we should expect β_1 to be zero when contrasting T4 with T3.

Column (1) in Table 6 summarizes the results of regressing the total surplus in each of the treatments separately without controls. Clustered standard errors computed at the pair of subjects matched in a given session are shown in parentheses. Clustered standard errors computed at the group of 4 subjects matched to play among them in a given session are shown in brackets. The total surplus in T2 is not significantly different from the total surplus in T1, though the coefficient associated with the treatment is indeed positive, in keeping with the model's prediction. Operating under the parameters in T3 (or T4) rather than under T2 (or T1) induces a significant rise in the total surplus which goes from 0.426 to 0.499 points, which is an increase of approximately 30% over the counterfactual. Thus, as predicted by our model, higher commitment opportunities tend to lead to better social outcomes. Finally, total surpluses in T3 and T4 are not significantly different, which suggests that the subjects understood that, in order to take advantage of the citizen's commitment opportunities, in T4 the ruler must relinquish.

In Table 6 column (2), we report the results when the same analysis is performed once the entire set of controls, as described above, was included. As the table shows, the results do not change in any meaningful way.[5] To sum up, large improvements in commitment opportunities (from T1 to T3 or T4) have a significant positive effect on the social surplus, while a small change (from T1 to T2) has no more than a small positive (and statistically non-significant) effect.

See Table 6: Regression Analysis Total Surplus

[4]Though we could have pooled all the data together for estimation, that would have not changed the analysis at all in our main specification where we do not include control variables. For expositional purposes we choose to present the results by contrasting pairs of experimental groups.

[5]Including the controls by round within each session does not alter the results either.

Reallocation of Political Power: In order to formally test the hypothesis that the probability of a reallocation of political power is higher when the citizen has more commitment opportunities, we use the following regression model:

$$r_{gps} = \delta + \gamma_1 D\mathrm{T} + \gamma_2 X_{gps} + \sum_{s=1}^{9} \gamma_{3s} D\theta_s + \gamma_{4s} Q_{gps} + \epsilon_{gps}, \tag{4}$$

where g indexes a particular pairing of subject partners, $p = 1, 2, ..., 4$ indexes experimental rounds, and $s = 1, 2, ..., 10$ indexes experimental sessions. Relinquish (r) is the dependent variable. Explanatory and control variables are the same as in the regression model (3).[6]

According to our theoretical predictions, we should expect γ_1 to be positive when comparing T4 to T1, T2 or T3, while we should expect γ_1 to be zero when contrasting T1 with T2, or T3 and T1 with T2.

Table 7 summarizes the results of a regression analysis of the decision to relinquish power in each of the treatments separately. Clustered standard errors are shown in parentheses and brackets. As predicted by our model, operating under the parameters in T4 rather than in T1, T2 or T3 induces a positive and statistically significant increase in the decision to relinquish power. The proportions of rulers that relinquish power is not significantly different in T1 than in T2, or in T2 than in T3. Although the proportion of rulers who relinquish power is significantly higher in T1 than in T3, the magnitude of the difference is very low.

See Table 7: Regression Analysis Reallocation of Political Power

Distribution of the Surplus: In order to formally test the hypotheses that the payoff for a player increases with his/her own commitment opportunities and does not vary or decrease in line with the other player's commitment opportunities, we use the following regression model:

$$v_{igps} = \zeta + \eta_1 D\mathrm{T} + \eta_2 X_{gps} + \sum_{s=1}^{9} \eta_{3s} D\theta_s + \eta_{4s} Q_{gps} + \epsilon_{gpsi}, \tag{5}$$

where g indexes a particular pairing of subject partners, $p = 1, 2, ..., 4$ indexes experimental rounds, $s = 1, 2, ..., 10$ indexes experimental sessions. The payoff for player i (v_i) is the dependent variable. Explanatory and control variables are the same as in the regression model (3).

According to our theoretical predictions for the ruler, we should expect η_1 to be positive when comparing T3 to T1, T2 or T4 and when comparing T2 to T1 or T4, while we should expect η_1 to be zero when contrasting T1 with T4. For the citizen, we should expect η_1 to be positive when comparing T4 to T1, T2 or T3, while we should expect η_1 to be zero when contrasting T1 with T2 or T3.

Tables 8 and 9 summarize the results of regressing the payoffs for rulers and citizens, respectively, in each of the different treatments. The corresponding clustered standard errors are shown in parentheses and brackets. As predicted by our model, operating under the parameters in T3 rather than in T1, T2 or T4 induces a positive and statistically significant effect on the payoff for rulers (one-tailed test). Contrary

[6]Note that in our main model (without controls), the comparison of means binary variables is also identified by means of a linear regression model since the conditional mean is saturated in that specification. Adding control variables does not change the estimate of the group treatment effect since those are balanced by randomization.

to what is suggested by our model, however, operating under the parameters in T4 rather under T1 or T2 induces a positive and statistically significant effect on the payoff for rulers. As predicted by our model theory, operating under the parameters in T4 rather than in T1, T2 or T3 also induces a positive and statistically significant rise in the payoff for citizens, while operating under the parameters in T1 rather than in T2 has no effect on citizens' payoff. Contrary to our model, operating under the parameters of T3 induces a positive and statistically significant increase in the payoff for citizens.

See Tables 8 and 9: Regression Analysis Payoffs

6 Beyond Comparative Statics

As mentioned in Section 3, we should not expect subjects in the laboratory to exactly replicate equilibrium outcomes. In this section we try to identify the departures from the benchmark provided by our model and briefly explore how and why they occur.

Table 5 shows that for T1 there is more cooperation than predicted by our calibrated model, while the opposite is true for T2, T3 and T4. Big improvements in commitment opportunities (from T1 to T3 or T4) have a significant effect on the social surplus, while smaller changes (from T1 to T2) do not. It is also evident that rulers do not always relinquish in T1, T2, and T3 and only 51% of the rulers relinquish in T4. Nor is the distribution of the surplus exactly as predicted. Sometimes one of the players systematically obtained more or less than theoretical predictions would indicate. Next, we take a closer look at subjects' behavior by exploring their decisions at key nodes of the game. Table 10 breaks down the results shown in Table 5 into relevant nodes.[7]

See Table 10: Beyond Comparative Statics

Under T1, the total surplus is, on average, larger than the amount predicted by the theory (1.513 points versus 1 point). In part, this difference is accounted for a small fraction of rulers who selected $r = 1$. As Table 10 shows, in T1 rulers relinquished 8 times, which accounts for 10% of the observations. Note that, conditional on $r = 1$, it is optimal for the citizens to select the maximum possible level of effort, i.e., $e = 1$, inducing a total surplus of 2.5 points. Indeed, we can observe from Table 10 that, within this group, the average level of effort was 0.95 and the average surplus was 2.445 points. However, contrary to theoretical predictions, citizens did not keep all the surplus for themselves. The average payoff for the rulers was 0.594 points. Note that when the citizen is selecting S the game is just a dictator game. The citizen can appropriate the entire surplus or share a portion of it with the ruler, but the ruler does not have the opportunity to accept or reject the distribution selected by the citizen. Thus, it seems that citizens are being altruistic in their dealings with the rulers that selected $r = 1$. Alternatively, we can interpret that citizens are rewarding rulers that trusted them. The ruler acts "kindly" by offering the citizen the possibility of producing a high surplus and then setting a transfer, and the citizen "reciprocates" by making a transfer.

Not all of the difference between the predicted and actual total surpluses under T1 can be attributed to rulers who selected $r = 1$. Indeed, as Table 10 shows, the average total surplus for the 72 observations

[7]Note that when the ruler relinquishes there are no promises made or decisions taken about T. Similarly, when the ruler does not relinquish, decisions about S are never made. In Table 10 we indicate these situations with NA (non-applicable).

in which rulers selected $r = 0$ is 1.409 points, while the average level of effort is 0.257. Thus, citizens are putting in more effort than the model predicts (0.257 versus 0) and, as a consequence, they are losing on average 0.069 points (they achieved an average of 0.431 points, whereas they could have obtained 0.5 points by selecting $e = 0$). This generates an extra payoff of 0.478 points for the rulers (on average, they received 0.978 points when they could have obtained 0.5 points if $e = 0$).[8] Summing up, on average, citizens are giving up 0.069 points and rewarding the rulers with 0.478 points. Again, one possibility is citizens are being altruistic. Although now the extra payoff for the rulers is costing the citizens only 0.069 points. Alternatively, there might be two effects working in opposite directions. On the one hand, when the ruler does not relinquish it is cheaper for citizens to be altruistic with the rulers. On the other hand, citizens might also want to punish the rulers for not trusting them.

Similar effects have been found to exist in previous laboratory experiments and they've been attributed to altruism, reciprocity and fairness. Andreoni and Miller (1993) find stable levels of cooperation of around 10%-15% when the model predicted no cooperation and they attribute this to altruism. Camerer and Weigelt (1988) estimate that the subjective prior that an opponent would play cooperatively is about 17% and McKelvey and Palfrey (1992) estimate that the proportion of altruists is between 5% and 10%. Keeneth and Martin (2001) investigate how fairness concerns influence individual behavior in social dilemmas using a sequential prisoner's dilemma experiment. They also find that the proportion of altruists is between 12.5% and 26%. More generally, Fehr and Gachter (2000) review several experiments that test for reciprocity in different settings, Andreoni and Miller 2002 find that subjects exhibit a consistent preference for altruism, while Charness and Haruvy (2002) test competing theories of non-pecuniary motives. They find that reciprocity, fairness, and altruism all play an important role in subjects' decisions.

Finally, it is important to remark that rulers who selected $r = 1$ obtained, on average, a lower payoff (0.594 points) than those that selected $r = 0$ (0.978 points); although the difference is not statistically different from zero.[9] Nevertheless, relinquishing does not appear to be the result of comprehension problems. Actually, the average quiz grade is higher for those who selected $r = 1$.[10] Rather, we posit that such behavior could be explained by the fact that some rulers were acting out of a desire to see how kindly citizens would behave if they select relinquish.

Under T2, the total surplus is, on average, lower than predicted (1.531 versus 1.875 points).[11] As in the case of T1, there are some rulers who selected $r = 1$ (only 3 observations, in this case), which marginally increases the average total surplus. If we focus on the 77 observations in which rulers selected $r = 0$, the average total surplus is 1.508 points. Why was the total surplus lower than predicted (1.508 versus 1.875 points)? The answer is that citizens put in less effort than they should have. The average e was 0.322, whereas it should have been 0.5. As a consequence, citizens lost an average of 0.053 points while rulers got an average of 1.061 points when they could have obtained 1.375 points. But the real

[8] The null hypothesis 'social surplus in T1 given that the ruler did not relinquish equals 1' is rejected both with and without controls (p-value = 0.000). The null hypothesis 'effort in T1 given that the ruler did not relinquish equals 0' is rejected without controls (p-value = 0.00) and with controls (p-value = 0.0089). The payoff of rulers and citizens in T1 given that the ruler did not relinquish are also significantly different from 0.5 (p-value = 0.000).

[9] The p-value for the test is 0.1760 without controls and 0.4509 with controls.

[10] We cannot reject the hypothesis that the marks of the quizzes for rulers that select $r = 0$ and $r = 1$ in T1 are the same. The p-value for this test is 0.4091 without controls and 0.692 with controls.

[11] The difference is statistically significant. The test total surplus in T2 equals to 1.875 is rejected (p-value = 0.000) with and without controls.

problem is that some of the rulers promised too high a tax. As Table 10 shows, the average T_{prom} was 1.128 points, when it should have been 1.[12]

As in the case of T2, under T3 the average surplus is also lower than predicted. On average, the total surplus was 1.957 points for all the observations and 1.943 points when rulers selected $r = 0$, versus a prediction of 2.5 points. Under T3 there are only two observations in which the rulers selected $r = 1$. Why is the total surplus lower than predicted? As in the case of T2, under T3 citizens also put in less effort than they should have. Conditional on $r = 0$ the average e was 0.618 but it should have been 1. However, this time the problem is not that the rulers promised too high a tax. On the contrary, they promised a tax that was lower than predicted. As Table 10 shows, on average T_{prom} is 1.424 versus a prediction of 1.83. The key problem is that citizens put in too little effort. As a consequence, rulers obtained less than predicted (an average of 1.267 points versus a prediction of 2 points), and citizens obtained more (an average of 0.676 points versus a prediction of 0.5 points).[13]

Under T4, the total surplus is, on average, lower than predicted (2.012 points versus 2.5). Most of the difference is accounted for by rulers who did not relinquish. Indeed, if we focus on the 51% (41 observations) of the rulers who selected $r = 1$, then the total surplus is, on average, 2.480 points and the average e is 0.99. Rulers that selected $r = 1$ obtained, on average, more than predicted (1.191 points versus 0.5 points), while citizens obtained less (1.298 points versus 2 points).[14] If the ruler decided $r = 0$ (this happened in 39 observations, i.e., 49% of the time), then the subsequent game is as in T1 when $r = 0$. If subjects believe that the game will be played as in T1, then rulers and citizens should expect to obtain an average of 0.978 and 0.431 points, respectively. In reality, they obtained an average of 1.023 and 0.487 points, respectively.[15]

Thus, it seems that rulers realized that citizens were being altruistic and internalized this preference in their relinquish decision. Indeed, when rulers selected $r = 1$, citizens offered an average of 1.191 points, whereas, when rulers selected $r = 0$, citizens offered an average of 0.738 points. In other words, this suggests that rulers relinquished only when citizens offered a transfer S that takes into account the fact that citizens will behave altruistically. Note that it is also possible that the logic of the ultimatum game applies here.[16] If a citizen shows that he is too greedy by offering a very low S, then the ruler will punish him by selecting $r = 0$ which will significantly reduce the payoff for the citizen. Indeed, rulers are not losing a great deal when they do not relinquish, while citizens are paying the total cost of this decision.

[12] The hypothesis 'total surplus in T2 equals to 1.875', 'payoff of the ruler in T2 equals to 1.375', 'payoff of the citizen in T2 equals to 0.5', 'effort in T2 equals to 0.5' (in all cases given that the ruler did not relinquish) are all rejected (p-value = 0.000 for each test) with controls. Finally, the hypothesis 'T_{Prom} in T2 equals 1 given that the ruler did not relinquish' is also rejected (p-value = 0.0009) with controls.

[13] The hypothesis 'total surplus in T3 equals 2.5' is rejected (p-value = 0.000) with controls, both for all observations and conditioning in $r = 0$. The hypothesis 'payoff of the citizen in T3 equals 0.5', 'payoff of the ruler in T3 equals 2 in T3', are also rejected (p-value = 0.000) with controls conditioning in $r = 0$. The hypothesis 'effort equals 1 conditional on $r = 0$ for T3' is rejected with p-value = 0.000. Finally, the hypothesis 'T_{prom} in T3 equals 1.83 conditioning in $r = 0$' is also rejected.

[14] The hypothesis 'payoff of the ruler in T4 given $r = 1$ equals 0.5' and 'payoff of the citizen in T4 given $r = 1$' are both rejected.

[15] The p-values for these tests are 0.7816 and 0.3884, respectively, indicating that conditional on $r = 0$, the equality of individual payoffs across treatments 1 and 4 cannot be rejected.

[16] There are numerous papers that report results from ultimatum games. See, for example, Forsythe, Horowitz, Savin and Sefton (1994); Hoffman, McCabe, Shachat, Smith (1994); Hoffman, McCabe and Smith (1996); and Slonim and Roth (1998).

7 Conclusions

In this paper we have developed a simple model with limited commitment and have tested it by means of a laboratory experiment. Overall, the experiment provides support for the hypothesis that more commitment opportunities lead to better social outcomes. Indeed, we find that this link is valid when commitment opportunities are asymmetrically distributed between players and even when a reallocation of political power is required to take advantage of new commitment opportunities. However, we also find that, at low levels of commitment, there is more cooperation than strictly predicted by our model, while the opposite is true at high levels of commitment. Furthermore, only large improvements in commitment opportunities have a significant effect on the social surplus, while small changes do not. It seems that the presence of pro-social preferences accounts for some, but not all, of the differences between laboratory results and theoretical predictions.

Our results might have interesting implications for political economy. The bad news is that the experiment suggests that small changes in commitment opportunities might not improve social outcomes. This result could be important for a long-standing debate in political economy about the relative virtues of gradual and radical reforms (see, among others, Popov 2000 and Roland 2000). The good news is that subjects seem to understand that sometimes, in order to take advantage of new commitment opportunities, political power must be reallocated to the agents who can credibly commit. This result could also be important to understand the political economy of reforms. Indeed, some historical reforms can be explained as the outcome of a redistribution of political power between parties with different abilities to commit. For example, Galiani and Torrens (2014) shows that the Repeal of the Corn Laws in Great Britain in 1846 can be understood as the result of a reallocation of political power within the British elite that increases the power of the new pro-free-trade commercial and industrial elite at expense of the old protectionist aristocracy.

References

[1] Acemoglu, Daron, 2003, "Why not a Political Coase Theorem? Social Conflict, Commitment, and Politics", *Journal of Comparative Economics*, 31(4).

[2] Andreoni, James and John H. Miller, 1993, "Rational Cooperation in the Finitely Repeated Prisoner's Dilemma: Experimental Evidence", *The Economic Journal*, 103(418), 570-585.

[3] Andreoni, James and John H. Miller, 2002, "Giving According to GARP: An Experimental Test of the Consistency of Preferences for Altruism", *Econometrica*, 70(2), 737-753.

[4] Bracht, Juergen, and Nick Feltovich, 2008,"Efficiency in the Trust Game: an Experimental Study of Precommitment", *International Journal of Game Theory*, 37, 39-72.

[5] Coase, Ronald, 1960, "The Problem of Social Cost", *Journal of Law and Economics*, 3, 1-44.

[6] Cooper, Russell, Douglas Dejong, Robert Forsythe, and Thomas Ross, 1990, "Selection Criteria in Coordination Games: Some Experimental Results", *American Economic Review*, 80, 218-233.

[7] Camerer, Colin, and Keith Weigelt, 1988, "Experimental Tests of a Sequential Equilibrium Reputation Model", *Econometrica*, 56, 1-36.

[8] Charness, Gary, and Ernan Haruvy, 2002, "Altruism, Equity, and Reciprocity in a Gift Exchange Experiment: An Encompassing Approach", *Games and Economic Behavior*, 40, 203-231.

[9] Druckman, James N., Donald P. Green, James H. Kuklinski and Arthur Lupia (editors), 2011, *Cambridge Handbook of Experimental Political Science*, Cambridge University Press, Cambridge, MA.

[10] Fehr, Ernst, and Simon Gachter, 2000, "Fairness and Retaliation: The Economics of Reciprocity", *Journal of Economic Perspectives*, 14(3), 159181

[11] Fischbacher, Urs, 2007, "z-Tree: Zurich Toolbox for Ready-made Economic Experiments", *Experimental Economics* 10(2), 171-178.

[12] Friedman, Daniel, 2010 "Preferences, Beliefs and Equilibrium: What Have Experiments Taught Us?", *Journal of Economic Behavior & Organization*, 73, 29-33.

[13] Forsythe R, Horowitz JL, Savin NE, Sefton M, 1994 "Fairness in Simple Bargaining Experiments", *Games and Economic Behavior.*

[14] Galiani, Sebastian, and Gustavo Torrens, 2014, "Autocracy, Democracy, and Trade Policy", Fothcoming in *Journal of International Economics.*

[15] Gerber, Alan S., and Donald P. Green, 2012, *Field Experiments: Design, Analysis, and Interpretation.*

[16] Harrison, Glenn W., and Michae McKeel, 1985, "Experimental Evaluation of the Coase Theorem", *Journal of Law and Economics*, 28, 653-670.

[17] Hoffman, Elisabeth, Kevin A. McCabe and Vernon L. Smith, 1994,"On Expectations and the Monetary Stakes in Ultimatum Games", *International Jorunal of Game Theory*, 25(3), 289-301.

[18] Hoffman, Elizabeth, Kevin A. McCabe, Keith Shachat and VernonSmith, 1994, "Preferences, Property Rights, and Anonymity in Bargaining Games," Games and Economic Behavior, Elsevier, vol. 7(3), 346-380.

[19] Hoffman, Elizabeth, and Matthew L. Spitzer, 1982, "The Coase Theorem: Some Experimental Tests", *Journal of Law and Economics*, 25, 73-98.

[20] McAdams, Richard H., 2000 *Experimental Law and Economics.* In vol. 1 of Encyclopedia of Law and Economics, edited by Boudewijn Bouckaert and Gerrit De Geest. Edward Elgar.

[21] McKelvey, Richard M., Palfrey, Thomas R., 1992, "An Experimental Study of the Centipede Game", *Econometrica*, 60(4), 803-836.

[22] Ben-Ner, Avner, and Louis Puttermanb, 2009,"Trust, Communication and Contracts: An Experiment", *Journal of Economic Behavior and Organization* 70, 106-121.

[23] North, Douglass, 1981, *Structure and Change in Economic History.* Norton, New York, NY.

[24] North, Douglass, 1990, "A Transaction Cost Theory of Politics", *Journal of Theoretical Politics* 2(4), 355-367.

[25] North, Douglass and Weingast, Barry, 1989, "Constitution and Commitment: The Evolution of Institutional Governing Public Choice in Seventeenth-Century England", *The Journal of Economic History*, 49(4), 803-832.

[26] Osborne, M. J., and A. Rubinstein, 1994, *A Course in Game Theory*. The MIT Press, Cambridge, MA.

[27] Palfrey, Thomas R., 2009, "Laboratory Experiments in Political Economy",*Annual Review of Political Science*, 12, 379-388

[28] Popov, V., 2000, "Shock Therapy versus Gradualism: The End of the Debate (Explaining the Magnitude of the Transformational Recession)". *Comparative Economic Studies*, 42(1), 1-57.

[29] Roland, G. 2000, *Transition and Economics. Politics, Markets and Firms*. The MIT Press, Cambridge, MA.

[30] Smith, Vernon, 2010, "Theory and Experiment: What are the Questions?", *Journal of Economic Behavior & Organization*, 73, 3-15.

[31] Slonim, Robert and Alvin E. Roth, 1998, "Learning in High Stakes Ultimatum Games: An Experiment in the Slovak Republic", *Econometrica*, 66(3), 569-596.

[32] Van Huyck, John B., Raymond C. Baitalio, and Mary F. Walters, 1995, "Commitment versus Discretion in the Peasant-Dictator Game", *Games and Economic Behavior*, 10, 143-170.

Appendix 1: Proof of Proposition 1

Suppose that $r = 1$. Then, the citizens optimal choice is $e = 1$ and, therefore, the payoffs of the citizen and the ruler will be $v_C = 2.5 - \rho S$ and $v_R = \rho S$, respectively.

Suppose that $r = 0$. In that case, we have to distinguish between two possible situations. If $\pi < \frac{1}{2}$, then the optimal choice for the citizen is $e = 2\pi$. The expected payoffs for the citizen and for the ruler, as a function of the T promised by the ruler, are therefore given by:

$$\mathbf{E}[v_C] = \begin{cases} \pi\left(\frac{1}{2} + 4\pi - T\right) + \frac{1-4\pi^2}{2} & \text{if } T \leq \frac{1}{2} + 2\pi, \\ \frac{1}{2} & \text{if } T > \frac{1}{2} + 2\pi, \end{cases}$$

and

$$\mathbf{E}[v_R] = \begin{cases} \pi T + (1-\pi)\left(\frac{1}{2} + 4\pi\right) & \text{if } T \leq \frac{1}{2} + 2\pi, \\ \frac{1}{2} & \text{if } T > \frac{1}{2} + 2\pi, \end{cases}$$

respectively. Hence, the ruler prefers to promise $T = \frac{1}{2} + 2\pi$, and we therefore have $\mathbf{E}[v_R] = \frac{1}{2} + 4\pi - 2\pi^2$ and $\mathbf{E}[v_C] = \frac{1}{2}$. If $\pi \geq \frac{1}{2}$. In this case, the optimal choice for the citizen is $e = 1$. The expected payoff for the citizen and for the ruler -as a function of the T promised by the ruler- are therefore given by:

$$\mathbf{E}[v_C] = \begin{cases} \pi(2.5 - T) & \text{if } T \leq 2.5 - \frac{1}{2\pi}, \\ \frac{1}{2} & \text{if } T > 2.5 - \frac{1}{2\pi}, \end{cases}$$

and

$$\mathbf{E}[v_R] = \begin{cases} T\pi + (1-\pi)\, 2.5 & T \leq 2.5 - \frac{1}{2\pi}, \\ \frac{1}{2} & \text{if } T > 2.5 - \frac{1}{2\pi}, \end{cases}$$

respectively, Hence, the ruler prefers to promise $T = 2.5 - \frac{1}{2\pi}$ and, therefore, $\mathbf{E}[v_R] = 2.0$ and $\mathbf{E}[v_C] = \frac{1}{2}$.

Finally, we must consider the ruler's relinquish decision. If $\pi < \frac{1}{2}$, the ruler prefers to relinquish if and only if $\rho S > \mathbf{E}[v_R] = \frac{1}{2} + 4\pi - 2\pi^2$. Therefore, if $\rho = 0$, then $r = 0$, $T = \frac{1}{2} + 2\pi$, $e = 2\pi$, $\mathbf{E}[v_R] = \frac{1}{2} + 4\pi - 2\pi^2$ and $\mathbf{E}[v_R] = \frac{1}{2}$; while if $\rho = 1$, then $r = 1$, $S = \frac{1}{2} + 4\pi - 2\pi^2$, $e = 1$, $v_R = \frac{1}{2} + 4\pi - 2\pi^2$ and $v_C = 2 - 4\pi + 2\pi^2$. If $\pi \geq \frac{1}{2}$, the ruler prefers to relinquish if and only if $\rho S > \mathbf{E}[v_R] = 2$. Therefore, if $\rho = 0$, then $r = 0$, $T = 2.5 - \frac{1}{2\pi}$, $e = 1$, $\mathbf{E}[v_R] = 2$ and $\mathbf{E}[v_R] = \frac{1}{2}$; while if $\rho = 1$, then $r = 1$, $S = 2$, $e = 1$, $v_R = \frac{1}{2}$ and $v_C = 2$. This completes the proof of Proposition 1. ■

Appendix 2: Script for General Instructions, PowerPoint Presentation, Quiz, Screens with Rules of the Game, and Questionnaire

In this appendix we present the script for the general instructions, the quiz, and a sample of the rules of the game as observed by the subjects.

Appendix 2.1: Script for General Instructions

We would like to welcome everyone to this experiment. This is an experiment in decision making, and you will be paid for your participation in cash, at the end of the experiment. Different subjects may earn different amounts. What you earn depends partly on your decisions, partly on the decisions of others, and partly on chance.

The entire experiment will be conducted through computer terminals, and all interaction between participants will take place through the computers. Partitions between workstations ensure your anonymity. It is important for you not to talk or to try in any way to communicate with other subjects during the experiments.

In your workstation you will find a pencil, a paper with equations, a paper with a decision tree, and scratch paper. During the experiment you can use the scratch paper to make calculations. You will also find a receipt that we will use to pay you at the end of the experiment.

We will now start with a brief instruction period. During the instruction period, you will be given a complete description of the experiment. If you have any questions during the instruction period, please raise your hand and your question will be answered so everyone can hear. If any difficulties arise after the experiment has begun, raise your hand, and an one of the persons conducting the experiment will come and assist you.

You are one of 16 students who have been randomly assigned to one of two groups. As you can see, all students have been assigned a computer. Each member of Group 1 will be randomly matched with a member of Group 2.

In each round, each pair of students will play the computer game that will appear on the screen, with the member of Group 1 playing the role of player 1 and the member of Group 2 playing the role of player 2. You will be told what your player number is, as it will appear on the screen at the beginning of the game. The matching-up of partners will be repeated after each game is played, so that you will play not every with the same person in two consecutive rounds. At the beginning of each round, the rules of the game will appear on the screen, as well as the timing and payoffs.

The experiment you are participating in is broken down into two unpaid practice rounds and four separate paid rounds. At the end of the last round, you will be paid the total amount you have accumulated during the course of the last four rounds. Everyone will be paid in private and you are under no obligation to tell others how much you earned. Your earnings are denominated in POINTS. Your PESO earnings are determined by multiplying your earnings in points by a conversion rate. In this experiment, the conversion rate is 25 to 1.[17]

Please turn your attention to the screen at the front of the room. We will explain the rules of the game.

[17] 25 Argentine pesos were equivalent to approximately 5.18 U.S dollars at the time.

Appendix 2.2: PowerPoint Presentation

Please find attached the Power Point presentation we employed to explain the rules of the game.

Appendix 2.3: The Quiz

After a general explanation of the rules of the game provided with the help of a PowerPoint presentation, subjects take the following quiz in order to make it possible to gauge their understanding of the rules of the game.

1. Player 2's [citizen] income (Y) before taxes is at least: [5 options]
2. Which of the following levels of effort maximizes player 2's payoff? (not to be confused with income) [5 options]
3. What is the maximum possible payoff for player 1 [ruler]? [5 options]
4. Suppose player 1 has not relinquished her right to impose a tax T and has decided to impose $T = min\{3, Y\}$. Which of the following levels of effort maximizes player 2's payoff? [5 options]
5. Suppose player 1 has not relinquished and has pledged to charge a tax $T = min\{0.4, Y\}$. This promise must be fulfilled with a probability equal to $1/4 = 0.25$. Which of the following maximizes player 2's payoff? [5 options]

Appendix 2.4: Rules of the Game

The following depiction provides a sample of the rules as seen by a subject in the role of the ruler (first practice round, treatment 1).

- **Screen 1**: You are about to play a game with another player. Player 1 (yourself) has the power to select taxes, and Player 2 (your partner) decides what level of effort to put into his or her work.

  Player 2's monetary income is given by: $Y = 0.5 + 2e$, where e is the effort level chosen.

  Effort can take any of the following values: 0, 0.2, 0.5, 0.8 or 1.

  The game also has another important feature: players are able to make promises about taxes and/or transfers.

- **Screen 2.a: The game has four stages (Stage 1: Promise about S)**

  Your partner (Player 2) selects and promises to transfer you $S = min\{S_0, Y\}$, where $S_0 \geq 0$. This promise will be relevant in the last stage of the round in case you have decided to give up your power to tax. In this case, your partner will have the chance to decide what amount to transfer you in the last stage. *Her promise is not enforceable.*

- **Screen 2.b: The game has four stages (Stage 2: Relinquish Decision)**

  You decide whether to relinquish ($r = 1$) or not ($r = 0$) your power to tax.

  If you relinquish, your partner will make you a transfer at the end of the round.

  If you do not relinquish, you have to select and promise to set a tax $T = min\{T_0, Y\}$, where $T_0 \geq 0$. This promise will be relevant in the last stage of the round. In this situation, two things can happen:

  1. with a probability $= 1/4 = 0.25$, your partner will be charged the amount T that you have promised, or
  2. with a probability $= 3/4 = 0.75$, you will be asked to select a new tax $T = min\{T_0, Y\}$, where $T_0 \geq 0$ to charge your partner in the last stage.

- **Screen 2.c: The game has four stages (Stage 3: Working Decision)**:

  If you decide to relinquish ($r = 1$), then your partner will select a level of effort (e) and a new transfer $S = min\{S_0, Y\}$, where $S_0 \geq 0$ which will be issued in the last stage.

  If you decide not to relinquish ($r = 0$), then your partner will select a level of effort (e). Then we will let you know whether your promise is enforceable. If not, you will be asked to select a new tax $T = min\{T_0, Y\}$, where $T_0 \geq 0$ to charge your partner.

- **Screen 2.d: The game has four stages (Stage 4: Payments)**:

  Player 1: $rS + (1 - r)T$

  Player 2: $Y - rS - (1 - r)T + \frac{1-e^2}{2}$

- **Screen 3: Probability of Enforcement**:

  Promises about T: 0.25

  Promises about S: 0

Appendix 2.5: The Questionnaire

Thank you for participating in this experiment! Please complete the following questionnaire before leaving.

Question 1: Gender (male/female)

Question 2: Age (in years)

Question 3: Nationality

Question 4: Fluent in English

Question 5: Racial Group (White/Black/White (Hispanic)/Asian)

Question 6: Have you ever taken a course in game theory / microeconomics? (Yes/No)

Question 7: Current Studies (Graduate/Undergraduate)

Question 8: Current Studies (Junior/Senior)

Appendix 3: Tables and Figures

In this Appendix we present all the tables and figures.

Table 1: Treatments and Predicted Outcomes

	$TSur$ (1)	e (2)	r (3)	$\mathbf{E}[v_R]$ (4)	$\mathbf{E}[v_C]$ (5)	T_{Prom} (6)	S_{Prom} (7)
T1 (0,0)	1.000	0.00	0	0.500	0.500	0.500	Any
T2 (0.25,0)	1.875	0.50	0	1.375	0.500	1.000	Any
T3 (0.75,0)	2.500	1.00	0	2.000	0.500	1.830	Any
T4 (0,1)	2.500	1.00	1	0.500	2.000	Any	0.500

Note: Tj (π, ρ) indicates Treatment j for which the ruler's promises are enforced with probability π and the citizen's promised are enforced with probability ρ; Column (1): Total surplus $TSur$; Column (2): effort e; Column (3): relinquish decision r; Column (4): Expected payoff of the ruler $\mathbf{E}[v_R]$; Column (5): Expected payoff of the citizen $\mathbf{E}[v_C]$; Column (6): Tax that the ruler promises to charge T_{Prom}; Column (7): Transfer that the citizen promises to pay if the ruler relinquish S_{Prom}.

Table 2: Balance Across Player Role (Rulers vs. Citizens)

	All Subjects			Ruler			Citizen				
	N	Mean	S.d	N	Mean	S.d	N	Mean	S.d	Dif	p
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Characteristics of Subjects											
Gender (male=1)	160	0.61	0.49	80	0.66	0.48	80	0.56	0.50	0.10	0.197
Age	160	20.53	2.43	80	20.44	2.49	80	20.61	2.37	-0.17	0.650
Nationality (Argentine=1)	160	0.93	0.26	80	0.93	0.27	80	0.93	0.27	0.00	1.000
Fluent in English (=1)	160	0.99	0.08	80	1.00	0.00	80	0.99	0.11	0.01	0.319
Race (White=1)	160	0.75	0.43	80	0.73	0.45	80	0.78	0.42	-0.05	0.468
Studied Game Theory (=1)	160	0.51	0.50	80	0.48	0.50	80	0.55	0.50	-0.07	0.346
First Half of Undergraduate Studies (=1)	160	0.47	0.50	80	0.51	0.50	80	0.43	0.50	0.08	0.270
First Half of Graduate Studies (=1)	160	0.01	0.11	80	0.03	0.16	80	0.00	0.00	0.03	0.157
Understanding of the Experiment											
Quiz-Mark	160	10.91	2.96	80	11.06	2.97	80	10.76	2.97	0.30	0.523
Answered correctly: question 1	160	0.89	0.31	80	0.90	0.30	80	0.89	0.32	0.01	0.799
Answered correctly: question 2	160	0.81	0.40	80	0.84	0.37	80	0.78	0.42	0.06	0.320
Answered correctly: question 3	160	0.89	0.31	80	0.89	0.32	80	0.90	0.30	-0.01	0.799
Answered correctly: question 4	160	0.78	0.42	80	0.81	0.39	80	0.74	0.44	0.07	0.259
Answered correctly: question 5	160	0.27	0.44	80	0.25	0.44	80	0.29	0.46	-0.04	0.595

Note: N is the number of observations, Mean is the sample mean and S.d is the standard deviation for the corresponding variable in each line. Entries in columns (1)-(3) indicate the values for the complete sample, in columns (4)-(6) for the subjects that played the role of rulers, and in columns (7)-(9) for the subjects that played the role of citizens. Entries in column (10) indicate the mean difference between subjects that play the role of rulers and those that played the role of citizens, while column (11) shows the p-value of the difference of means test (between subjects that played the role of rulers and subjects that played the role of citizens). * indicates that the test is significant at 10%; ** significant at 5%; *** significant at 1%.

Table 3: Balance Across Treatments I

	All Subjects			T1		T2		T3		T4	
	N	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.	Mean	S.d.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Characteristics of Subjects											
Gender (male=1)	160	0.61	0.49	0.73	0.45	0.70	0.46	0.50	0.51	0.53	0.51
Age	160	20.53	2.43	20.90	2.77	20.20	1.99	20.63	2.68	20.38	2.22
Nationality (Argentine=1)	160	0.93	0.26	0.88	0.33	0.88	0.33	0.98	0.16	0.98	0.16
Fluent in English (=1)	160	0.99	0.08	0.98	0.16	1.00	0.00	1.00	0.00	1.00	0.00
Race (White=1)	160	0.75	0.43	0.68	0.47	0.80	0.41	0.83	0.38	0.70	0.46
Studied Game Theory (=1)	160	0.51	0.50	0.57	0.50	0.50	0.51	0.40	0.50	0.57	0.50
First Half of Undergraduate Studies (=1)	160	0.47	0.50	0.40	0.50	0.55	0.50	0.43	0.50	0.50	0.51
First Half of Graduate Studies (=1)	160	0.01	0.11	0.03	0.16	0.00	0.00	0.03	0.16	0.00	0.00
Understanding of the Experiment											
Quiz-Mark	160	10.91	2.96	10.5	2.88	10.80	3.52	11.18	2.72	11.18	2.72
Answered correctly: question 1	160	0.89	0.31	0.93	0.27	0.90	0.30	0.90	0.30	0.85	0.36
Answered correctly: question 2	160	0.81	0.40	0.75	0.44	0.88	0.33	0.78	0.42	0.83	0.38
Answered correctly: question 3	160	0.89	0.31	0.88	0.33	0.85	0.36	0.93	0.27	0.93	0.27
Answered correctly: question 4	160	0.78	0.42	0.70	0.46	0.73	0.45	0.85	0.36	0.83	0.38
Answered correctly: question 5	160	0.27	0.44	0.25	0.44	0.25	0.44	0.28	0.45	0.30	0.46

Note: N is the total number of observations, Mean is the sample mean and S.d. is the standard deviation for the corresponding variable in each line. Entries in columns (1)-(3) indicate the values for the complete sample, in columns (4)-(5) for the subjects that played treatment 1, columns (6)-(7) for those that played treatment 2, columns (8)-(9) for those that played treatment 3, and columns (10)-(11) for those that played treatment 4. Note that There were 40 observations of each of variable in each treatments.

Table 4: Balance Across Treatments II

	T1/T2 (1)	T2/T3 (2)	T3/T4 (3)	T1/T3 (4)	T1/T4 (5)	T2/T4 (6)
Characteristics of Subjects						
Gender (male=1)	0.808	0.069*	0.826	0.039**	0.066*	0.111
Age	0.198	0.423	0.650	0.653	0.352	0.711
Nationality (Argentine=1)	1.000	0.092*	1.000	0.092*	0.092*	0.092*
Fluent in English (=1)	0.320	1.000	1.000	0.320	0.320	1.000
Race (White=1)	0.209	0.778	0.194	0.124	0.812	0.308
Studied Game Theory (=1)	0.507	0.375	0.120	0.120	1.000	0.507
First Half of Undergraduate Studies (=1)	0.148	0.269	0.507	0.823	0.375	0.659
First Half of Graduate Studies (=1)	0.320	0.320	0.320	1.000	0.320	1.000
Understanding of the Experiment						
Quiz-Mark	0.678	0.595	1.000	0.284	0.284	0.595
Answered correctly: question 1	0.697	1.000	0.505	0.697	0.294	0.505
Answered correctly: question 2	0.156	0.245	0.582	0.796	0.419	0.537
Answered correctly: question 3	0.749	0.294	1.000	0.462	0.462	0.294
Answered correctly: question 4	0.808	0.176	0.765	0.111	0.194	0.290
Answered correctly: question 5	1.000	0.802	0.808	0.802	0.622	0.622

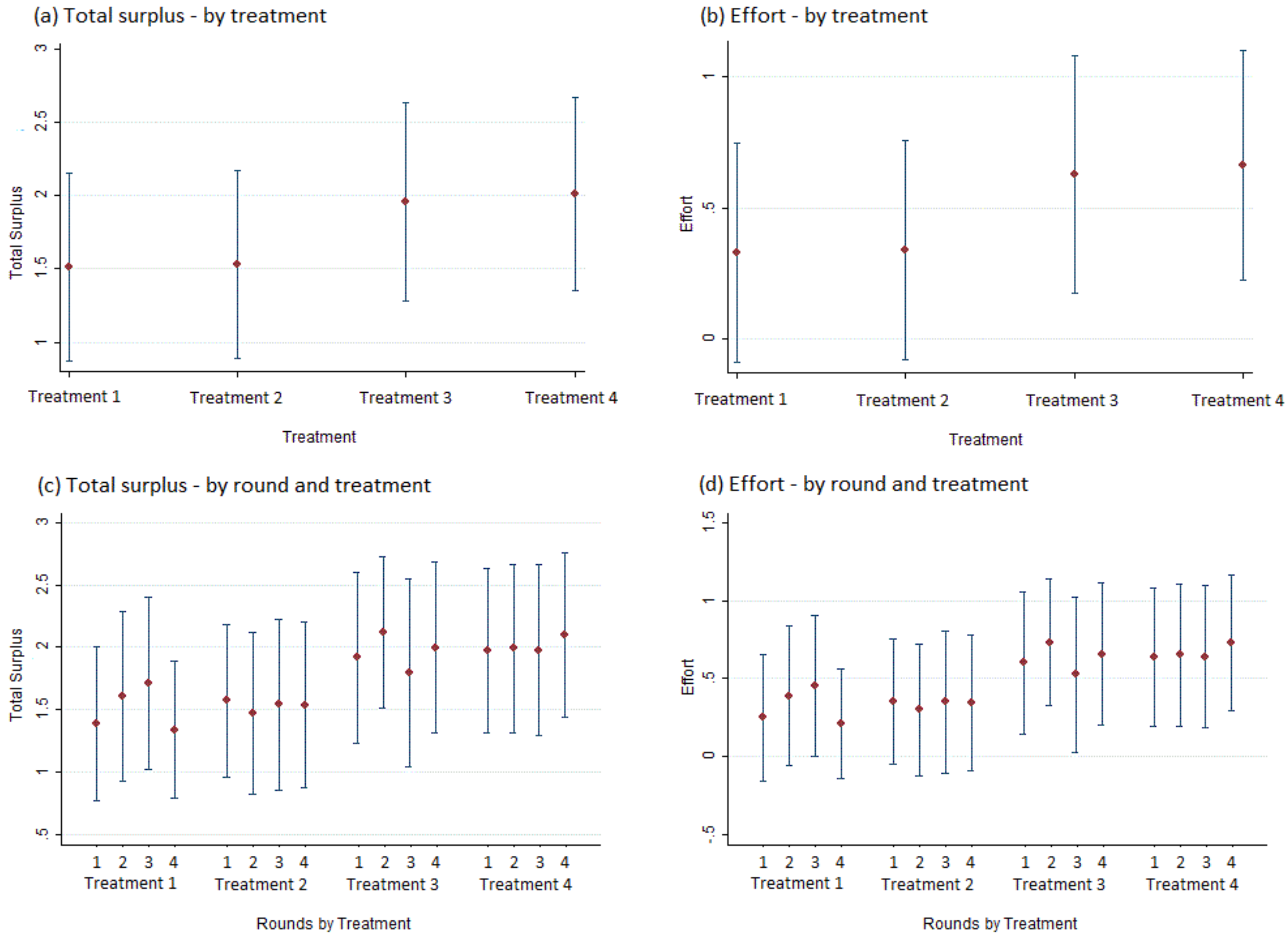
Note: Each entry indicates the mean difference between the two treatments in the column for the corresponding variable in each line. * indicates that the difference of means test is significant at 10%; ** significant at 5%; *** significant at 1%.

Table 5: Decisions Across Treatments (Descriptive Statistics)

	$TSur$	e	r	v_R	v_C	T_{Prom}	S_{Prom}	T_{Decid}	S_{Decid}	T	S
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
T1											
Model Prediction	1.000	0.00	0	0.500	0.500	0.500	Any				
N	80	80	80	80	80	72	80	72	8	72	8
Mean	1.513	0.326	0.100	0.940	0.573	0.887	7.650[a]	1.771	0.594	1.771	0.594
S.d.	0.642	0.420	0.302	0.735	0.532	0.736	28.797[a]	1.992	0.696	1.992	0.696
T2											
Model Prediction	1.875	0.50	0	1.375	0.500	1.000	Any				
N	80	80	80	80	80	77	80	58	3	77	3
Mean	1.531	0.338	0.038	1.037	0.494	1.128	1.625	1.835	0.417	1.658	0.417
S.d.	0.639	0.421	0.191	0.794	0.410	0.630	0.737	1.274	0.722	1.168	0.722
T3											
Model Prediction	2.500	1.00	0	2.000	0.500	1.830	Any				
N	80	80	80	80	80	78	80	24	2	78	2
Mean	1.957	0.628	0.025	1.236	0.722	1.424	1.561	2.138	0.000	1.656	0.000
S.d.	0.679	0.453	0.157	0.731	0.550	0.389	0.730	0.646	0.000	0.595	0.000
T4											
Model Prediction	2.500	1.00	1	0.500	2.000	Any	0.500				
N	80	80	80	80	80	39	80	39	NA	39	41
Mean	2.012	0.664	0.513	1.109	0.903	1.367	0.970	1.612	NA	1.612	1.191
S.d.	0.657	0.440	0.503	0.545	0.496	3.918	0.404	1.110	NA	1.110	0.302

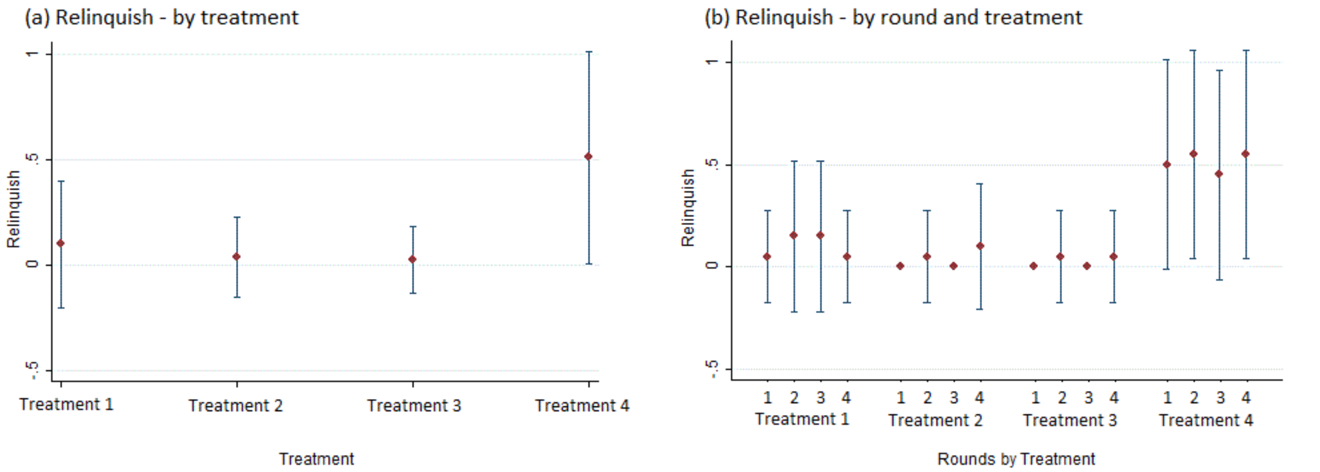
Note: Column (1): Total surplus $TSur$; Column (2): effort e; Column (3): relinquish decision r; Column (4): Payoff of the ruler v_R; Column (5): Payoff of the citizen v_C; Column (6): Tax that the ruler promises to charge T_{Prom}; Column (7): Transfer that the citizen promises to pay if the ruler relinquish S_{Prom}; Column (8) The tax that the ruler sets if his/her promise ends up being non enforceable T_{Decid}; Column (9): The transfer that the citizen sets if his/her promise ends up being non enforceable S_{Decid}; Column (10): The tax that the ruler actually charges T; Column (11): The transfer that the citizen actually pays S. For each treatment N indicates total number of observations, Mean is the sample mean and S.d. is the standard deviation for the corresponding variable in each column. NA indicates non applicable (Note S_{Decid} for T4 is NA because promises of citizens are enforceable with certainty). To facilitate comparison we also repeat the model prediction for each treatment (in the case of the payoffs, the model prediction refers to the expected payoffs). a In T1 there was one subject that selected extremely high values of S_{Prom}. If we take this subject out of the sample the unconditional mean and standard deviation of S_{Prom} are 1.473 and 0.631, respectively.

Figure 1: Total Social Surplus and Effort by Treatment and Round



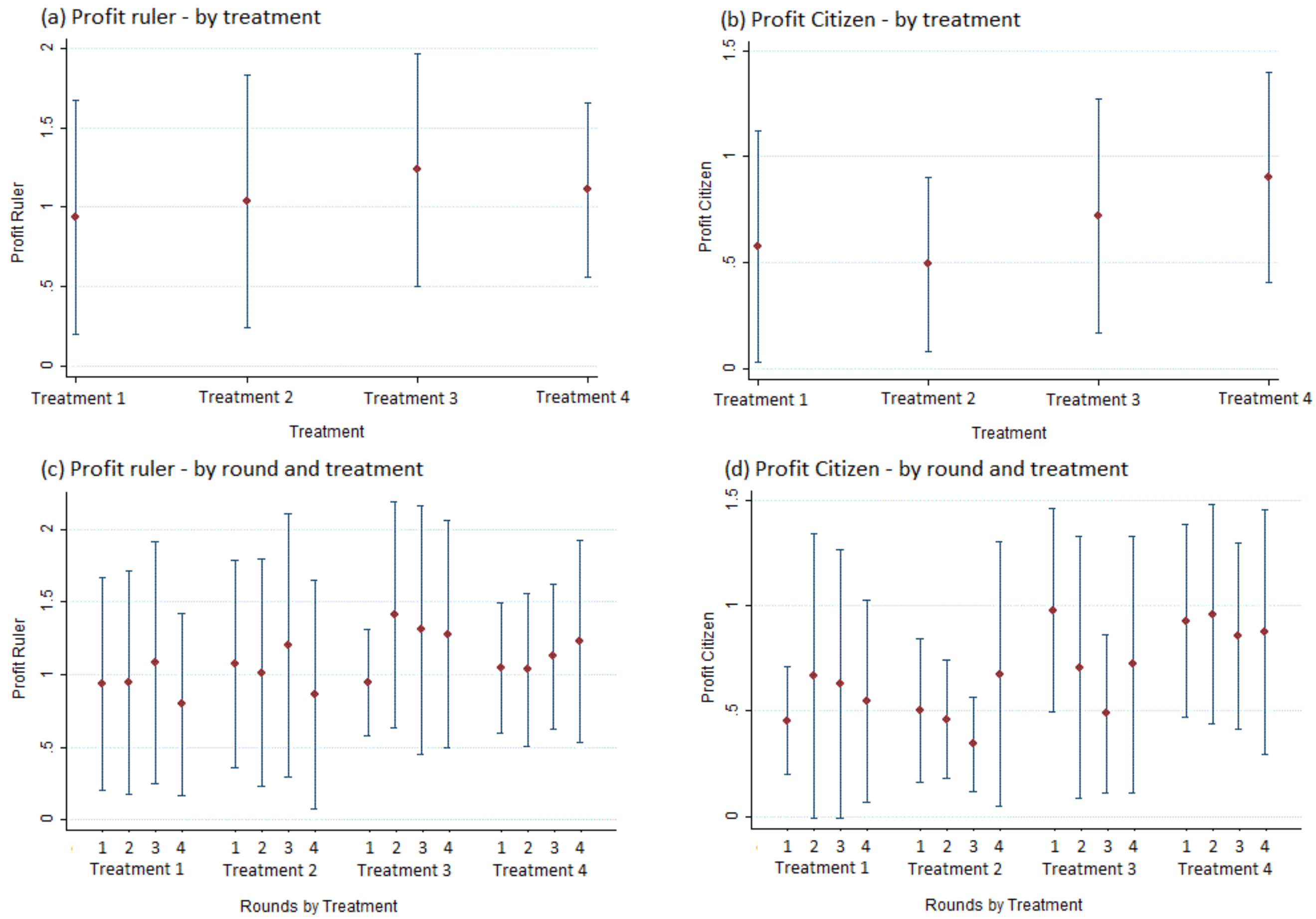
Note: Red diamonds denote average value of the variable per treatment/round within treatment. Blue bars indicate one standard deviation from the mean, calculated in standard form.

Figure 2: Relinquish Decision by Treatment and Round



Note: Red diamonds denote average value of the variable per treatment/round within treatment. Blue bars indicate one standard deviation from the mean, calculated in standard form.

Figure 3: Payoffs by Treatment and Round



Note: Red diamonds denote average value of the variable per treatment/round within treatment. Blue bars indicate one standard deviation from the mean, calculated in standard form.

Table 6: Regression Analysis: Total Surplus

	(1)	(2)
Treatment 1 (=0) vs Treatment 2 (=1)		
$\hat{\beta}_1$	0.018	0.012
S.e. clustered by pair of subjects	0.042	0.127
S.e. clustered by group of 4 subjects	0.130	0.101
R-squared	0.000	0.130
Treatment 1 (=0) vs Treatment 3 (=1)		
$\hat{\beta}_1$	0.444***	0.473***
S.e. clustered by pair of subjects	0.124	0.160
S.e. clustered by group of 4 subjects	0.117	0.159
R-squared	0.103	0.201
Treatment 1 (=0) vs Treatment 4 (=1)		
$\hat{\beta}_1$	0.499***	0.530***
S.e. clustered by pair of subjects	0.109	0.127
S.e. clustered by group of 4 subjects	0.117	0.108
R-squared	0.130	0.292
Treatment 2 (=0) vs Treatment 3 (=1)		
$\hat{\beta}_1$	0.426***	0.402***
S.e. clustered by pair of subjects	0.123	0.143
S.e. clustered by group of 4 subjects	0.123	0.118
R-squared	0.096	0.282
Treatment 2 (=0) vs Treatment 4 (=1)		
$\hat{\beta}_1$	0.481***	0.636***
S.e. clustered by pair of subjects	0.107	0.131
S.e. clustered by group of 4 subjects	0.123	0.121
R-squared	0.122	0.291
Treatment 3 (=0) vs Treatment 4 (=1)		
$\hat{\beta}_1$	0.054	0.105
S.e. clustered by pair of subjects	0.118	0.127
S.e. clustered by group of 4 subjects	0.109	0.109
R-squared	0.002	0.136
Controls	NO	YES
N	160	160

Note: * significant at 10%; ** significant at 5%; *** significant at 1% (using standard errors clustered by pair of subjects). Controls: (i) Individual characteristics X_{gps}: for both subjects in group g gender, age, nationality, racial group, whether s/he has ever taken a course in game theory, whether s/he is a graduate or a junior or senior undergraduate student; (ii) Level of understanding of the game Q_{gps}: mean quiz-mark per group and quiz-mark of the citizen in each group; and (iii) Fixed effects by session $D\theta_s$.

Table 7: Regression Analysis: Reallocation of Political Power

	(1)	(2)
Treatment 1 (=0) vs Treatment 2 (=1)		
$\hat{\gamma}_1$	-0.063	-0.037
S.e. clustered by pair of subjects	0.042	0.045
S.e. clustered by group of 4 subjects	0.048	0.038
R-squared	0.015	0.142
Treatment 1 (=0) vs Treatment 3 (=1)		
$\hat{\gamma}_1$	-0.075*[a]	-0.097**
S.e. clustered by pair of subjects	0.044	0.045
S.e. clustered by group of 4 subjects	0.050	0.045
R-squared	0.024	0.229
Treatment 1 (=0) vs Treatment 4 (=1)		
$\hat{\gamma}_1$	0.413***	0.425***
S.e. clustered by pair of subjects	0.066	0.074
S.e. clustered by group of 4 subjects	0.069	0.062
R-squared	0.200	0.318
Treatment 2 (=0) vs Treatment 3 (=1)		
$\hat{\gamma}_1$	-0.013	-0.025
S.e. clustered by pair of subjects	0.033	0.023
S.e. clustered by group of 4 subjects	0.031	0.021
R-squared	0.001	0.295
Treatment 2 (=0) vs Treatment 4 (=1)		
$\hat{\gamma}_1$	0.475***	0.520***
S.e. clustered by pair of subjects	0.059	0.079
S.e. clustered by group of 4 subjects	0.056	0.066
R-squared	0.283	0.366
Treatment 3 (=0) vs Treatment 4 (=1)		
$\hat{\gamma}_1$	0.488***	0.571***
S.e. clustered by pair of subjects	0.060	0.050
S.e. clustered by group of 4 subjects	0.058	0.039
R-squared	0.302	0.434
Controls	NO	YES
N	160	160

Note: * significant at 10%; ** significant at 5%; *** significant at 1% (using standard errors clustered by pair of subjects). [a] not significant with s.e. clustered by group. Controls: (i) Individual characteristics X_{gps}: for both subjects in group g gender, age, nationality, racial group, whether s/he has ever taken a course in game theory, whether s/he is a graduate or a junior or senior undergraduate student; (ii) Level of understanding of the game Q_{gps}: mean quiz-mark per group and quiz-mark of the citizen in each group; and (iii) Fixed effects by session $D\theta_s$.

Table 8: Regression Analysis: Payoff of the Ruler

	(1)	(2)
Treatment 1 (=0) vs Treatment 2 (=1)		
$\hat{\eta_1}$	0.097	0.088
S.e. clustered by pair of subjects	0.142	0.152
S.e. clustered by group of 4 subjects	0.167	0.126
R-squared	0.004	0.157
Treatment 1 (=0) vs Treatment 3 (=1)		
$\hat{\eta_1}$	0.296**^{a}	0.295*^{c}
S.e. clustered by pair of subjects	0.134	0.174
S.e. clustered by group of 4 subjects	0.162	0.188
R-squared	0.040	0.211
Treatment 1 (=0) vs Treatment 4 (=1)		
$\hat{\eta_1}$	0.169^b	0.187^b
S.e. clustered by pair of subjects	0.117	0.138
S.e. clustered by group of 4 subjects	0.137	0.147
R-squared	0.017	0.211
Treatment 2 (=0) vs Treatment 3 (=1)		
$\hat{\eta_1}$	0.199^b	0.195^b
S.e. clustered by pair of subjects	0.140	0.141
S.e. clustered by group of 4 subjects	0.170	0.119
R-squared	0.017	0.247
Treatment 2 (=0) vs Treatment 4 (=1)		
$\hat{\eta_1}$	0.723	0.218^b
S.e. clustered by pair of subjects	0.124	0.142
S.e. clustered by group of 4 subjects	0.146	0.152
R-squared	0.003	0.176
Treatment 3 (=0) vs Treatment 4 (=1)		
$\hat{\eta_1}$	-0.126	-0.152^b
S.e. clustered by pair of subjects	0.115	0.116
S.e. clustered by group of 4 subjects	0.139	0.103
R-squared	0.010	0.184
Controls	NO	YES
N	160	160

Note: * significant at 10%; ** significant at 5%; *** significant at 1% (using standard errors clustered by pair of subjects). a significant at 10% with s.e. clustered by group. b significant at 10% in one-sided test with s.e. clustered by pair. c not significant with s.e. errors clustered by group. Controls: (i) Individual characteristics X_{gps}: for both subjects in group g gender, age, nationality, racial group, whether s/he has ever taken a course in game theory, whether s/he is a graduate or a junior or senior undergraduate student; (ii) Level of understanding of the game Q_{gps}: mean quiz-mark per group and quiz-mark of the citizen in each group; and (iii) Fixed effects by session $D\theta_s$.

Table 9: Regression Analysis: Payoff of the Citizen

	(1)	(2)
Treatment 1 (=0) vs Treatment 2 (=1)		
$\hat{\eta_1}$	-0.079	-0.076
S.e. clustered by pair of subjects	0.081	0.068
S.e. clustered by group of 4 subjects	0.096	0.019
R-squared	0.007	0.144
Treatment 1 (=0) vs Treatment 3 (=1)		
$\hat{\eta_1}$	0.149	0.178**[c]
S.e. clustered by pair of subjects	0.094	0.085
S.e. clustered by group of 4 subjects	0.094	0.038
R-squared	0.019	0.153
Treatment 1 (=0) vs Treatment 4 (=1)		
$\hat{\eta_1}$	0.330***	0.343***[b]
S.e. clustered by pair of subjects	0.086	0.092
S.e. clustered by group of 4 subjects	0.081	0.047
R-squared	0.094	0.216
Treatment 2 (=0) vs Treatment 3 (=1)		
$\hat{\eta_1}$	0.227***	0.207***
S.e. clustered by pair of subjects	0.082	0.068
S.e. clustered by group of 4 subjects	0.074	0.003
R-squared	0.053	0.240
Treatment 2 (=0) vs Treatment 4 (=1)		
$\hat{\eta_1}$	0.408***	0.418***
S.e. clustered by pair of subjects	0.073	0.092
S.e. clustered by group of 4 subjects	0.056	0.006
R-squared	0.169	0.204
Treatment 3 (=0) vs Treatment 4 (=1)		
$\hat{\eta_1}$	0.181**[a]	0.258***[b]
S.e. clustered by pair of subjects	0.087	0.075
S.e. clustered by group of 4 subjects	0.053	0.036
R-squared	0.029	0.124
Controls	NO	YES
N	160	160

Note: * significant at 10%; ** significant at 5%; *** significant at 1% (using standard errors clustered by pair). *a* Significant at 1% with s.e. clustered by group. *b* significant at 10% with s.e. clustered by group. *c* not significant with s.e. clustered by group. Controls: (i) Individual characteristics X_{gps}: for both subjects in group g gender, age, nationality, racial group, whether s/he has ever taken a course in game theory, whether s/he is a graduate or a junior or senior undergraduate student; (ii) Level of understanding of the game Q_{gps}: mean quiz-mark per group and quiz-mark of the citizen in each group; and (iii) Fixed effects by session $D\theta_s$.

Table 10: Beyond Comparative Statics

		$TSur$ (1)	e (2)	v_R (3)	v_C (4)	S_{Prom} (5)	T_{Prom} (6)	S_{Decid} (7)	T_{Decid} (8)	S (9)	T (10)	$Quiz_C$ (11)	$Quiz_R$ (12)
	All (N)	80	80	80	80	80	72	8	72	8	72	80	80
	Mean	1.513	0.326	0.940	0.573	7.650[a]	0.887	0.594	1.771	0.594	1.771	11.550	9.450
	S.d.	0.642	0.420	0.735	0.532	28.797[a]	0.736	0.696	1.992	0.696	1.992	2.574	2.746
T1	$r=0$ (N)	72	72	72	72	72	72	NA	72	NA	72	72	72
	Mean	1.409	0.257	0.978	0.431	8.320[a]	0.887	NA	1.771	NA	1.771	9.375	11.375
	S.d.	0.591	0.383	0.734	0.236	30.301[a]	0.736	NA	1.992	NA	1.992	2.801	2.564
	$r=1$ (N)	8	8	8	8	8	NA	8	NA	8	NA	8	8
	Mean	2.445	0.950	0.594	1.851	1.616	NA	0.594	NA	0.594	NA	10.125	13.125
	S.d.	0.102	0.093	0.696	0.749	0.397	NA	0.696	NA	0.696	NA	2.232	2.232
	All (N)	80	80	80	80	80	77	3	58	3	77	80	80
	Mean	1.531	0.338	1.037	0.494	1.625	1.128	0.417	1.835	0.417	1.658	10.800	10.800
	S.d.	0.639	0.421	0.794	0.410	0.737	0.630	0.722	1.274	0.722	1.168	3.622	3.361
T2	$r=0$ (N)	77	77	77	77	77	77	NA	58	NA	77	77	77
	Mean	1.508	0.322	1.061	0.447	1.607	1.128	NA	1.835	NA	1.658	10.831	10.831
	S.d.	0.631	0.415	0.792	0.319	0.736	0.630	NA	1.274	NA	1.168	3.412	3.679
	$r=1$ (N)	3	3	3	3	3	NA	3	NA	3	NA	3	3
	Mean	2.127	0.733	0.417	1.710	2.083	NA	0.417	NA	0.417	NA	10.000	10.000
	S.d.	0.647	0.462	0.722	0.687	0.722	NA	0.722	NA	0.722	NA	1.732	1.732
	All (N)	80	80	80	80	80	78	2	24	2	78	80	80
	Mean	1.957	0.628	1.236	0.722	1.561	1.424	0.000	2.138	0.000	1.656	10.650	11.700
	S.d.	0.679	0.453	0.731	0.550	0.730	0.389	0.000	0.646	0.000	0.595	3.089	2.113
T3	$r=0$ (N)	78	78	78	78	78	78	NA	24	NA	78	78	78
	Mean	1.943	0.618	1.267	0.676	1.524	1.424	NA	2.138	NA	1.656	11.769	10.846
	S.d.	0.682	0.455	0.712	0.476	0.702	0.389	NA	0.646	NA	0.595	2.095	2.870
	$r=1$ (N)	2	2	2	2	2	NA	2	NA	2	NA	2	2
	Mean	2.500	1.000	0.000	2.500	3.000	NA	0.000	NA	0.000	NA	9.000	3.000
	S.d.	0.000	0.000	0.000	0.000	0.000	NA	0.000	NA	0.000	NA	0.000	0.000
	All (N)	80	80	80	80	80	39	NA	39	41	39	80	80
	Mean	2.012	0.664	1.109	0.903	0.970	1.367	NA	1.612	1.191	1.612	11.250	11.100
	S.d.	0.657	0.440	0.545	0.496	0.404	3.918	NA	1.110	0.302	1.110	2.313	3.033
T4	$r=1$ (N)	41	41	41	41	41	NA	NA	NA	41	NA	41	41
	Mean	2.489	0.990	1.191	1.298	1.191	NA	NA	NA	1.191	NA	11.415	10.756
	S.d.	0.048	0.044	0.302	0.273	0.302	NA	NA	NA	0.302	NA	2.863	2.596
	$r=0$ (N)	39	39	39	39	39	39	NA	39	NA	39	39	39
	Mean	1.510	0.321	1.023	0.487	0.738	1.367	NA	1.612	NA	1.612	10.769	11.769
	S.d.	0.625	0.407	0.712	0.293	0.368	3.918	NA	1.110	NA	1.110	3.207	1.870

Note: Column (1): Total surplus $TSur$; Column (2): effort e; Column (3): Payoff of the ruler v_R; Column (4): Payoff of the citizen v_C; Column (5): Transfer that the citizen promises to pay if the ruler relinquish S_{Prom}; Column (6): Tax that the ruler promises to charge T_{Prom}; Column (7): The transfer that the citizen sets if his/her promise ends up being non enforceable S_{Decid}; Column (8): Tax that the ruler sets if his/her promise ends up being non enforceable T_{Decid}; Column (9): Transfer that the citizen actually pays S; Column (10): Tax that the ruler actually charges T; Column (11): Quiz grade for the citizen $Quiz_C$; and Column (12): Quiz grade for the ruler $Quiz_R$. For each treatment N indicates the number of observations conditional on an specific relinquish decision (all, $r = 0$ and $r = 1$), Mean is the sample mean and S.d. is the standard deviation conditional on the relinquish decision for the corresponding variable in each column. NA indicates non applicable (S_{Decid} for T4 is NA because promises of citizens are enforceable with certainty.) a In T1 there was one subject that selected extremely high values of S_{Prom}. If we take this subject out of the sample the unconditional mean and standard deviation of S_{Prom} are 1.473 and 0.631, respectively; while the mean and standard deviation of S_{Prom} conditional on $r = 0$ are 1.616 0.397, respectively.

INSTRUCTIONS

PCT LAB EXPERIMENT

RULES OF THE GAME

- **Player 1**: has the power to set taxes **T**, but he/she can give up this power in exchange for a transfer **S** from player 2.

- **Player 2**: decides how much **effort** he/she wants to put. Higher effort means more income, but income can be taxed.

FOUR STAGES

Stage 1: Player 2 promises to pay a transfer S

Stage 2: Relinquish decision

- In case of not relinquishing:

  Player 1 promises a tax T.

Stage 3: Working decision

Stage 4: Payoffs

Please, let me know if you have questions.

STAGE 1: Promise about <u>**S**</u>

Player 2 selects a transfer S: He/she promises to pay a transfer S to Player 1 if Player 1 relinquishes.

Promises are **NOT** necessarily binding/enforceable. S is binding with probability ρ.

- For some players, S is always binding (ρ=1)
- For some players, S is never binding (ρ=0)

STAGE 2: Relinquish decision

Player 1 observes the outcome if stage 1, and decides whether to:

Relinquish his power to set taxes: ($r = 1$)

In this case, Player 2 will make a transfer to Player 1.

or…

Not relinquish his power to set taxes: ($r = 0$)

In this case, Player 1 charges a tax on Player 2.

STAGE 2: Not Relinquish

If Player 1 has not relinquished (r = 0), then he promises to charge a tax **T**.

Promises are **NOT** necessarily binding/enforceable. T is binding with probability π.

- With a certain probability π, the promise is binding and Player 1 must do as promised in the last stage.

- With a certain probability $1-\pi$, the promise is NOT binding and Player 1 may tax a different amount in last stage.

STAGE 3: Working decision

Player 2 selects an effort level (**ε = 0 / 0.20 / 0.5 / 0.8 / 1**)

1) If Player 1 relinquished ($r = 1$): Player 2 is either forced to transfer the promised S to player 2, or allowed to make a different transfer, depending on probability ρ.

2) If Player 1 has NOT relinquished ($r = 0$): Player 1 may or may not be forced to tax Player 2 the promised T, depending on probability π.

STAGE 4: PAYOFFS

In this stage, **all players collect their payoffs.**

Player 1: $v_1 = rS + (1-r)T$

Player 2: $v_2 = Y - rS - (1-r)T + \frac{(1-\varepsilon^2)}{2}$

INCOME

$Y = 0.5 + 2\varepsilon$

TAX / TRANSFER

NOT TAXABLE

Note: The values of S and T finally executed MIGHT OR MIGHT NOT coincide with promises in stage 1. Recall that promises about T and S must only be respected with probabilities π and ρ, respectively.

TAX / TRANSFER

Taxes (T) and **transfers (S)** are as follows:

$$T = min\{T_0, Y\}$$
$$S = min\{S_0, Y\}$$

$\boldsymbol{T_0}$ and $\boldsymbol{S_0}$ can be any non-negative number

Y is Player 2's income: ***Y* = 0.50 +2*ε***

ε is the effort that Player 2 puts into his/her work:

(ε = 0, 0.2, 0.5, 0.8, 1).

FINAL PAYMENTS

r = 1 means that Player 1 relinquished.

r = 0 means that Player 1 did NOT relinquish.

Payoff Player 1: $$v_1 = rS + (1 - r)T$$

Player 1 receives either a Transfer or a Tax.

Payoff Player 2: $$v_2 = Y - rS - (1 - r)T + \frac{(1 - \varepsilon^2)}{2}$$

Player 2 receives total income (Y = 0.50 +2ε) less transfers (S) or taxes (T) plas a gain/satisfaction associated to leisure.

PLEASE CAREFULLY check about the following three things:

1) Who is the player that offers a deal and who is the player that must accept or reject it.

2) The probability that a tax promise T in the bargaining stage must be respected (π).

3) The probability that a transfer promise S in the bargaining stage must be respected (ρ).

The first two rounds are unpaid, so that you can familiarize with the game. The following 4 rounds are paid rounds.

EQUATIONS

Payoff Player 1: $v_1 = rS + (1-r)T$

Payoff Player 2: $v_2 = Y - rS - (1-r)T + \frac{(1-\varepsilon^2)}{2}$

$$v_2 = Y - rS - (1-r)T + 0.5 - 0.5(\varepsilon^2)$$

Income (only Player 2): $Y = 0.5 + 2\varepsilon$

Taxes / Transfers: $T = min\{T_0, Y\}$

$$S = min\{S_0, Y\}$$