An experimental inquiry into the effect of yardstick competition on corruption [1]

Angelino Viceisza [**]

International Food Policy Research Institute, 2033 K Street NW, Washington, DC 20006, United States of America, a.viceisza@cgiar.org

Abstract

This study reports theory-testing laboratory experiments on the effect of yardstick competition on corruption. On the incumbent side, yardstick competition acts as a corruption-taming mechanism if the incumbent politician is female. On the voter side, voters focus on the difference between the tax rate in their own jurisdiction and that in another. If the tax rate is deemed unfair compared to the one in another jurisdiction, voters re-elect less. The findings support the claim by Besley and Case (1995) that incumbent behavior and tax setting are tied together through the nexus of yardstick competition. This renders generalizability to these laboratory experiments and addresses some concerns raised by Levitt and List (2007).

Key words: Corruption, Yardstick Competition, Political Agency, Asymmetric and Private Information, Experiments
JEL Codes: C92, D72, D73, D82, H73

[1] I thank Robin Boadway, James C. Cox, Susan K. Laury, Jorge Martinez-Vazquez, Charles N. Noussair, Andreas Ortmann, Ragan A. Petrie, Vjollca Sadiraj, Lise Vesterlund, Mary Beth Walker, Sally Wallace, Frans van Winden and Yongsheng Xu for invaluable comments. I am indebted to Krawee Ackaramongkolrotn, Artidiantun Adji, Roberta Calvet, Urs Fischbacher, Ben Greiner, Daniel Hall, William Holmes, Christine Moloi, Carmit Segal and Amanda Wilsker for their assistance with programming and/or conducting the experiments. I also thank seminar participants at the 2007 World Meeting of the Public Choice Societies, the Federal Reserve Bank of Atlanta, the Georgia State University (GSU) Economics Department and the International Food Policy Research Institute for helpful comments. All errors remain my own. Financial support from the GSU Dissertation Grant Program and the Experimental Economics Center (ExCEN) at GSU is greatly appreciated.
[**]International Food Policy Research Institute, 2033 K Street NW, Washington, DC 20006, USA; Phone: 1-202-862-8196, Fax: 1-202-467-4439.
Email address: a.viceisza@cgiar.org (Angelino Viceisza).
URL: http://www.ifpri.org/srstaff/Viceiszaa.asp (Angelino Viceisza).

1 Introduction

According to the Organization for Economic Co-operation and Development, the Nigerian President recently estimated the total cost of corruption at a quarter of Africa's total income. [1] Nigeria is not alone. As suggested by Transparency International's indices, many countries are coping with this problematic phenomenon. [2] Like other forms of economic behavior, corrupt behavior can be seen as the result of an interaction between an agent's environment and the institution in which such agent makes decisions. [3] Thus, it is interesting to study how institutions affect corrupt behavior. The nature and causes of corruption have been studied in many contexts. [4] Recently, economists have become interested in almost self-correcting mechanisms for political corruption that fall under the category of fiscal decentralization (e.g. Bardhan and Mookherjee, 2005). This study addresses the efficacy of one such mechanism; namely, whether an experimental institution with yardstick competition has an effect on incumbent and voter behavior and if so, in what direction. Theoretically, the link between yardstick competition and corruption has been addressed in several studies including Besley and Smart (2007) and Belleflamme and Hindriks (2005). Empirically, the link between yardstick competition and corruption has been addressed using field data, e.g. Besley and Case (1995), Dincer et al. (2006) and Jonhson (2006). Experimentally, as reviewed by Abbink (2005a) and Dusek et al. (2005), no studies have addressed this question. [5] This study thus forms a contribution in the following ways. First, it provides experimental evidence on the effect of yardstick competition on incumbent and voter behavior and welfare. In other words, we study the efficacy of yardstick competition as a mechanism, e.g. Shleifer (1985) and Holmstrom

[1] The World Bank (1997) defines corruption as: "The abuse of public office for private gain." Jain (2001) proposes a similar definition.

[2] Corruption is termed problematic because there is generally widespread consensus among (social) scientists across many disciplines that corruption is costly to society, since it hinders economic growth and thus promotes poverty and income inequality. See for example Mauro (1995).

[3] There are many ways to define the terms environment and institution. This paper adopts definitions proposed by Smith (1989). The environment consists of the collection of all agents' characteristics; that is, tastes and technology, which in traditional economics are represented by utility or preference functions, resource endowments and production or cost functions. The institution defines the language (messages or actions) of communication. The institution also specifies the order in which economics agents move, or that there is no form (moves are free form), and the rules under which messages become contracts and thus allocations.

[4] Some general survey pieces include Martinez-Vazquez et al. (2007), Aidt (2003), Tanzi (1998) and Bardhan (1997).

[5] Potters et al. (2004) experimentally address the question of collusion under yardstick competition.

(1982). This is relevant both in economics (public economics, industrial organization, mechanism design etc.) and in other fields such as political science. Secondly, we address internal versus external validity of yardstick competition as a mechanism by comparing the results of this study with the seminal study by Besley and Case (1995). This gets to issues of generalizability as discussed in Levitt and List (2007). Third, by conducting a theory-testing experiment, we are able to identify a key area in which the theory can be improved to capture baseline agent behavior. Finally, the study contributes to the experimental literature on signaling. This includes studies on (1) the plausibility of sequential equilibria in signaling games (e.g. Brandts and Holt, 2005), (2) the lemon's phenomenon (e.g Miller and Plott, 1985) and (3) signaling in miscellaneous contexts including voluntary contributions (e.g. Potters et al., 2007), reputation building (Grosskopf and Sarin, 2006) and board composition and behavior (e.g. Gillette et al., 2003). The main findings are the following. The study finds strong evidence of yardstick competition. On the incumbent side, yardstick competition acts as a corruption-taming mechanism if the incumbent politician is female. Also, yardstick competition makes incumbents more aware of their choices, since beliefs matter more in the presence of yardstick competition. On the voter side, voters focus on the difference between the tax rate in their own jurisdiction and the tax rate in another jurisdiction when making their re-election decisions. In particular, as the discrepancy between the tax rates increases, voters are less likely to re-elect. This suggests that voters use yardstick information in a specific manner; namely, the overall difference between tax rates matters. The experiments also find some partial effects that determine corruptibility. On the incumbent side, the unit cost of the public good matters. If the unit cost is high, incumbents are more likely to charge low taxes. Also, gender is important when interacted with the unit cost. Female incumbents are more likely to divert rents and charge maximum taxes when the unit cost is high. Last, incumbents' beliefs towards re-elections are important. Incumbents weigh the likelihood of re-election heavily when choosing tax rates. On the voter side, beliefs are captured by the tax rate, which is a signal about the incumbent's nature. In particular, voters re-elect more frequently as taxes are lower. Also, female voters behave more consistently with the pooling equilibrium in the Besley and Smart (2007) model. Overall, the findings support the seminal study by Besley and Case (1995) in their claim that incumbent behavior (i.e. vote-seeking) and tax setting are tied together through the nexus of yardstick competition. This renders generalizability to the findings of these laboratory experiments, addressing some of the issues raised by Levitt and List (2007). The remainder of the paper is organized as follows. The following section discusses the design of the experiments. Then, we turn to the main treatment effects. Finally, we conclude.

2 Experimental design

2.1 The experimental game

The theoretical model that underlies the experimental design is a variant of Besley and Smart (2007). It is a game-theoretic model of elections that is cast in a principal-agent framework. There are two "active" players, i.e. a principal (the voter) and a first agent (the first-period incumbent) and one "passive" player, i.e. a second agent (the challenger). An agent's type (i) can be good (g) or bad (b).[6] Each incumbent knows his own type; however, the voter and the other incumbent do not. The game comprises two periods. At the end of the first period, there is an election. If the first-period incumbent is re-elected, he remains in office during the second period. On the other hand, if the first-period incumbent is voted out, the challenger is elected to office during the second period. The challenger is "passive" in the sense that he only plays a role during elections and during the second period if elected. This context describes a dynamic game of incomplete information. It is dynamic because it encompasses two periods and decisions made in the first period affect outcomes in the second period. There is incomplete information, since at least one player is uncertain about another player's payoff. Following Harsanyi (1967), Harsanyi (1968a) and Harsanyi (1968b), the game is transformed into a dynamic game of imperfect information by introducing moves of nature that determine types. The baseline experimental game consists of ten repetitions of this one-shot dynamic signaling game. Incumbent and voter subjects face alternatives that are "on the equilibrium path" according to the main proposition (i.e. lemma 1) derived by Besley and Smart (2007).[7] In order to minimize repeated-game effects, subjects are guaranteed that they will not interact with the same player for more than one repetition. The following steps describe one full repetition. In the first period of the repetition:

(1) Nature moves and the incumbent's type is determined to be good or bad.
(2) Nature moves again and the unit cost of public good provision (θ) is determined to be high or low.
(3) The incumbent's type – which is private information – is determined by

[6] We maintain the definitions of good and bad types discussed by Besley and Smart (2007). The good type always behaves non-strategically by setting rent (i.e. corruption) equal to zero. Under these circumstances, both the incumbent and the voter achieve the same level of welfare (i.e. payoffs). This is also the maximum level of welfare for the voter. On the other hand, the bad type always diverts some rent and thus sets his own welfare (i.e. payoff) above that of the voter. We maintain these definitions in order to infer the proportions of types within the subject population.
[7] Besley and Smart (2007) show existence and uniqueness of the sequential equilibria associated with this signaling game. See also Lockwood (2005).

the two components (i, θ), i.e. whether he is by nature good or bad and whether the unit cost of public good provision is by nature high or low.

(4) Having observed his type and knowing his preferences, the incumbent chooses a payoff pair from a set of alternatives. Each alternative represents a tax-public good signal (t, G).

(5) The voter observes his payoff and the range of possible payoffs for the incumbent.

(6) The voter re-elects the incumbent or votes him out.

In the second period of the repetition:

(1) If the incumbent is voted out, he gets a payoff of zero and the challenger's type is determined by a lottery. [8] The result of the lottery determines the voter's second-period payoff.

(2) If the incumbent is re-elected, he is in office again.

(3) The second-period alternatives that he chooses from are affected by another draw of the unit cost of the public good, determined by nature. Payoffs for both players are determined by the incumbent's choice.

(4) There is no election.

These two periods together constitute one repetition. This entire process is repeated ten times. To operationalize this game, some questions of experimental design need to be addressed. First, we need to address the question whether or not types and preferences are to be induced. Voters are allowed to behave according their homegrown preferences, since part of what we want to learn from conducting the experiments is how voters react to different information sets. On the incumbent side, the unit cost of the public good is induced, since this is consistent with the theoretical game. Benevolence, however, is not induced. While the primary aim of the experiment is to test a model of corruption and yardstick competition that rests on the assumption that there is a positive proportion of good types in the population, a secondary aim is to test whether this assumption is empirically valid. In fact, in accordance with parallelism to the naturally occurring world, it is sensible to ask whether this assumption is plausible. Furthermore, since one of the questions we seek to answer is the effect of yardstick competition on incumbent behavior, it would not make much sense to induce types. [9] Secondly, we must address parameterization of the experiment and how this affects the experimental payoff functions. The experimental payoffs are derived from explicit functional forms

[8] Since the challenger is a move of nature and therefore a "passive" player in the theoretical game, the challenger is experimentally operationalized as a lottery.

[9] Allowing incumbents to behave according to their homegrown types has an important implication: Contrary to the assumption that theoretical voters know the exact distribution of types within the incumbent population, the voters in the experiment do not. In particular, voter subjects must create their own beliefs and update accordingly as the game progresses.

of the welfare functions discussed by Besley and Smart (2007). The bad incumbent's per-period payoff function is characterized by $r = t - \theta G_\theta$, where r is rent diversion (i.e. corruption), t is the tax collection, θ is the unit cost of the public good and G_θ is the level of the public good associated with the unit cost. The good incumbent's per-period payoff function (which is the same as the voter's) is characterized by $W(G,t) = G - \mu t$, where G and t are as defined previously and μ is an exogenous parameter that indexes the marginal cost of public funds.[10] [11] Finally, the following parameterizations apply to the experimental treatments discussed below: (1) with equal probability, the unit cost of the public good is high ($\theta = 2$) or low ($\theta = 1$), (2) the marginal cost of public funds is one quarter ($\mu = \frac{1}{4}$) and (3) there are a maximum level of tax collection of $T = 4$ and a minimum level of public good provision of $G_T = 1$ associated with the experimental game.[12]

2.2 Treatments

The experimental design is based on two treatments, a baseline treatment (treatment B) and a yardstick treatment (treatment Y). Treatment B is designed to capture baseline behavior and treatment Y is designed to capture behavior in the presence of yardstick competition. The treatments differ subtly by what is observed by the voter. In treatment Y, the voter observes additional information; namely, another voter's first-period payoff (i.e. tax rate)

[10] Browning (1976) defines the marginal cost of public funds as the direct tax burden plus the marginal welfare cost produced in acquiring the tax revenue.

[11] Besley and Smart (2007) assume that the tax-cost function is strictly convex. Assuming a linear tax-cost function has no particular complications for their results. In particular, the good-type's optimization problem becomes $\max_G [G - \mu\theta G]$, which reduces to $\max_G [G(1 - \mu\theta)]$. If the parameter $(1 - \mu\theta) \geq 0$, then $G^* > 0$. If not, then $G^* = 0$. In order to have an interior solution, we need μ and θ to satisfy the following relation: $\theta \leq \frac{1}{\mu}$. This condition is satisfied in all parameterizations. A linear tax-cost function also leaves the relationships between G and θ and G and μ unchanged. Namely, suppose the condition $\theta \leq \frac{1}{\mu}$ holds. Now, consider a slight increase in either θ or μ such that θ becomes strictly greater than $\frac{1}{\mu}$. Then, G^* goes from strictly positive to zero in a discrete manner. Thus, the negative relationships are maintained.

[12] This requirement is not discussed by Besley and Smart (2007). It is introduced here to avoid negative voter payoffs. By formulation of the voter's welfare function, if the minimum level of the public good is zero (i.e. $G_T = 0$), the voter automatically has negative payoffs when the incumbent diverts maximum rents. So, we introduce this additional restriction in the empirical game. Suppose the minimum level of the public good is equal to δ (i.e. $G_T = \delta$), then as long as δ satisfies the following condition, existence of the sequential equilibria derived by Besley and Smart (2007) is unaffected: $\frac{(1-\beta)T - r_1}{H(1-\beta)} \leq \delta \leq \frac{(1-\beta)T - r_1}{L(1-\beta)}$.

Table 1
First-period alternatives faced by incumbent in treatment B

If Unit Cost is Low ($\theta = 1$)		If Unit Cost is High ($\theta = 2$)	
Alternative	Payoff pair	Alternative	Payoff pair
A_1	$(1.50, 1.50)$	A_1	$(0.75, 0.75)$
A_2	$(1.50, 0.75)$		not applicable*
A_3	$(2.35, 0.16)$	A_2	$(1.35, 0.16)$

**Notes:* There is no comparable alternative.

drawn at random from the distribution of first-period payoffs in the baseline treatment. The experiments are between-subjects in the sense that any given pair of subjects (incumbent-voter) participate in one and only one of these two treatments. Comparisons are then made across treatments to draw conclusions regarding treatment effects. In both treatments, incumbents are faced with the same set of potential choices. Each choice alternative represents a payoff pair. The first component belongs to the incumbent and the second to the voter. These payoff pairs (which are summarized in table 1) are derived using the payoff functions and parameter values discussed previously. [13] If an incumbent chooses alternative 1, we can infer that he is behaving as if he is good, since this alternative leads to equal payoffs since it is based on a zero-rent behavioral assumption. On the other hand, if an incumbent chooses alternative 2 when the unit cost of the public good is low, we can infer that he is playing according to the pooling equilibrium, since he is diverting rents in a restrained manner, in an attempt to fool the voter and gain re-election. Namely, when the voter observes a first-period payoff of 0.75, it is uncertain whether this payoff came from a good type that got a high cost or a bad type that is facing a low cost. Finally, if an incumbent chooses any other alternative, he is playing according to the separating equilibrium since he is pushing the voter to the lowest possible payoffs by diverting maximum rents.

2.3 Implementation

The experiments took place in the experimental laboratory at the Experimental Economics Center (ExCEN) at Georgia State University. Subjects were recruited using ExCEN's online recruiter system, which contains names of students taking courses in many different areas including but not limited to accounting, actuarial science, biology, business administration, chemistry, economics, finance, geology, geography, mathematics, nursing, political sci-

[13] The second-period payoff alternatives are a subset of the first-period alternatives; see subject instructions.

ence and sociology. The experiments were programmed and conducted with the software z-Tree (Fischbacher, 2007). In each experiment session, half of the subjects were incumbents and half of the subjects were voters. Each experiment session had at least twenty subjects in order to guarantee that a given politician and voter were never paired for more than one repetition. In other words, a given politician did not interact with a given voter for more than one repetition and vice versa. So, re-matching during the experiment was predetermined according to the "two ships passing in the night" design. Subjects were informed accordingly. Each experiment session consisted of instructions, a five-minute trial, a quiz, a summary of the treatment, the treatment and a post-questionnaire. The experiments lasted an average of ninety minutes. Average payoffs in the baseline treatments were \$15.22 for the incumbent and \$13.59 for the voter. In the yardstick treatments these were \$15.25 and \$13.08 respectively. Subjects were paid \$5.00 for showing up, \$3.00 for completing the trial and quiz and \$2.00 for completing the post-questionnaire. The procedures during the experiment were as follows. Subjects were assigned a number at sign-in. These numbers were used to randomly enter subjects into the experimental laboratory. Random entry also determined random assignment to a fixed role and different pairs during treatment B or treatment Y. After entering the experimental laboratory, subjects were handed paper-based instructions, which were read out loud by the experimenter. They were then put through the five-minute trial. This was an opportunity for subjects to interact with the software and practice making decisions that did not affect their payoffs. Thereafter, subjects were put through the quiz. Contrary to the trial, the quiz was not timed. Subjects were informed that they would earn three dollars for completing the trial and quiz regardless of how many questions they answered correctly. They were asked to pay attention to the screens observed in the trial and to the questions asked in the quiz. During the quiz, the software informed subjects whether or not they answered a particular question correctly. In both cases, the quiz gave an overview of the correct answer and referred subjects to the instructions. After the quiz was completed, subjects were asked whether they had any questions that they wanted clarified in private. If so, those were clarified. The experimenter then summarized the treatment. In doing so, the experimenter elaborated on the main issues and any particular issues that seemed to be problematic based on the responses in the quiz. The problematic issues were generally uniform across sessions. Once this process was completed, subjects were ready to start the main treatment. At this point, subjects were informed whether they would be Player X (a politician) or Player Z (a voter). Subjects were told to note that the room was divided into two "sections" by means of a blank column of computer stations running from front to back. This blank column divided the room into incumbents and voters. Everyone on one side of the room was randomly assigned to be an incumbent and everyone on the other side was randomly assigned to be a voter. Within each section, subjects were separated by means of dividers. This guaranteed that they could make decisions in private. Furthermore, the

voter side had higher dividers such that it would be impossible for a particular voter to observe the result of the coin toss. This also reinforced the nature of the information asymmetry in the game. Subjects were told not to communicate with each other during the experiment and they were reminded that they would keep the same role throughout the entire treatment. At the beginning of each repetition and period within a repetition, the experimenter flipped a coin. This represented the move by nature that determined the unit cost of the public good. The coin toss took place in front of the first two politicians. They observed and verified the coin toss and its result. Upon verification, the experimenter input the result into the computer and the respective period was conducted. Upon conclusion of the two periods, the process was repeated until all ten repetitions were concluded. After these repetitions, subjects completed the post-questionnaire.

3 Results

3.1 Aggregation and demographics

A total of four experiment sessions were conducted: two B-sessions and two Y-sessions. The main choice (endogenous) variables of interest are first-period choices by incumbents (this variable is named Choice) and re-election (acceptance) decisions by voters (this variable is named Accept). Full distribution tests (Kolmogorov-Smirnov, KS and Epps-Singleton, ES) suggest that there are no statistically significant differences between the sessions for these main variables of interest. [14] Aggregation across experiment sessions leads to the following number of subjects for the respective treatments: treatment B has a total of thirty politicians and thirty voters and treatment Y has a total of twenty eight politicians and twenty eight voters. The following descriptive statistics describe the average subject profile. Female subjects constitute 56.03% of the sample. The average age is 22 years with a standard deviation of 4 years. About 13% of subjects are economics majors, 12% are biology majors, 8% are accounting majors and 4% are political science majors. The remaining subjects comprise miscellaneous majors including film and video, sociology, gerontology, chemistry and english. Finally, the maximum self-reported annual income range is $30,001 to $60,000 with a mode annual income range of $0 to $1,000.

[14] All KS p-values are greater than 0.358. All ES p-values are greater than 0.13.

Table 2
Incumbents' average first-period choices by treatment

	Unit Cost Low		Unit Cost High	
Treatment	Equalize	Not Equalize	Equalize	Not Equalize
B	64.00%	36.00%	66.67%	33.33%
Y	58.27%	41.73%	57.52%	42.48%

Notes: $N_B = 30$, $N_Y = 28$ and $T = 10$.

3.2 Average behavior

3.2.1 Incumbent and voter behavior

To get an impression of the percentage of good (i.e. not corrupt) versus bad (i.e. corrupt) incumbents, we look at the average percentage of incumbents that equalize first-period payoffs within a treatment conditional on the draw of the unit cost of the public good (table 2). Incumbents in treatment Y equalize first-period payoffs at similar rates regardless of the level of the unit cost of the public good. These rates tend to be lower than in treatment B. On average, 64% of subjects equalize first-period payoffs, which suggests that 36% of incumbents are corrupt. This percentage, however, only accounts for first-period choices. If we want to get an estimate of the proportion of good incumbents based on the definition proposed in the Besley and Smart (2007) model, we must consider both first- and second-period payoffs. After all, a good incumbent is defined as one who does not divert rent in either period. Table 3 shows the percentage of incumbents that behave as if bad in the second period conditional on having behaved as if good in the first period. According to the definitions of types in the Besley and Smart (2007) model, these percentages should be zero. The table shows that this is not the case. In fact, the percentage of incumbents that equalize payoffs in the first period and divert maximum rents in the second period are relatively high (i.e. more than 50% in all cases). These findings mitigate empirical validity of the aforementioned definitions and indicate clear strategic behavior on the part of incumbents. We term these choices "theoretically inconsistent" in the context of the Besley and Smart (2007) model, since they signal a particular type in the first period and a different type in the second. It should be noted that this behavior is different from the pooling equilibrium derived by Besley and Smart (2007). Finally, we turn to average voter behavior. Table 4 draws this comparison, where 1.50, 0.75 and 0.16 represent voters' first-period payoffs as discussed in table 1.

Table 3
Theoretically inconsistent choices by treatment

Treatment	Unit Cost Low	Unit Cost High
B	56.94%	68.75%
Y	74.58%	71.67%

Notes: $N_B = 30$, $N_Y = 28$ and $T = 10$.

Table 4
Voters' re-election decisions

Treatment	1.50	0.75	0.16
B	61.89	19.52	11.80
Y	76.99	21.25	15.48

Notes: $N_B = 30$, $N_Y = 28$ and $T = 10$.

3.3 Treatment effects

This section focuses on the main treatment effects. First, we specify the econometric model by expressing the estimation equation and by defining the variables of interest. Then, we perform estimation and inference. Three comments are noteworthy. First, throughout the discussion we focus on random-effects (RE) estimators. This choice is mainly dictated by the fact that all covariates of interest are individually invariant. Thus, if estimation were to occur via FE, it would wipe out the main partial effects (Wooldridge, 2002). Secondly, a *potentially* relevant variable – viz. history – is not completely controlled for in the regressions below. History of play should not be a major explanatory variable in our regressions, since the experiments were carefully designed to avoid repeated-play effects. Nonetheless, we might expect subjects to play fictitiously.[15] If this is the case, we should control for history accordingly. This would require instrumental-variable (IV) estimation, since the history of incumbents' and voters' choices are endogenous. Since there are no proper instruments, we partially control for history by including (1) time and (2) the history of the unit cost (θ_{t-1}) as explanatory variables in the regressions. The time trend is intended to capture any upward or downward trends that may be present in the data. The history of the unit cost acts as an imperfect proxy for history of choices in both the incumbent and voter regressions. [16] Finally,

[15] The modern game-theoretic usage of the term "fictitious play" can be different from the definitions first discussed by Brown (1951) and Robinson (1951). Here we use the term to indicate a situation in which subjects play as if they are paired with the same player every repetition, even though they are told they are not.

[16] The historical value of the unit cost is exogenous, since it is randomly determined by the experimenter by means of a coin flip. Furthermore, it directly impacts incumbents' past first-period choices and thus it indirectly affects the set of signals

since we have designed the experiments according to a specific structure (i.e. a logical game form), we know to a great extent the data generating processes. In particular, since the data are generated according to a *sequential* game of incomplete information by construction (i.e. by design of the experiments), we can rule out with certainty the possibility that decisions are being made simultaneously by the incumbent and the voter. Thus, we estimate separate equations of interest for each treatment – one for the first mover (the incumbent) and one for the second mover (the voter).

3.3.1 Incumbent Behavior

An individual incumbent's (i) main choice variable in any given repetition (t) is his first-period choice, $Choice_{it}$. There are two possible ways to define this variable. The first is to let $Choice_{it}$ be a dummy variable that is equal to one if the incumbent chooses to equalize first-period payoffs (i.e. behave as if good). The second alternative is to let $Choice_{it}$ have a range of three possible values according to equalization, pooling or separation. We report findings based on the first formulation, since it lends itself to easier interpretations of the coefficients given the main question.[17] We estimate the following equation:

$$Choice_{it} = \beta_0 + \beta_1 gen_i + \beta_2 \theta_t + \beta_3 d_Y + \beta_4 bel_{it} + \beta_5 \theta_{t-1} + \beta_6 t + \beta_7 I + \varepsilon_{it}, \quad (1)$$

where $i = 1, ..., 58$, $t = 1, ..., 10$, β_0 represents a constant term, gen_i represents a dummy variable that is equal to one if the individual is female, θ_t represents the current draw of the unit cost of the public good, d_Y represents a dummy variable that is equal to one if the individual is part of the treatment group, θ_{t-1} is the past lag of the unit cost as a proxy for history, t is a time trend, I is a set of interaction terms consisting of $(gen_i * \theta_t)$, $(gen_i * d_Y)$, $(bel_{it} * d_Y)$, $(\theta_{t-1} * d_Y)$, $(gen_i * \theta_t * d_Y)$ and $(\theta_t * \theta_{t-1})$ and ε_{it} has a one-way error component structure of the form $\varepsilon_{it} = \alpha_i + \nu_{it}$, where α_i represents a vector of unobserved individual heterogeneities and ν_{it} satisfies the following strict exogeneity assumption, $E(\nu_{it}|X_{it}, \alpha_i) = 0$, where X_{it} is the set of explanatory variables in expression 1. We also include a proxy variable for beliefs, bel_{it},

---

that were observable by the voter in the previous period. So, in this sense, the history of the unit cost satisfies both the conditions for an IV or proxy variable (Wooldridge, 2002). However, contrary to the history of incumbents' and voters' choices, the history of the unit cost is not individually variant. In other words, the history of the unit cost is imperfect either as a proxy variable or as an IV. So, the question becomes whether to use this variable as a weak instrument or as an imperfect proxy. We choose the latter approach, since we expect past choices to have relatively little effect on incumbent and voter behavior due to the re-pairing design discussed previously.

[17] The main treatment effects are robust to the alternative formulation.

on the right-hand side of equation 1. It is expected that when an incumbent makes a first-period choice, he anticipates a particular re-election decision to follow such choice. In other words, an incumbent has some belief towards re-election that is associated with taxes charged. This belief is likely to influence his first-period choice. So, beliefs can be a statistically significant explanatory variable in the incumbent regressions. Since beliefs are an inherent characteristic of the incumbent, they are unobserved. Furthermore, there is no reason to expect beliefs to be correlated with any of the other explanatory variables in equation 1. So, the correct approach to control for beliefs is to use a proxy variable that satisfies the following conditions: (1) irrelevance for explaining the dependent variable and (2) correlation with beliefs (Wooldridge, 2002). We claim that $Accept_{it}$ (i.e. voters' actual re-election decisions) constitutes a valid proxy for beliefs. First, this variable satisfies requirement (1) since it is chosen by the voter and thus, it should not be relevant in explaining the incumbent's choice. Secondly, the variable satisfies requirement (2) since it is highly correlated with beliefs. This is testable. Under the assumption that the regressors in equation 1 are exogenous and capture all other unexplained variation in $Choice_{it}$, the least-squares residuals from the estimation of equation 1 represent consistent estimates of the error term ε_{it} and thus of beliefs. A regression of these estimates on $Accept_{it}$ finds that this variable is strongly significant (p-value 0.000). Table 5 reports the estimates of a linear probability model (LPM). [18] In general, incumbents behave in a more corrupt manner in the presence of yardstick competition; however, female incumbents are more likely to equalize first-period payoffs in the presence of yardstick competition. Furthermore, female incumbents are less likely to equalize when the unit cost is high. Finally, beliefs are an important determinant of incumbents' choices.

3.3.2 Voter Behavior

An individual voter's (i) main choice variable in any given repetition (t) is his re-election decision, $Accept_{it}$, which is a dummy variable that is equal to one if the voter re-elects (accepts) the incumbent. We estimate the following equation:

$$Accept_{it} = \beta_0 + \beta_1 gen_i + \beta_2 pool_{it} + \beta_3 yardstick_{it} + \beta_4 \theta_{t-1} + \beta_5 I + \varepsilon_{it}, \quad (2)$$

where $i = 1, ..., 28$, $t = 1, ..., 10$, β_0 represents a constant term, gen_i represents a dummy variable that is equal to one if the individual is female, $pool_{it}$ represents a dummy variable that is equal to one if the voter observes a pooling first-period payoff, θ_{t-1} is the past lag of the unit cost as a proxy for history, I is an interaction term ($gen_i * pool_{it}$) and ε_{it} has a one-way error component structure of the form $\varepsilon_{it} = \alpha_i + \nu_{it}$, where α_i represents a vector of unobserved

[18] The results are robust to alternative estimation methods (i.e. logit and probit).

Table 5
Incumbent Regression (Treatment Y)

Dependent Variable: $Choice_{it}$ (1=equalize, 0=not equalize)	
$constant$	0.485 (0.039)
gen_i	0.112 (0.421)
θ_t	0.119 (0.343)
d_Y	-0.313 (0.023)
$beliefs_{it}$	0.372 (0.000)
$history$ (θ_{t-1})	-0.097 (0.441)
t	-0.004 (0.523)
$gen_i * \theta_t$	-0.146 (0.077)
$gen_i * d_Y$	0.330 (0.052)
$beliefs_{it} * d_Y$	0.145 (0.049)
$\theta_{t-1} * d_Y$	-0.008 (0.915)
$gen_i * \theta_t * d_Y$	-0.056 (0.561)
$\theta_t * \theta_{t-1}$	0.037 (0.618)
overall R^2	0.2883
Notes: p-values in parentheses. $N = 58$ and $T = 10$.	

individual heterogeneities and ν_{it} satisfies the following strict exogeneity assumption, $E(\nu_{it}|X_{it}, \alpha_i) = 0$, where X_{it} is the set of explanatory variables in expression 2. There are some subtleties in equation 2. First, since the yardstick sessions are identical to the baseline sessions except for one minor modification, the immediate question arises how to model this subtle change if we want to compare the baseline and yardstick data. One alternative is to include a dummy that equals one if the voter subject is in the treatment group. This formulation leads to the conclusion that yardstick competition has no effect. Another alternative is to define a variable that takes into account the nature of the yardstick information. In this case we are unable to collapse the data from treatment B and treatment Y, since any yardstick variable that incorporates the nature of the yardstick signal is undefined in treatment B. In other words, if we want to appeal to the yardstick nature of the Y-sessions by defining a yardstick variable, the above regression will be based only on data from the Y-sessions (i.e. $N = 28$). So, the question we ask in this formulation is whether the yardstick variable captures any treatment effects given the assumption that treatments B and Y are otherwise identical. This assumption is reasonable, since the baseline and yardstick treatments were designed to be identical except for the presence of yardstick information. So, if equation 2 controls

for the same set of covariates as a baseline regression, the yardstick variable should capture consistent treatment effects. The second question is how to define the yardstick variable. The findings are robust to two alternative definitions. In particular, treatment effects arise whether we define the yardstick variable as (1) the difference between the tax rate in one's own jurisdiction and the tax rate in the other jurisdiction or (2) a dummy that equals one if the tax rates are the same. We report results for the former. We estimate equation 2 for three separate cases (when the signal is 1.50, 0.75 and 0.16). Each estimation is based on a subset of repetitions, since the yardstick signal can only take on a particular value in the first period of any given repetition. Table 6 reports the estimates of a linear probability model (LPM) for the three cases.[19] The main result arising from this table is that there is strong evidence of yardstick competition. In particular, voters pay attention to the difference between the tax rate in another jurisdiction and their own. Furthermore, any deviation from the signal that indicates that they are worse off in comparison to the other jurisdiction is punished by voting out (rejecting) the incumbent. To interpret the yardstick competition coefficient, recall that the yardstick variable is defined as the difference between the yardstick signal and the voter's first-period payoff. So, suppose the signal equals 1.50. Then, the yardstick variable has a nonnegative range. In particular, the variable can take on the values 0.00 (1.50 minus 1.50), 0.75 (1.50 minus 0.75) or 1.34 (1.50 minus 0.16). So, an increase in the yardstick variable from zero to positive automatically tells us that the first-period payoff is below the yardstick signal. Thus a negative sign on the $yardstick_{it}$ coefficient tells us that the voter is more likely to vote out the incumbent when first-period welfare is below that of the other jurisdiction. A similar reasoning holds when the signals are 0.75 and 0.16 respectively.

3.4 Summary and related literature

The previous section discussed treatment and partial effects. This section summarizes these findings and relates them to previous findings by Besley and Case (1995). We start with an overview of the main treatment effects. In general, incumbents in treatment Y behave more corruptly; however, female incumbents in treatment Y behave less corruptly. Thus, yardstick competition acts as a corruption-taming mechanism if the incumbent is female. Voters are less likely to re-elect in the presence of yardstick competition conditional on the draw. In particular, voters in treatment Y fixate on the difference between the tax rate from the other jurisdiction and the tax rate in their own jurisdiction. As this difference increases (i.e. as the voter's own tax rate is deemed more unfair compared to the benchmark), voters are more likely to

[19] These findings are robust to alternative estimation methods (i.e. logit and probit).

Table 6
Voter Regression (Treatment Y)

Dependent Variable: $Accept_{it}$ (1=re-elect, 0=vote out)			
	1.50 ($T = 2$)	0.75 ($T = 4$)	0.16 ($T = 4$)
constant	0.889 (0.000)	0.098 (0.526)	-0.136 (0.399)
gen_i	-0.081 (0.500)	0.067 (0.593)	-0.039 (0.766)
$pool_{it}$	0.087 (0.685)	0.129 (0.384)	0.121 (0.255)
$yardstick_{it}$	-0.549 (0.000)	-0.409 (0.000)	-0.511 (0.000)
history (θ_{t-1})	-0.033 (0.792)	0.253 (0.004)	0.181 (0.018)
$gen_i * pool_{it}$	0.167 (0.537)	0.164 (0.393)	0.078 (0.586)
overall R^2	0.4585	0.3368	0.3205
Notes: p-values in parentheses. $N = 28$.			

vote out the incumbent. This finding is robust to an alternative definition of the yardstick variable. This suggests that voters process yardstick information in a specific manner by focusing on the difference between tax rates. This finding supports Besley and Case (1995). This suggests that the findings of these laboratory experiments do generalize to the external naturally occurring environment. Finally, it should be noted that incumbents in Treatment Y care more about beliefs than incumbents in the baseline. So, the presence of a benchmark makes incumbents more concerned about their choice of tax rates. In particular, the sign of the coefficient (which is positive) suggests that incumbents are more likely to choose a particular tax rate in the presence of yardstick competition if they believe it to be associated with a higher likelihood of re-election. So, even though yardstick competition does not enter on the incumbent's side, it affects incumbent's perspectives for re-election and thus their choices. This gives support to the Besley and Case (1995) claim that *incumbent behavior (i.e. vote-seeking) and tax setting are tied together through the nexus of yardstick competition.* Next, we discuss to the main partial effects. These partial effects hold overall for pooled data from the baseline and the yardstick treatment. On the incumbent's side, three variables affect corruptibility: (1) the unit cost, (2) gender interacted with the unit cost and (3) beliefs. In particular, the unit cost of the public good (θ) pushes incumbents towards equalization when it takes on a high value. This is likely due to the fact that incumbents have the ability to fool the voter by pooling when the unit cost is low. When the unit cost is high, the incumbent must either equalize or separate. If the incumbent chooses the latter by diverting maximum rents and charging maximum taxes, the voter is likely to punish by voting him out. So, an incumbent that behaves strategically and is forward-looking would rather divert zero rents and equalize payoffs when the unit cost is high in an attempt to be re-elected. Gender in combination with the unit cost is

also significant. When the unit cost is high, female incumbents are less likely to equalize. This is likely because female incumbents anticipate being voted out anyway when the unit cost is high. Namely, if the incumbent diverts zero rent when the unit cost is high, there is a still a chance that the voter does not re-elect, since the voter may think that the incumbent is trying to fool him by pooling. So, female incumbents may separate at a higher rate in anticipation of rejection. Finally, beliefs are important. Incumbents choose tax rates that they believe are associated with a higher likelihood of re-election. On the voter's side, mainly the tax rate matters. A lower tax rate (i.e. higher payoff) is associated with a higher re-election rate.

4 Conclusion

This study reports theory-testing laboratory experiments on the effect of yardstick competition on corruption. On the incumbent side, yardstick competition acts as a corruption-taming mechanism if the incumbent politician is female. Also, yardstick competition makes incumbents more aware of their choices, since beliefs matter more in the presence of yardstick competition. On the voter side, voters focus on the difference between the tax rate in their own jurisdiction and the tax rate in another jurisdiction when making their re-election decisions. If the tax rate in their own jurisdiction is deemed unfair compared to the one in another jurisdiction, voters are less likely to re-elect. This suggests that voters use yardstick information in a specific manner; namely, the overall difference between tax rates matters. The experiments also find some partial effects that determine corruptibility. On the incumbent side, the unit cost of the public good matters. If the unit cost is high, incumbents are more likely to charge low taxes. Also, gender is important when interacted with the unit cost. Female incumbents are more likely to divert rents and charge maximum taxes when the unit cost is high. Last, incumbents' beliefs towards re-elections are important. Incumbents weigh the likelihood of re-election heavily when choosing tax rates. On the voter side, beliefs are captured by the tax rate, which is a signal about the incumbent's nature. In particular, voters re-elect more frequently as taxes are lower. Overall, the findings support the seminal claim by Besley and Case (1995) that incumbent behavior (i.e. vote-seeking) and tax setting are tied together through the nexus of yardstick competition. This renders generalizability to the findings of these laboratory experiments and addresses some of the concerns raised by Levitt and List (2007).

5 Appendix: subject instructions

See *http://viceisza.googlepages.com/SubjectInstructions.pdf*. This is an active link.

References

Abbink, K., 2005a. Laboratory experiments on corruption. In: Rose-Ackerman, S. (Ed.), Handbook of Economic Corruption (forthcoming). Edward Elgar Publishers, Northhampton, US, and Cheltenham, UK.

Aidt, T. S., 2003. Economic analysis of corruption: a survey. The Economic Journal 113 (491), F632–F652.

Bardhan, P., 1997. Corruption and development: A review of issues. Journal of Economic Literature 35 (3), 1320.

Bardhan, P., Mookherjee, D., 2005. Decentralization, corruption and government accountability: An overview. In: Rose-Ackerman, S. (Ed.), Handbook of Economic Corruption (forthcoming). Edward Elgar Publishers, Northhampton, US and Cheltenham, UK.

Belleflamme, P., Hindriks, J., 2005. Yardstick competition and political agency problems. Social Choice & Welfare 24 (1), 155–169.

Besley, T., Case, A. C., 1995. Incumbent behavior: Vote-seeking, tax-setting, and yardstick competition. The American Economic Review 85 (1), 25–45.

Besley, T., Smart, M., 2007. Fiscal restraints and voter welfare. Journal of Public Economics 91 (3-4), 755–773.

Brandts, J., Holt, C. A., 2005. Naive bayesian learning and adjustment to equilibrium in signaling games.

Brown, G. W., 1951. Iterative solutions of games by fictitious play. In: Koopmans, T. C. (Ed.), Activity Analysis of Production and Allocation. Wiley, New York.

Browning, E. K., 1976. The marginal cost of public funds. Journal of Political Economy 84 (2), 283.

Dincer, O., Ellis, C., Waddell, G., 2006. Corruption, decentralization and yardstick competition.

Dusek, L., Ortmann, A., Lizal, L., 2005. Understanding corruption and corruptibility through experiments. Prague Economic Papers 14 (2).

Fischbacher, U., 2007. z-tree: Zurich toolbox for ready-made economic experiments. Experimental Economics (forthcoming) 10 (2), 171–178.

Gillette, A. B., Noe, T. H., Rebello, M. J., 2003. Corporate board composition, protocols, and voting behavior: Experimental evidence. Journal of Finance 58 (5), 1997–2032.

Grosskopf, B., Sarin, R., 2006. Is reputation good or bad? an experiment.

Harsanyi, J. C., 1967. Games with incomplete information played by "bayesian" players, i-iii. part i. the basic model. Management Science 14 (3).

Harsanyi, J. C., 1968a. Games with incomplete information played by "bayesian" players, i-iii. part ii. bayesian equilibrium points. Management Science 14 (5).

Harsanyi, J. C., 1968b. Games with incomplete information played by "bayesian" players, i-iii. part iii. the basic probability distribution of the game. Management Science 14 (7).

Holmstrom, B. R., 1982. Moral hazard in teams. Bell Journal of Economics and Management Science 13 (2), 324–340.

Jain, A. K., 2001. Corruption: A review. Journal of Economic Surveys 15 (1), 71.

Jonhson, L. C., 2006. Yardstick competition: An empirical investigation using state taxes and media markets.

Levitt, S. D., List, J. A., 2007. What do laboratory experiments measuring social preferences reveal about the real world? The Journal of Economic Perspectives 21, 153–174.

Lockwood, B., 2005. A note on the hybrid equilibrium in the besley-smart model.

Martinez-Vazquez, J., Arze del Granado, J., Boex, J., 2007. Fighting Corruption in the Public Sector. Contributions to Economic Analysis. Elsevier.

Mauro, P., 1995. Corruption and growth. Quarterly Journal of Economics 110 (3), 681.

Miller, R. M., Plott, C. R., 1985. Product quality signaling in experimental markets. Econometrica 53 (4), 837–872.

Potters, J., Rockenbach, B., Sadrieh, A., van Damme, E., 2004. Collusion under yardstick competition: an experimental study. International Journal of Industrial Organization 22 (7), 1017–1038.

Potters, J., Sefton, M., Vesterlund, L., 2007. Leading-by-example and signaling in voluntary contribution games: an experimental study. Economic Theory 33 (1), 169–182.

Robinson, J., 1951. An iterative method of solving a game. Annals of Mathematics 54, 296–301.

Shleifer, A., 1985. A theory of yardstick competition. RAND Journal of Economics 16 (3), 319–327.

Smith, V. L., 1989. Theory, experiment and economics. The Journal of Economic Perspectives 3 (1), 151–169.

Tanzi, V., 1998. Corruption around the world: Causes, consequences, scope, and cures. International Monetary Fund. Staff Papers - International Monetary Fund 45 (4), 559.

The World Bank, P., 1997. Helping countries combat corruption: The role of the world bank. Tech. rep., The World Bank.

Wooldridge, J. M., 2002. Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge, MA.