Appendix

Contents

A Lab-in-the Field Experiment

A.1 Experimental Design

The experimental framework follows the design in Martin (2016), which analyzed micro-level effects of taxation on citizen behavior. This section first describes Martin's basic game, then how we adjust it to examine the effects of visibility and tax modality. In Martin's basic game there are two players, one Citizen and one Leader. The Citizen receives an endowment, and the Leader receives a group fund of 10 MU. The Leader is asked to allocate the government budget between his own salary and a transfer to the Citizen. At the same time, the Citizen decides, for every decision the Leader could make, whether or not to pay 1 MU in order to fine the Leader 4 MU; any money lost in punishment simply disappears from the game. There are two basic versions of the game: one in which the Citizen gets a 5 MU endowment and is not taxed, and one in which the Citizen gets a 10 MU endowment but must pay a 5 MU "direct tax," which is doubled and given to the Leader as the group fund.

We designed treatments that replicate several desirable features of this original structure. First, in choosing whether to punish rent-seeking, citizens must decide if the returns to doing so exceed the costs of action. The lack of bargaining between the Leader and Citizen reflects a stylized vision of taxation as seen by an individual citizen, namely that they are both exogenous and mandatory. Budgets in the game are constant and observable, allowing us to control for the level of information citizens have.

Following Martin, we refer to the Leader's salary as the group fund to signal that Citizens have some degree of discretion over its disbursement. As in Martin, we use single-shot interactions to avoid the possibility that citizens will use punishment in earlier rounds as a form of signaling about future behavior. All subjects play 6 rounds—1 practice round and 5 data-live rounds. Leaders are paired with three citizens in every round, such that Leaders effectively play three instances of the game in each round—one for each Citizen. Citizen-Leader pairs are built such that, while a Citizen may play with the same Leader more than once, he never plays with the same Leader two rounds in a row. The game is also double-blind in the sense that neither Citizens nor Leaders are aware of whether they have played with their counterparts previously.

A.1.1 VAT Modification

Since VAT is levied on purchases, testing the effect of direct versus indirect taxes requires the addition (compared to Martin (2016)) of a purchasing phase of the game. To avoid conflating the effect of purchasing with the effect of paying VAT, all treatments have a purchasing phase. In each round, respondents choose one of five real items to buy: candles, soap, posho (maize meal), cooking oil, and rice. These items were extensively pre-tested, and are all staple items that most households buy. It is also possible to procure all of them in fine increments. This allowed us to choose the quantities such that the total purchase price of each good was identical. Piloting revealed that respondents have extremely accurate priors on the true market price of these goods, and they were able to correctly the guess their price even when they were sold in less common quantities. Moreover, because of their ubiquity, the market prices varies minimally or not at all across Kampala. Quantities of

all items were chosen such that the value of the goods purchased by respondents would be worth the equivalent of 5 MU (approximately 500 shillings) in actual markets and shops in Kampala.

This yields a modified version of the basic Martin game in which subjects purchase a good regardless of their treatment condition. This modified game is given in the main text of the paper. Note that the Citizen purchases a real item that they take home with them at the end of the game; in the next section we discuss the purchasing phase in more detail. Table A.1 breaks down how much each player has at each stage of the game for the different treatments. If certain treatments delivered a greater economic benefit than others, treatment effects could be due to this difference rather than only to differences in the mode or visibility of taxation. Holding constant the economic value of the game is therefore a key feature of the game.

Direct Tax and Windfall. These treatments are identical to those in the Martin games, except for the initial phase. Now the Citizen's wage is 15 MU (not 10) in the Direct Tax game, and 10 MU (not 5) in the Windfall game. In both conditions, the Citizen gets her wage, then pays 5 MU to purchase a small item.

Visible and Hidden VAT. In both VAT conditions, the Citizen receives a wage of 15 MU. She then purchases a small item for 10 MU. Of that 10 MU, 5 MU represents the base cost of the item, and 5 MU is the VAT. For both VAT conditions, Stages 4 and 5 are identical to the Direct Tax and Windfall conditions. The only difference between the two VAT treatments is the way the tax is handled in Stage 3.

In the Visible VAT condition, the Citizen buys the good, and the 10 MU paid is placed on a "shop" tile on the gameboard. Subjects are told during group training that half of the cost of the good is a VAT and that the group fund over which allocation occurs is built from this tax. In the one-on-one example, subjects (both Citizens and Leaders) are also told that the group fund comes from tax money.

In the Hidden VAT condition, the Citizen is told only the final price of the good and that there is a tax; they are never told the exact amount of the tax. Moreover, the connection between the group fund and the tax is made explicit only in the group training and not in the one-on-one example that all subjects receive before actual play begins. After the group training the tax is levied much as it is outside the lab: Citizens know they pay a tax but are not reminded about the exact amount, and its connection to the group fund (government budget) is obscured. In both Hidden and Visible VAT, Leaders have full information about the source of the government budget and the size of the tax.

The key difference between the Visible and Hidden VAT treatments is transparency: in the Visible VAT game, subjects are told the amount of the tax implicitly because they receive information about both total cost of the good and the relative percentage of the tax (50% of the total cost). In the Hidden VAT treatment they are told only the total cost of the good and that there is a tax. Because the base price of all goods is the market price, they can (but may not necessarily) infer the size of the tax. To control for such inferences, our post-game survey includes measures of respondents' beliefs about the true market price of each good and about the amount of VAT they are paying in the game. The Visible VAT is thus designed to imitate as closely as possible the general equilibrium phase of VAT in which citizens are aware of they pay a tax but are forced to infer a good's "true price"—and thus the effective amount of taxation—based on their priors about what the good cost previously.

Stage	**Direct Tax Game**	**VAT Games**	**Windfall Game**
1	15 MU	15 MU	10 MU
2	10 MU + item	5 MU + item	5 MU + item
3	5 MU + item	5 MU + item	5 MU + item
4	Leader decides how to allocate group fund		
5	Citizen decides whether to punish		

Table A.1: Citizen Endowments. Each cell shows the total endowment possessed by Citizens at each stage of each revenue treatment prior to decision-making. The goal of the design is that Citizens should have identical final endowments in Stage 3 (i.e. before the Leader makes his allocation) despite differences in the initial endowment. Differences between the Direct Tax and VAT games in Stage 2 are driven by the VAT paid on the purchase. This difference is equalized in Stage 3 when subjects in the Direct Tax game are forced to pay a direct tax of 5 MU on their wages.

The second way in which visibility is altered across the games is the degree to which the connection is made between the taxes a citizen pays and the group fund that the Leader receives. In the Direct Tax condition, citizens see real coins taken as a tax on their wage and transferred directly to the group fund, which is then given to the Leader. Thus, the link between taxes and the government's budget is clear. This replicates what actually happens when a citizen pays income or property taxes. In the Visible VAT condition, the price of the good is initially moved to the Shop tile; 500 UGX is then taken, doubled, and passed to the Leader as the group fund. Thus, the link between the budget and the price of the good is made visible. In the Hidden VAT treatment citizens are still told that the group fund comes in part from the taxes they paid on goods (although the citizens do not observe this process), but during gameplay that link is not explicitly modeled, and they are never told the exact amount of the tax.

Note that the Windfall game has variants in which the group fund also comes from aid and oil. This yields three non-tax versions of the game and 6 total treatment conditions: Direct Tax, Visible VAT, Hidden VAT, and our three non-Tax conditions, collectively referred to as the Windfall game. To preserve power for analysis of the VAT and Direct Tax games, the three non-tax conditions are together allocated the same portion of the sample as the remaining 3 treatments. The difference in the starting wages across the conditions are driven by the need to assess a tax either during purchasing (Visible and Hidden VAT) or on Citizen wages (Direct Tax). The higher starting salaries for these conditions ensure that at the allocation stage of the game all subjects have equivalent amounts of remaining resources. Table A.1 gives an overview of the endowments possessed by both Citizens and Leaders at each stage of the game in our three main treatment conditions.

A.2 Implementation

A.2.1 Subjects

Within each game session, we randomly assigned subjects to the role of Citizen or Leader at a ratio of 3 Citizens per Leader. In the first round, each Citizen was randomly assigned to a play with a Leader, resulting in three Citizen-Leader pairs for each round. During the game, each citizen received the transfer decided by the Leader to whom he or

she was assigned. Leaders thus made transfer decisions for each Citizen separately. In each subsequent round, the subjects' roles remained the same, but Citizen-Leader pairs were re-randomized. Citizens never played with the same Leader two rounds in a row. Though Citizens could be paired with the same Leader more than once throughout the activity, neither Citizens nor Leaders knew in any given round whether they had played with their counterpart in earlier rounds. The randomization algorithm used to generate Citizen-Leader pairs also minimized the average number of repeated pairings.

A.2.2 Sampling

Enumeration was conducted at field sites located in high-density, lower-income areas around Kampala. Each day, a local political ("LC1") unit was identified by the enumerators for mobilization. A convenience sample was then recruited from the LC1 unit. Because of natural attrition and no-shows, recruiters were instructed to obtain 20 confirmed attendees for each session of 16. Subjects were recruited one day prior to their participation and told which session to attend. Potential participants were notified that we were conducting a research study about how people make decisions involving money in groups, and we would like to invite them to participate. If they did wish to participate, it would be arranged for them to come to the session offices where they would complete a survey and take part in some activities, and that the activities in the sessions would use real money, and are about public finance, and about what Ugandan citizens think about how governments spend money. Potential participants were informed of the estimated session length as well as the minimum and anticipated average compensation for participation, which was equivalent to about a day's wage for the target population. More specifically, each respondent received a small amount simply for showing up, plus their payout from all 5 rounds of the game. This payout included the goods "bought" in each round. Thus, respondents received an amount of cash that depended on their gameplay (and that of the Citizens or Leaders they were paired with) plus the 5 items they purchased. We only recruited participants who were age 20 and older, as these respondents are more likely to have exposure to taxation and to understand the political nature of the games. Consent was obtained when the participant arrived at the project site prior to the beginning of the games. Consent was obtained by the enumeration team leader. The team leader gave the consent form to the participant to read and, if the participant requested it, provided assistance understanding any confusing portions. Participants then gave their first name and signature on a sign in sheet to confirm the receipt and comprehension of the consent form and their willingness to participate. The use of a central sheet collected and kept by our survey firm was designed to facilitate record-keeping and also to remove the necessity of participants carrying around the consent form, including carrying it off-site. Participant could request a copy of the consent form to carry off-site if they so-chose.

A.3 Outcomes of Interest

Punishment Threshold. Our key outcome of interest is the smallest transfer made by the Leader at which the Citizen will not punish the leader. This quantity represents what the subject considers the acceptable transfer, and thus by extension the acceptable level of

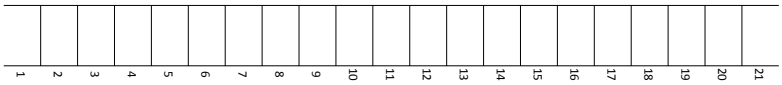

Figure A.1: Utility Ladder

rent-seeking from the Leader. Because Leader transfers below this threshold will result in punishment, which is costly for the Citizen, higher average thresholds represent a greater willingness to punish rent-seeking. Citizens could choose any feasible threshold from 0 — corresponding to a decision to never punish the Leader — to 10, a value that would result in punishment unless the Leader allocated the entire budget to the Citizen.

Loss. To measure loss, we use a "utility ladder" to capture participants' subjective utility at four distinct points in the game. We set the citizen's reference point mid-way up the ladder when the wage for the round is received. We then ask the citizen to update their ladder (1) once the citizen purchases a good, (2) when the leader is given the group fund, (3) after the Citizen observes the Leader's allocation, and (4) at the end of the round (e.g. after any punishment has occurred and final payouts realized). Only the first ladder measurement is directly relevant to discerning whether visibility affects the loss subjects incur from paying the tax at the time of purchase. Ladders 3 and 4 enable us to test whether punishment generates expressive benefits for citizens. The utility ladder is shown in Figure A.1. The ladder is a 21-point scale ranging from 0 to 20. At the beginning of the round, subjects are anchored at a value of ten by the enumerator. The explanation script is as follows:

> Now, this is a picture of a ladder with 21 rungs. Someone at the bottom rung is very unhappy or not well off. Someone at the top rung is very happy or well off. Now, suppose that after you get your wage, you are here at rung 10.

After the Citizen makes a purchase, he or she is asked the following:

> Now, look at the ladder again. Now that you have bought the [ITEM], where are you on the ladder?

The subsequent ladder measurements use similar phrasing, and occur at the relevant part of the round as described above.

This version of the utility ladder was run in Uganda in a prior experiment and piloted extensively. Analysis of both the prior experiment and the one described here demonstrates that subjects' ladder position increases monotonically (in an approximately linear fashion)

with the Leader's transfer. A picture of the ladder as it was shown to respondents is shown in Figure A.1.

Ownership. Citizens also received an additional question to gauge their degree of ownership over the group fund. The answer options were a 10-point Likert scale. The question read:

1. How much do you agree with the following statement: I feel strong ownership over the group fund.

We pre-registered two expectations: first, that, averaging across conditions, ownership will be higher for Citizens in both the direct and indirect tax treatments relative to those in the Windfall condition. Second, we expected that ownership would be highest where the connection between the tax and the group fund is most explicit—in other words, in the Direct Tax condition. This is precisely the pattern we observe in Figure 3.

Manipulation Check. Subjects also received a manipulation check that asked them the source from which the group fund in the game was derived. Enumerators were instructed not to read any of the answer options and to select the source reported by the subject. Answers were then coded as correct or incorrect *ex post*. The purpose of this measure was to demonstrate that subjects failed the manipulation check more often in the conditions in which the tax was the least visible. The results, reported in the maintext, demonstrate that pass rates were nearly perfect in the Direct Tax condition, fell to approximately two-thirds in the Visible VAT condition and fell below a third in the Hidden VAT condition.

A.4 Game Boards

See Figure A.2.

A.5 Additional Results

Table 2 of the maintext shows only the comparisons that are directly relevant to our main hypotheses. We report the results of other comparisons of interest in Table A.2. Note that in the Direct Tax and Windfall conditions, no tax of any kind was paid before the ladder measurement. As a result, all comparisons that include either Direct Tax or Windfall against either the Hidden or Visible VAT conditions will show a substantial difference. In a subsequent ladder measurement taken after subjects paid a tax on their wage (in the Direct Tax condition only), there is a large drop in utility similar to that observed in the Hidden and Visible VAT conditions. Though not directly relevant for the analysis here, the effect of each treatment on the additional ladder measurements described in the implementation section are available on request.

While the specification that includes controls and fixed effects is preferred on precision grounds, we also pre-registered a simpler, treatment-only specification. This specification still utilizes the same clustered and robust standard errors as the main specification, but omits all others. These results are reported in Table A.3. Standard errors increase substantially, reflecting the fact that the enumerator and round fixed-effects soak up a substantial amount of variation and so their absence reduces precision. The main result on subject thresholds remains significant, though at the 5% level instead of the 1% level. The results on the

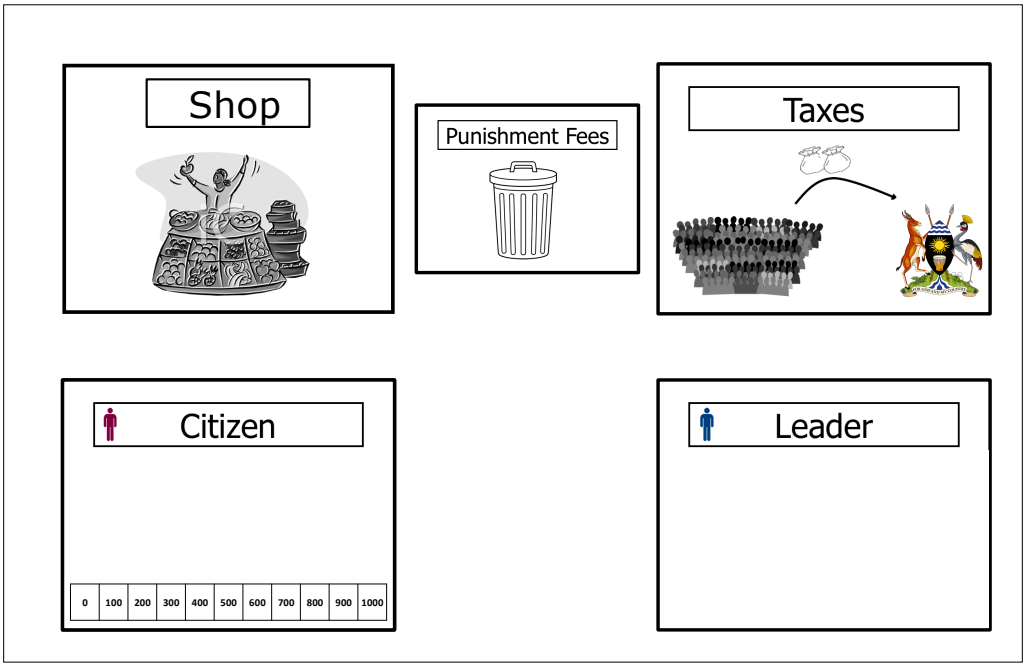

Figure A.2: Example Game Board, Tax as Source. All enumeration was conducted using 100 UGX coins. These were moved around the game board as the Citizen earns her wage, buys an item, pays any taxes, and makes a punishment decision. The “Tax" tile reminds the respondent of the source in the Direct Tax and Visible VAT conditions; it was left blank in the Hidden VAT condition and just labeled as “group fund".

	Dependent Variable		
	Threshold	Ownership	Ladder Position
Direct Tax - Hidden VAT	15.03	0.32	4.68^{***}
	(11.20)	(0.19)	(0.23)
Visible VAT - Hidden VAT	27.46^{**}	0.34^{*}	-0.44^{*}
	(10.70)	(0.19)	(0.23)
Visible VAT - Direct Tax	12.44	0.03	-5.12^{***}
	(11.11)	(0.19)	(0.22)
Direct Tax - Windfall	43.20^{***}	0.50^{**}	0.28
	(11.34)	(0.19)	(0.23)
Hidden VAT - Windfall	28.18^{**}	0.18	-4.40^{***}
	(10.90)	(0.19)	(0.24)
Visible VAT - Windfall	55.64^{***}	0.53^{***}	-4.84^{***}
	(10.75)	(0.19)	(0.24)
Round FE	✓	N/A	✓
Item FE	✓	N/A	✓
Covariates	✓	✓	✓
N	4150	829	4150

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.2: Treatment Effects on Punishment, Ownership, and Ladder Position. Increasing the visibility of an indirect tax increases subjects' punishment thresholds, budget ownership, and feelings of loss when purchasing the good (Row 1, Columns 1, 2 and 3 respectively). Columns 1 and 3 use subject-round data with subject-clustered (CR2) standard errors. Because ownership does vary across rounds, Column 2 uses subject-level data with robust (HC3) standard errors. Note that the total sample size includes all treatment conditions.

ownership and loss mechanisms are similar in magnitude but fall out of significance due to the loss of precision.

Figure A.3 shows the results of the lab experiments. Panel A of Figure A.3 plots the treatment means and 95% confidence intervals for Citizens' punishment thresholds. Panel B of Figure A.3 plots the by-treatment means and 95% confidence intervals for our subject-level ownership measure. To examine how tax visibility affects the subjective loss individuals feel from paying a tax, Panel C of Figure A.3 shows Citizens' average utility ladder values after purchasing. In both VAT treatments, utility decreased after purchasing the taxed good, relative to the pre-purchase anchor of 10.[1]

Table A.4 shows additional results on expressive benefits from punishment. As described above, we measured subjective utility for each citizen immediately after they found out how much of the group fund they had received from the Leader, but before any punishment took place (Ladder 3). We then measured utility again at the end of the round, after any punishment had taken place (Ladder 4). Recall that punishment occurred only if the transfer was lower than the punishment threshold set by the Citizen. If there are no expressive benefits to punishment, we should expect that Ladder 4 is *lower* than Ladder 3 for individuals who punished. If (Ladder 4 - Ladder 3) is on average higher for individuals who punished, this suggests that they must receive expressive benefits from the punishment. The first column of Table A.4 shows exactly this: controlling for the subject's threshold, the leader's transfer, and the transfer the citizen received in the previous round, we find that on average Citizens who punish see their utility increase by 5.24 ladder rungs. This provides strong evidence that punishment generates expressive benefits.

Table A.4 also shows that the size of these expressive benefits – again measured by (Ladder4-Ladder3) are higher when a citizen was more in the realm of losses from paying the tax. To test this, our independent variable, *Degree of Loss* is 10 minus Ladder 2, which measures how far below the initial utility benchmark of 10 a respondent was at the time the Leader was given the group fund. Limiting our analysis to those who punished, we expect that those with high values of Degree of Loss are associated with higher expressive benefits from punishment. We indeed find that this is the case: for each additional point of loss suffered earlier in the game, the utility gain from punishment is 0.39 rungs higher.

Leader Behavior: To ensure that Leader behavior is not an influential factor in our results, we ran several straightforward analyses; the results suggest that Leader behavior is not a confounding factor.

Figure A.4 plots the mean Leader transfer, by round, for each of the treatment conditions. The clear correlation of the paths and limited variance in their trendlines across all treatment conditions suggests that treatment group assignment does not have a notable association with Leader transfer decision.

Further, Table A.5 presents OLS estimator regression results for several derivations of a model reflecting mean Leader transfer, by treatment. Columns Two and Three include Leader fixed effects and Leader & Round fixed effects, respectively. All models cluster standard errors by Leader ID. The coefficients and their relatively large standard errors suggest no statistically significant relationship between treatment group and Leader transfer.

[1] The Direct Tax and Windfall conditions are omitted because respondents bought an untaxed good and thus ladder values are not comparable.

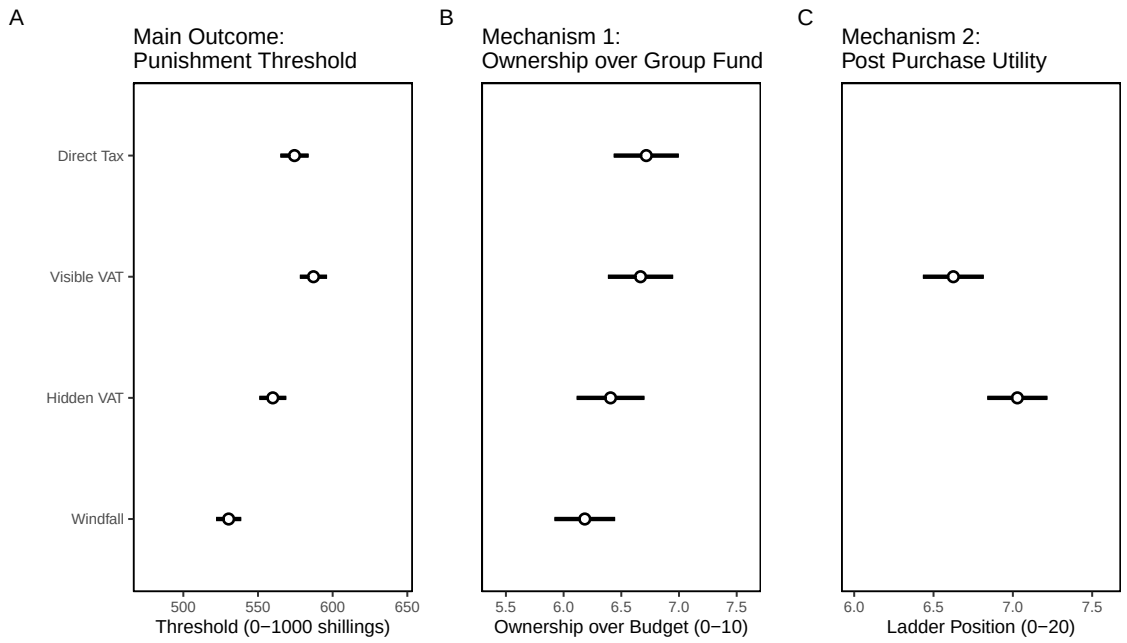

Figure A.3: Mean Punishment Thresholds, Ownership, and Ladder Position by Treatment Condition. Panel A shows average punishment thresholds; panels B and C show by-treatment averages for the ownership and loss mechanisms.

Additionally, to ensure that any relationship isn't obscured through the collapsing of Leader transfer to the mean, we run a similar set of models across all Leader transfer observations, as presented in Table A.6. Again, we included multiple extensions of the main model to ensure that our models are properly established. As with Table 1, there is no statistically significant relationship at any conventional level across all models, including when Leader fixed effects, round fixed effects, and Subject threshold as an additional covariate are introduced.

	Dependent Variable		
	Threshold	Ownership	Ladder Position
Direct Tax - Hidden VAT	14.52	0.31	4.65***
	(12.33)	(0.21)	(0.26)
Visible VAT - Hidden VAT	27.24**	0.26	−0.40
	(12.08)	(0.21)	(0.27)
Visible VAT - Direct Tax	12.72	−0.05	−5.05***
	(12.38)	(0.20)	(0.26)
Direct Tax - Windfall	44.08***	0.53***	0.18
	(11.90)	(0.20)	(0.26)
Hidden VAT - Windfall	29.56**	0.22	−4.47***
	(11.59)	(0.20)	(0.27)
Visible VAT - Windfall	56.80***	0.48**	−4.87***
	(11.64)	(0.20)	(0.26)
Round FE	✗	N/A	✗
Item FE	✗	N/A	✗
Covariates	✗	✗	✗
N	4150	829	4150

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.3: Effects with Treatment-Only Specification. Results are similar to the specification with fixed effects and subject- and round-level covariates, but standard errors are larger. Columns 1 and 3 use subject-round data with subject-clustered (CR2) standard errors. Because ownership does vary across rounds, Column 2 uses subject-level data with robust (HC3) standard errors. Note that the total sample size includes all treatment conditions.

	Dependent Variable:	
	Benefit from Punishment	
Punishment Observed	5.24^{***}	
	(0.22)	
Degree of Loss		0.39^{***}
		(0.04)
Visible VAT	0.26	-1.08^{***}
	(0.20)	(0.36)
Direct Tax	-0.02	-0.76^{**}
	(0.20)	(0.34)
Hidden VAT	0.23	-1.02^{***}
	(0.20)	(0.34)
Subject Threshold	0.00^{***}	0.00
	(0.00)	(0.00)
Previous Transfer	-0.07^{**}	-0.04
	(0.04)	(0.06)
Leader Transfer	0.16^{***}	0.49^{***}
	(0.05)	(0.08)
Round FE	✓	✓
Item FE	✓	✓
Covariates	✓	✓
N	3400	1721

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.4: Evidence of utility gains from punishment. Column 1 reflects the expressive benefits from punishment, where the benefit from punishment is the difference in the ladder position at point 3 (after the Citizen observes the Leader's allocation) and point 4 (at the end of the round after punishment has occurred and final payouts realized). Punishment Observed is an indicator variable denoting whether the citizen chose to punish the Leader. Given that choosing to punish the Leader has a negative economic effect to the Citizen, we would expect that there would be expressive benefits from the decision to punish. The positive and statistically significant coefficient of Punishment Observed affirms this expectation. Column 2 expands upon this, showing that the benefit from inducing punishment upon the leader is associated with the degree to which the Citizen experiences loss. The positive and statistically significant coefficient of Degree of Loss (10 - the ladder position at point 2) implies that larger losses subsequently lead to larger expressive utility gains from punishment.

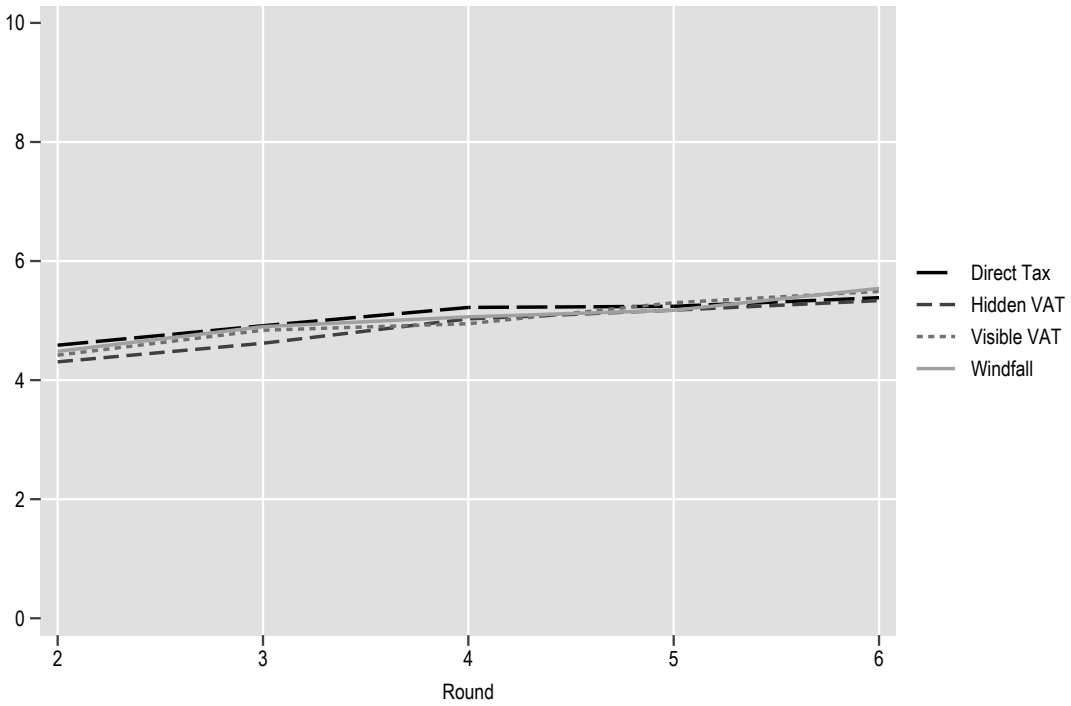

Figure A.4: Average Leader transfer by round for each treatment group

DV: Leader Transfer			
Hidden VAT	-0.224	-0.224	-0.224
	(0.158)	(0.158)	(0.159)
Visible VAT	-0.042	-0.042	-0.042
	(0.207)	(0.207)	(0.208)
Windfall	-0.021	-0.020	-0.023
	(0.221)	(0.222)	(0.223)
Constant	5.064***	5.595***	5.080***
	(0.389)	(0.115)	(0.077)
N	1442	1442	1442
Leader FE		✓	✓
Round FE			✓

$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

Table A.5: Mean Leader transfer and treatment group. Standard errors clustered by Leader. Columns (2) and (3) include Leader fixed effects and Leader & Round fixed effects, respectively.

DV: Leader Transfer				
Hidden VAT	-0.177	-0.177	-0.176	-0.176
	(0.151)	(0.150)	(0.152)	(0.151)
Visible VAT	-0.070	-0.068	-0.068	-0.068
	(0.217)	(0.217)	(0.215)	(0.220)
Windfall	-0.037	-0.036	-0.036	-0.036
	(0.201)	(0.200)	(0.199)	(0.185)
Subject Threshold				0.000
				(0.000)
Constant	5.069***	5.590***	5.046***	5.039***
	(0.389)	(0.113)	(0.069)	(0.226)
N	4148	4148	4148	4148
Leader FE		✓	✓	✓
Round FE			✓	✓

Table A.6: Dependent variable is Leader transfer. Standard errors clustered by Leader. Columns (2) and (3) include Leader fixed effects and Leader & Round fixed effects, respectively. Columnn (4) includes Subject's Punishment Threshold, Leader fixed effects, and Round fixed effects.

B Survey Experiment

B.1 Experimental Design

The VAT experiment has the following structure:

1. Respondents complete an "effort task" — in this case, a conjoint survey experiment — to earn money (2600 UGX) that is used in the VAT experiment. The amount of 2600 UGX is approximately US$0.75. This is significant enough in the local context that the payoffs are meaningful, and that respondents are fairly compensated for taking time to complete the survey, but not so high as to make the survey appear coercive.

2. Respondents are shown a 21-rung ladder that represents wellbeing; higher on the ladder corresponds with being better off. Respondents are initially placed in the middle of the ladder (rung 10).

3. Respondents are presented with a decision in which they must spent a portion of their earned income to buy one of two goods: soap or airtime. The goods cost the same amount; the actual amount can vary and was chosen based on the local costs of those items.[2]

4. Respondents then given one of three treatments:

   - The control group is simply told "this is ITEM that costs AMOUNT."
   - The "Hidden Tax" group is also reminded that the item includes tax, but is not told how much that tax is.

[2]In practice, the variation was extremely small — in all but a few cases the good price was either 500 or 600 UGX.

- The "Visible Tax" group is explicitly told which taxes are levied on that item, how much they are, and how much the item would cost without taxes.

5. The respondent is asked to update where they are on the 21-point ladder.

B.2 Text of Experiment

Earlier, we gave you some money in return for completing a task. We told you that you would be using some of this money to complete a task. We will be using some of that money we gave you for this task. Now, this is a picture of a 0-20 ladder with 21 points. Someone at the bottom rung is very poor or unhappy, and someone at the top rung is very wealthy or happy. To start, we will put you here in the middle, on rung 10.

Now I want to give you the opportunity to buy one of the following items: SOAP, AIRTIME. The SOAP costs BLANK shillings and the AIRTIME costs BLANK shillings.[3]

Q: What will you buy?
1. SOAP
2. AIRTIME

[Control Group]: "This is ITEM that costs AMOUNT Sh."
[Visible Tax]: This is ITEM that costs AMOUNT Sh. If there were no taxes on ITEM, it would cost BLANK– you could buy it and have BLANK Sh left over. But, because there is BLANK the total cost of the ITEM is BLANK Sh.
[Hidden Tax]: "This is ITEM that costs AMOUNT Sh. Remember that this price includes taxes levied by the government."

[All treatment groups]: Now, you must use the money to buy ITEM for AMOUNT Sh – it represents the money you spend in daily life to support yourself or your family.

Now that you have bought ITEM, where are on you on the ladder? [ENUMERATOR: Record response. Someone at the bottom rung (0) is very poor or unhappy, and someone at the top rung (20) is very wealthy or happy.]

B.3 Sampling Strategy

Our sampling strategy was a modified area probability sample in which we intentionally oversampled urban areas and allocated approximately half of the sample to be heads of household. We did so because another experiment fielded simultaneously was targeting its treatments towards this demographic. The characteristics of this population are such that the treatment effects we find in the survey experiment are likely to be lower bounds on the effects we would find in a more rural sample. This is because city-dwellers are more likely to be informed about government behavior, meaning that information on the rate of taxation may be less valuable. They are also relatively wealthier than more rural respondents and

[3]These amounts are filled in by the enumerator based on local prices that they get from their supervisor. This was done due to confusion in piloting when prices differed across locales.

thus may suffer less severe utility losses even when information about the amount of tax paid for their item of choice is not previously known.

We split our sample between municipalities—a special administrative designation reserved for urban areas—and non-municipalities. In each of ten districts, split across Uganda's four regions proportional to their respective shares of total population, we sampled one municipality and one non-municipality, both of which are counties. We thus achieve a more urban sample than would be obtained in a conventional area probability sample, but still retain a significant proportion of peri-urban and rural respondents. In the urban county (e.g. the municipality), our sampling frame was the universe of polling stations present in the 2016 elections. We then binned polling stations into quartiles according to the number of registered voters.[4] In municipalities, we drew 8 polling stations, taking 2 from each quartile to ensure that we cover the polling-station size distribution, which is highly correlated with an area's level of urbanization. In non-municipalities we drew 4 polling stations, one from each quartile. Sampling at the PSU level was random walk starting from the polling station.

Respondents were selected according to their head-of-household status, with the goal of an approximately equal split between head of households and non-head of households. We defined head of households for our purposes to be the male or female that is responsible for making financial decisions and/or handles household expenditures. Actual sampling was done using respondent cards, which are one of these two types and are shuffled at the start of each day. The remaining one-half of the sample is non-household heads, which were selected randomly after a full enumeration of all qualified residents of the household.

B.4 Tax and Certainty Modules

In order to measure tax prevalence and the certainty with which subjects could estimates their income burden for individual tax types, we used two related modules. In the first, we elicited the following information:

1. Their total individual and household income, including from secondary sources such as a small business
2. How much money they estimated having paid in all kinds of taxes as as a raw amount
3. Whether they paid an income tax (also called "pay as you earn" in Uganda), service tax, business licence tax, corporate tax, value-added tax, import tax, excise tax, or property tax

The binary measures in the final item were used to create the left panel of Figure 5. After this first module, a second "certainty" module was completed in which, for each of the taxes the subject indicated paying, the following was asked:

1. How much (in UGX) they paid for each individual tax
2. Their level of certainty that the number they gave was correct (1-10 scale)

[4]The number of registered voters is an excellent proxy for total population, which is not available at the polling-station level.

The full text of both batteries are available on request. The certainty measure was elicited using a ladder scale similar to that used in the lab experiments. Respondents were told that a value of 1 corresponded to being "not sure about the tax amount AT ALL" while a value of 10 meant that s/he was "completely certain about the tax amount." These measures were used to construct the right panel of Figure 5. Subjects that did not report paying a given tax were coded as missing on the corresponding certainty measure. If, for example, a subject said they did not pay an income tax, they were coded as missing on the certainty measure for income taxes. The mean certainty estimates for each tax are thus calculated using only subjects who reported paying that tax.

B.5 Additional Results

In Figure 5 of the main text, we presented results for the proportion of our survey experiment sample paying each of 5 major taxes: VAT, excises, local, business and income. We also showed respondents' confidence in estimating the amount of money paid in the form of each. A concern with this result was that the certainty measure for the direct taxes may be artificially high because only relatively wealthier respondents, who may be more certain about all tax types, pay direct taxes. To examine whether this was the case, we conducted the same analysis as in Figure 5 on the restricted sample of subjects who paid at least one direct tax (n = 552). These results are shown in Figure B.5, and demonstrate that the order of the results found in Figure 5 hold on this restricted sample: the two indirect taxes are those over which subjects are the most uncertain, while the 3 direct taxes are those over which subjects are the most certain.

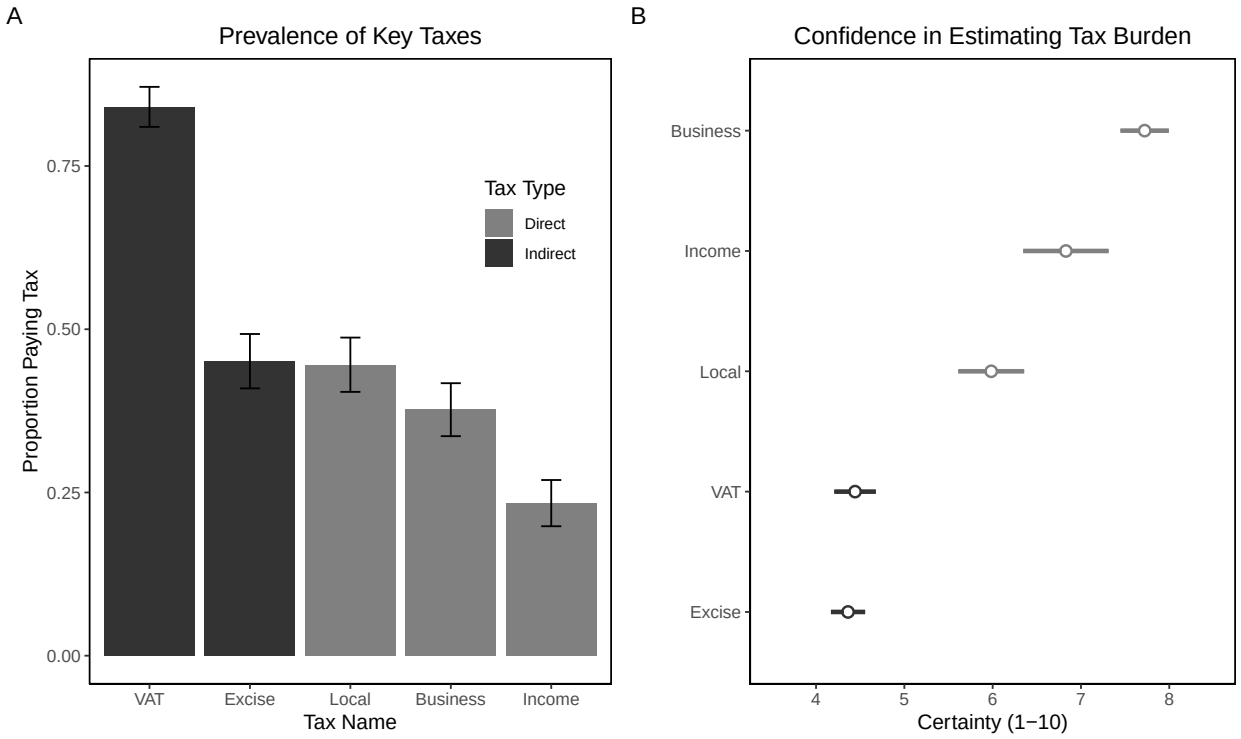

Figure B.5: Tax Prevalence and Certainty for Most Common Taxes for Restricted Sample. Included subjects are those who reported paying at least one direct tax (n = 552). Left pane shows the proportion of the sample that reports paying the most common direct taxes (gray) and indirect taxes (black). The right pane arranges these taxes according to how confident respondents were in reporting the amount of their income they paid in each tax. As expected, the two indirect taxes (VAT and Excise) are by far the least visible.

B.6 Supplementary Afrobarometer Analysis

The main text uses data from Afrobarometer to examine the proportion of Ugandans who believe that they pay income tax and value-added tax, respectively. This appendix describes the data analysis in more detail, and supports our claim that citizens are systematically under-reporting exposure to VAT but not income tax.

Afrobarometer Round 5 was enumerated in Uganda in December 2011-February 2012. The resulting dataset is a "nationally representative, random, stratified probability sample" of 2,400 Ugandans. More information on the survey, including codebooks and data, is available at https://afrobarometer.org/publications/uganda-round-5-summary-results.

For our analysis we draw on responses to three questions. To measure perceived consumption tax payments, we use Q73A, which asks " Regardless of whether you are able to pay them, are you required to pay each of the following, or haven't you been able to find out about this: [Insert either "General sales tax" or "Value added tax"] on the food or goods that you buy from shops or traders?" To measure perceived income tax payments, we use two questions. The first asks "Regardless of whether you are able to pay them, are you required to pay each of the following, or haven't you been able to find out about this: If you have paid employment, are you required to pay an income tax, that is, a tax deducted from your wages by your employer?" The second asks a similar question but for self-employment earnings: " Regardless of whether you are able to pay them, are you required to pay each of the following, or haven't you been able to find out about this: If you are self employed, are you required to pay a tax on the earnings from your business or job?" We count an individual as paying income tax if they respond yes to either question. Summary statistics show that 28.56% of Ugandan respondents report paying income or self-employment tax. In late 2011 / early 2012, when the Round 5 survey took place, the threshold for income tax payments was earning at least 130,000 UGX per month, or approximately 4,300 UGX per day [5]. While official income data are difficult to find, 2016 data from the World Bank reported that only about 30.1% of Ugandans were above the "lower middle income class poverty line" of 4228 UGX per day [6]. Assuming no rapid changes in income distribution between 2012 and 2016, this suggests that the 28.56% number of income tax payers is close to the expected number. Thus, income taxes are not drastically under-reported, suggesting high visibility.

In contrast, only 38.5% of Ugandan respondents report being required to pay value-added taxes on purchases. As almost all purchases are subject to VAT in Uganda, this suggests significant under-reporting of consumption tax liabilities, and is thus consistent with low visibility for indirect taxes.

One concern with our results would be if respondents know they pay consumption taxes, but are not familiar with the term "value-added tax". However, Afrobarometer's protocols include extensive pre-testing in the national and local languages of each country, along with extensive enumerator training. We believe that this should ensure that the question correctly conveyed the desired concept to respondents, and avoided this kind of misunderstanding.

[5] https://www.wider.unu.edu/sites/default/files/Publications/Working-paper/PDF/wp2021-11-effects-personal-income-tax-reform-Uganda.pdf

[6] https://databank.worldbank.org/data/download/poverty/33EF03BB-9722-4AE2-ABC7-AA2972D68AFE/Global_POVEQ_UGA.pdf

C Cross-National Analysis

C.1 Implementation Details

In this section we provide additional implementation details for the modified extreme bounds analysis used to produce our estimates of the effects of direct and indirect taxation on corruption. The full variable list, along with the data source and categories to which each variable was assigned, are included below. The "Revenue" category is reserved for our two independent variables of interest.

Like all core covariates, the additional covariates used to build each estimating equation are also lagged one period. The additional covariates were chosen based on two criteria: their potential to act as confounders or proxy for them, and their levels of non-missingness. The second criteria was necessary to prevent the models from being estimated on samples that were very small and thus substantially different from one another. The model with the most missingness was estimated on 1,810 country-years while the model with the least was estimated on 2,969 country-years. The ICTD revenue measures have by far the largest missingness — all exceed 25% — and are the limiting factor in data completeness. Of the 15 additional covariates, the mean missingness rate was 16.7%. Mean levels of missingness on the 11 core covariates is 13.6%. The data sources for all variables as well as their levels of missingness are included in Table C.1. The ability to manipulate the effective sample through the use of covariates strengthens the case for considering all plausible models in evaluating the effect of indirect and direct taxation on corruption.

Variable	Group	Source	% Missing
Civil War (binary)	key	PRIO	0.0
GDP growth (annual %)	key	WDI, 2019	9.7
Inflation, consumer prices (annual %)	key	WDI, 2019	19.8
Legislative or constituent assembly election	key	VDEM	9.2
Log - GNI per capita, Atlas method (current US dollar)	key	WDI, 2019	14.2
Log - Population Total (WDI)	key	WDI, 2019	3.0
Presidential election (binary)	key	VDEM	9.2
Regime Type (autocracy, anocracy, democracy) based on Polity2 Score	key	Polity	14.2
Rule of Law index	key	VDEM	9.2
Agriculture, value added (% of GDP)	additional	WDI, 2019	17.7
Exports of goods and services (% of GDP)	additional	WDI, 2019	15.4
Foreign direct investment, net inflows (% of GDP)	additional	WDI, 2019	13.2
Foreign direct investment, net outflows (% of GDP)	additional	WDI, 2019	30.7
Fuel exports (% of merchandise imports)	additional	WDI, 2019	33.9
Fuel imports (% of merchandise imports)	additional	WDI, 2019	31.1
General government final consumption expenditure (% of GDP)	additional	WDI, 2019	21.8
Interstate Conflict (binary)	additional	PRIO	0.0
Labor force participation rate, total (% of total population ages 15-64)	additional	WDI, 2019	28.0
Polity Authoritarian Transitions (binary)	additional	Polity	13.1
Polity Dem Transitions (binary)	additional	Polity	13.1
Population ages 0-14 (% of total)	additional	WDI, 2019	6.9
Population ages 65 and above (% of total)	additional	WDI, 2019	6.9
Trade (% of GDP)	additional	WDI, 2019	15.4
Consolidated Non-Tax Revenue (as share of GDP)	revenue	ICTD	32.0
Revenue excluding grants and social contributions	revenue	ICTD	29.0
Direct taxes excluding social contributions and resource revenue	revenue	ICTD	35.9
Non-resource and resource-derived indirect taxes (combined)	revenue	ICTD	35.4
Urban population (% of total)	additional	WDI, 2019	3.1

Table C.1: Variables Used for Extreme Bounds Analysis

C.2 Robustness Tests

We conducted three sets of robustness tests to investigate the internal and external validity of the cross-national analysis. First, we considered six alternative configurations of fixed effects and time trends: country fixed effects with a linear time trend, country fixed effects with a quadratic time trend, country fixed effects with a linear regional time trend, country and region-year fixed effects, country fixed effects with a quadratic regional time trend, and country and year fixed effects with a quadratic regional time trend.[7] For each alternative specification, we conduct the same extreme bounds procedure used to produce Figure 4.

The specifications were chosen to represent a range of theoretically defensible sources of variation off which to identify the model. The third specification, which includes separate year fixed-effects for each region, has the particularly desirable property that it allows for common temporal shocks to have heterogeneous effects by region. This is especially important for our panel because the temporal period includes several shocks — such as the end of the Cold War and multiple global recessions and regional recessions — that could have region-specific effects on levels of corruption. It is therefore especially encouraging that the pattern observed in Figure 4 persists even under this conservative model.

The use of region-specific time trends in Models 3, 5, and 6 are also motivated by the desire to account for regional-level temporal dynamics that could affect both the demand for and supply of accountability. The results for all six models are given in Figure C.1. While the shapes of the distribution change slightly across models, the overall pattern is identical regardless of specification: direct taxation has an unambiguously negative effect on corruption, while indirect taxation has a null or at best weakly negative effect, always with variably signed coefficients.

Our second robustness test was to consider alternative dependent variables. Because our main results use a measure of accountability as well as a measure of corruption, we considered measures that were similar to either. We used a set of criteria to select candidate measures from the Quality of Governance and Varieties of Democracy datasets. First, we selected any measure that was sufficiently close to the corruption and accountability measures reported in the main text. This resulted in 26 candidate measures, approximately half of which are VDEM measures, including two measures of corruption that were included as sub-indices in our broader corruption measure. Our preferred accountability measure is one of three sub-indices used to create a broader accountability index measure. We included in our list of candidate measures the broader index as well as the remaining two sub-indices. These sub-indices measure horizontal accountability, which measures a government's responsiveness to other state institutions, and diagonal accountability, which measures accountability to civil society and the media. While we view both as sufficiently distinct from the concept of accountability studied in this paper, we include them here for purposes of transparency and completeness.

We evaluated the missingness of these measures across the duration of our panel and

[7] The regions are as follows: East Asia and Pacific (n = 1,030 country years); Europe and Central Asia (n = 1,749); Latin America & the Caribbean (n = 1,282); Middle East and North Africa (n = 780); North America (n = 78), South Asia (n = 312), and Sub-Saharan Africa (n = 1,818).

dropped any measure with missingness exceeding 20 percent.[8] Because many of these measures are non-missing only for wealthy countries and for the most recent years in our panel, this criteria eliminated all but 10 measures. The full list of measures is included in Table C.2, along with a brief description of the measure, its variable name in the dataset, and its missingness. For ease of interpretation, we leave corruption measures in their original direction, such that higher values indicates more corruption. Thus, for accountability measures we should observe that the direct taxation coefficient distribution dominates the indirect taxation distribution, while the opposite should be the case for corruption measures. All variables are rescaled using the `scales` package in `R` so that, like our maintext measures, they are also between 0 and 100. Taking these 10 measures as additional dependent variables, we conduct the same analysis as in the maintext, estimating all 4,943 models using country and year fixed effects.

The results are given in Figure C.2. While the magnitude of the effects differs across the different measures, the relationship between direct and indirect taxation is identical to that shown in the maintext, with direct taxation associated with comparatively stronger good governance effects across all ten measures. We therefore view as unlikely the possibility that our results are driven by particular features of our chosen dependent variables.

In our third robustness test, we address concerns regarding the uncertain sequencing of changes in the dependent and independent variables by using their 5-year moving averages. To avoid misspecification stemming from missingness, we calculated each moving average variable as a nonmissing moving average (*i.e.* the sum of observations in years $t-2$ through $t+2$ divided by the number of years in which an observation was present). The results are given in Figure C.3. As with the previous robustness checks, the relationship between direct and indirect taxation mirrors the delineation shown in Figure 4.

Our final robustness test was to explore alternative ways of aggregating the coefficients from each model beyond the simple average. First, we weighted coefficients according to their corresponding models' R^2 statistic, giving greater weight to models with lower residual variance. Second, we weighted according to the proportion of non-missing data, giving greater weight to models whose coefficients were estimated on a larger and thus more representative sample. Third, we weighted according to the product of the R^2 statistic and the proportion of non-missing data. The change in the average coefficient and the proportion of time each were significant were trivial. Results using the simple average and the third weighting scheme are available as part of the model output in the replication archive.

[8] The 20% threshold was chosen because it represented a clear breakpoint: mean missingness was 74.4% among variables with a missingness of 20% or more.

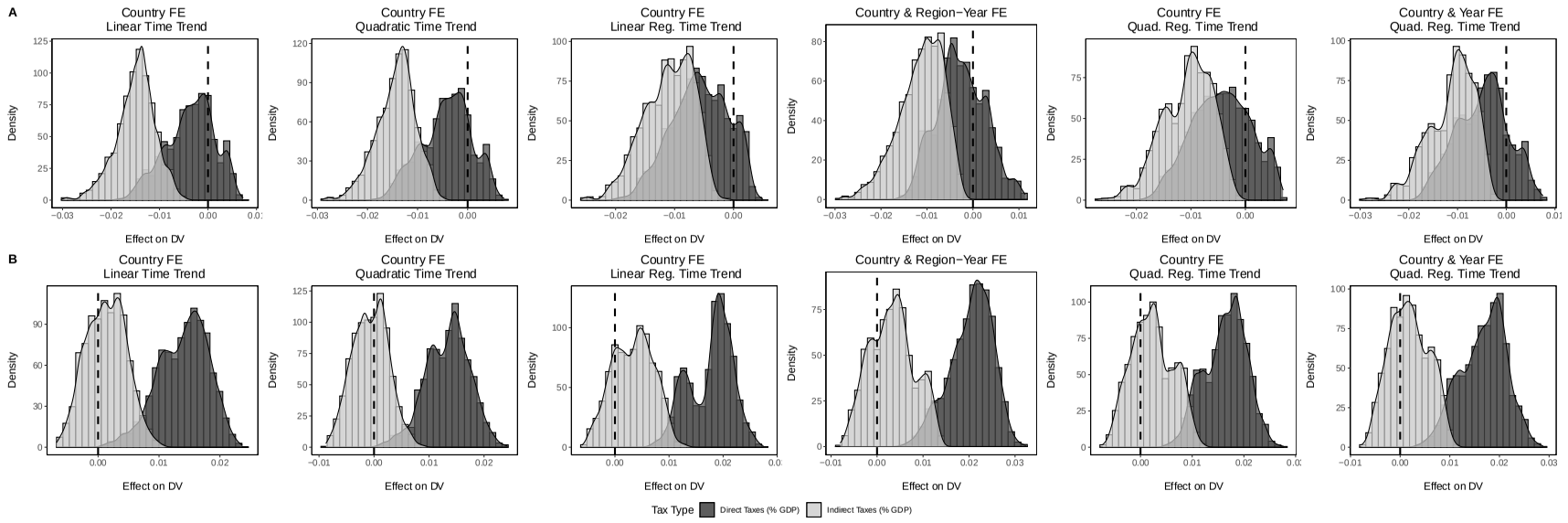

Figure C.1: Robustness to Alternative Modeling Choices. Histograms are produced using the estimates of direct and indirect taxation from each of the 4,943 total models made possible by the included base and additional covariates. Row A and Row B display the results in which the dependent variable is Vertical Accountability and Corruption, respectively. In all four alternative models for each measure, the pattern observed in Figure 4 remains the same. A Kolmogorov-Smirnov test for equality of distributions strongly rejects the null hypothesis of distributional equivalence ($p \approx 0$) for each of the four models.

Description	Variable Name	Source	%Missing
Accountability index	aii_acc	Global Integrity	96.2
Bayesian Corruption Indicator	bci_bci	Sherppa Ghent University	20.6
Governance performance	bti_gp	Bertelsmann - Stiftung	87.6
Public Services	ffp_ps	Fragile States Index	67.9
Indicator of Quality of Government	icrg_qog	International Country Risk Guide	34.5
Accountability index	iiag_acc	Ibrahim Index of African Governance	87.1
Corruption Perceptions Index	ti_cpi	Transparency International	82.9
Executive Corruption Index	vdem_execorr	VDEM	13.8
Public Sector Corruption Index	vdem_pubcorr	VDEM	13.8
Control of Corruption	wbgi_cce	The World Bank Group	49.4
Control of Corruption	wel_coc	Christian Welzel	64.2
Free and Fair Elections	bti_ffe	Bertelsmann - Stiftung	87.6
Contestation (as a dimension of polyarchy)	cam_contest	Coppedge, Alvarez, and Maldonado	49.5
Voter turnout in election	cpds_vt	Comparative Political Dataset	83.0
Trust in parliament	ess_trparl	European Social Survey	97.5
Trust in parties	ess_trpart	European Social Survey	97.9
Trust in politicians	ess_trpolit	European Social Survey	97.5
Participation	iiag_par	Ibrahim Index of African Governance	87.1
Index of Democratization	van_index	Tatu Vanhanen	13.7
Public Sector Corruption Exchanges	vdem_excrptps	VDEM	13.8
Legislature Corrupt Activities	vdem_gcrrpt	VDEM	18.9
Liberal Democracy Index	vdem_libdem	VDEM	14.2
Accountability Index	acc_index	VDEM	16.6
Horizontal Accountability Index	acc_horiz	VDEM	16.6
Diagonal Accountability Index	acc_diag	VDEM	16.6
Regime Corruption	v2xnp_regcorr	VDEM	9.2

Table C.2: Alternative Dependent Variables

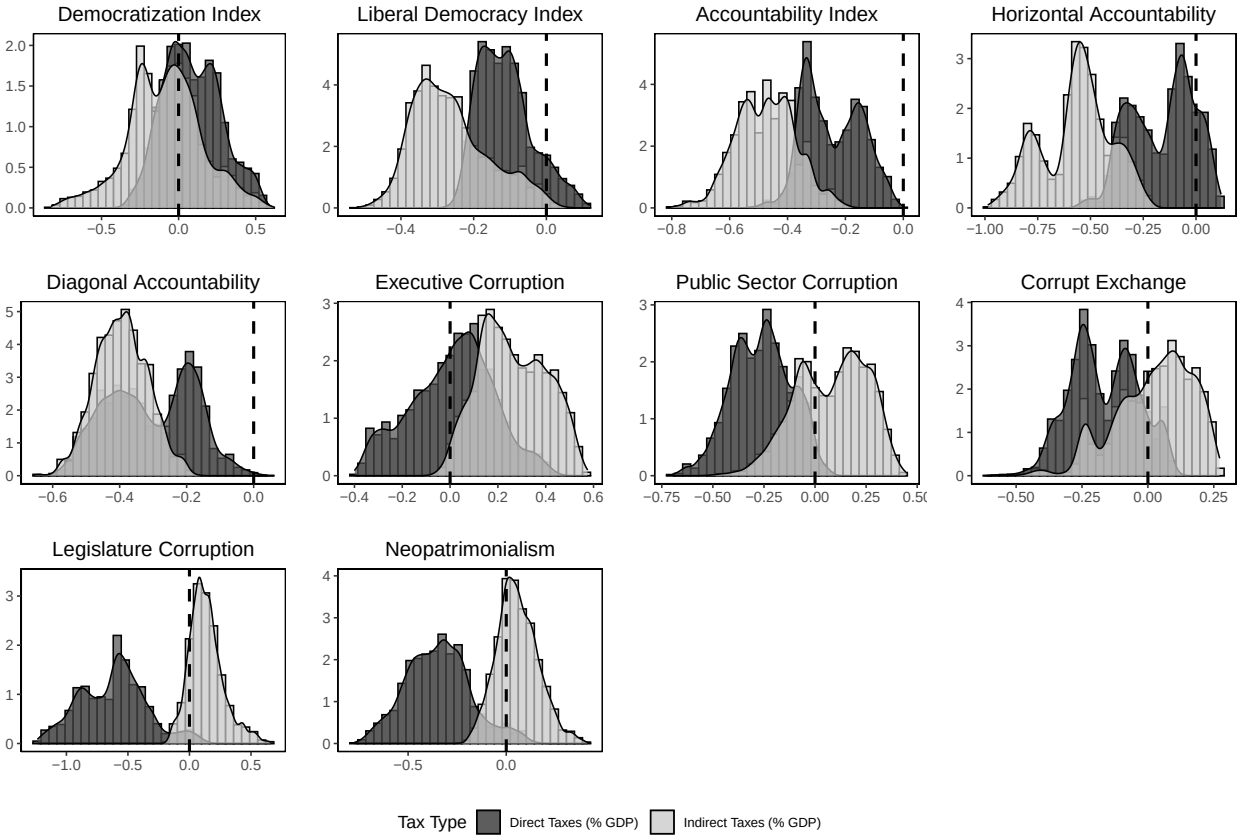

Figure C.2: Robustness to Alternative Dependent Variable. Histograms are produced using the estimates of direct and indirect taxation from each of the 4,943 total models made possible by the included base and additional covariates. While in some cases the direct tax coefficients are more dispersed relative to the measures reported in the maintext, the difference between the two distributions remains.

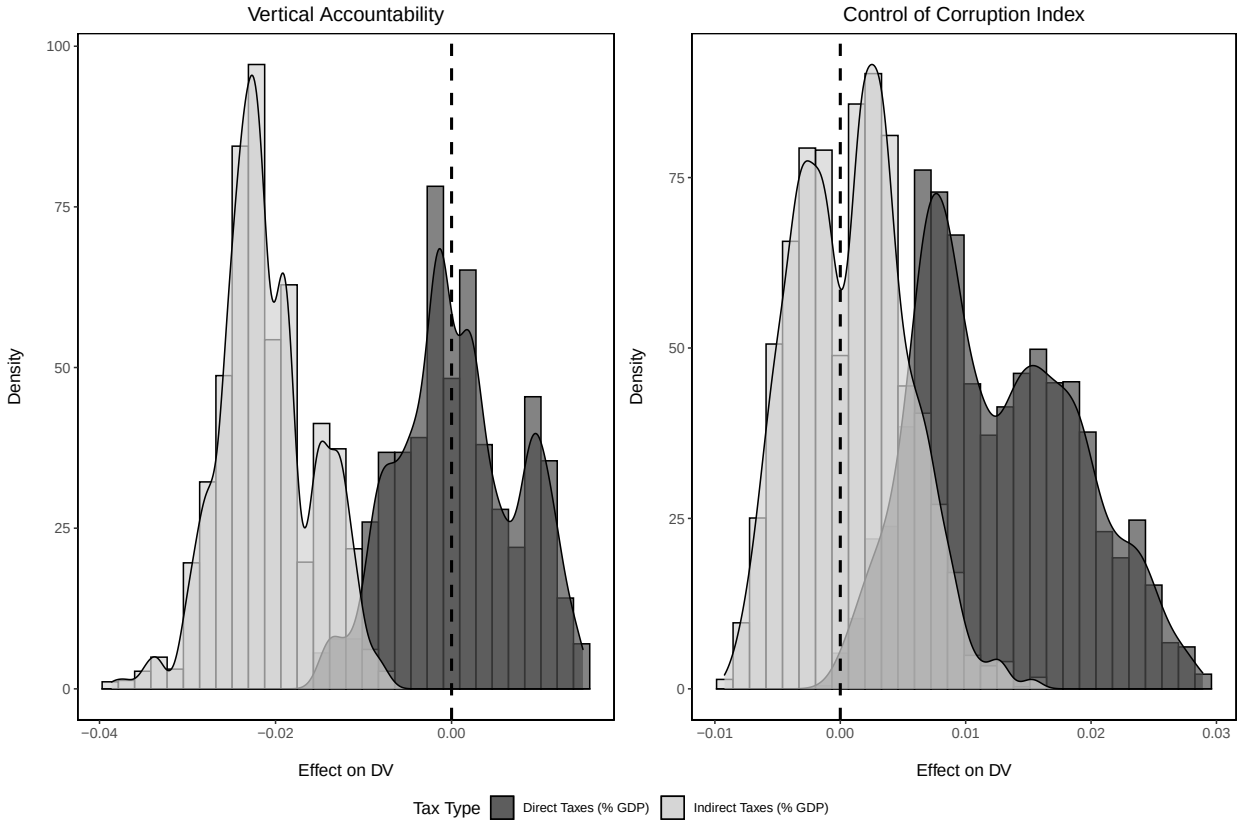

Figure C.3: Robustness to Moving Average of Dependent and Independent Variables. Histograms are produced using the estimates of direct and indirect taxation from each of the 4,943 total models made possible by the included base and additional covariates. In this alternative specification, the dependent and independent variables were replaced with their 5-year moving average. The patterns observed in this alternative specification largely reflects the same delineation between tax types shown in Figure 4.

D Balance Tables and Summary Statistics

D.1 Lab Experiment

Statistic	Mean	St. Dev.	Min	Max	N
Female	0.46	0.50	0.00	1.00	830
Age	29.40	9.22	20.00	69.00	830
Education	9.41	4.12	0.00	18.00	830
Quality of Local Schools	4.53	0.70	2.00	6.00	829
Quality of Local Clinics	4.15	0.78	2.00	6.00	830
Quality of Local Roads	3.68	0.91	2.00	6.00	830
Quality of Local Sewage/Sanitation	2.94	0.90	1.00	5.00	830
Trust in Member of Parliament	2.07	1.00	1.00	5.00	827
Paid Income Tax in Prev 6 months	0.13	0.33	0.00	1.00	821
Voted in Prior Election	0.69	0.46	0.00	1.00	830
Average Punishment Threshold in Game	546.31	117.80	100.00	950.00	830
Average Quality of Local Services	3.83	0.55	2.25	5.00	830

Table D.3: Sample Summary Statistics in Lab Experiment. Relatively higher rate of income tax payment and lower mean age compared to the survey experiment reflects the fact that the lab experiment sample was recruited entirely from urban areas.

	Direct Tax	Hidden VAT	Visible VAT	Windfall	F-test p	FDR q
Female	0.43	0.42	0.46	0.54	0.07	0.49
Age	29.04	29.93	29.40	29.24	0.79	0.79
Education	9.77	9.16	9.61	9.07	0.23	0.54
Trust in Member of Parliament	2.09	2.15	1.95	2.08	0.20	0.54
Paid Income Tax in Prev 6 months	0.14	0.10	0.13	0.12	0.64	0.75
Voted in Prior Election	0.73	0.68	0.66	0.71	0.44	0.62
Average Quality of Local Services	3.85	3.85	3.77	3.83	0.44	0.62

Table D.4: Balance Test for Successful Randomization of Lab Experiment Treatments. First four columns show the means of each covariate for each treatment condition. F-test p value generated by regressing the treatment variable on each covariate. Multiple-testing correction performed via the Benjamini-Hochberg procedure. High p-values are consistent with successful randomization. Covariates chosen to present range of demographic, socioeconomic and political measures that have maximum overlap with both the lab-in-the-field and survey experiments.

D.2 Survey Experiment

Statistic	Mean	St. Dev.	Min	Max	N
Female	0.45	0.50	0	1	1,673
Age	36.24	13.17	20	87	1,673
Education	8.92	4.58	0.00	17.00	1,652
Registered to Vote	0.92	0.28	0	1	1,673
Trust in Member of Parliament	2.19	1.15	1.00	4.00	1,654
Voted in Prior Election	0.76	0.43	0	1	1,673
Paid Income Tax in Prev 6 months	0.07	0.26	0.00	1.00	1,608
Average Quality of Local Services	2.87	0.63	1.00	4.40	1,670

Table D.5: Sample Summary Statistics in Survey Experiment. Relatively lower rate of income tax payment and higher mean age compared to the lab experiment sample reflects the fact that the lab experiment sample was recruited entirely from urban areas while the survey experiment merely oversampled them.

	Control	Hidden Tax	Visible Tax	F-test p	FDR q
Female	0.44	0.45	0.47	0.54	0.61
Age	36.75	35.97	35.99	0.53	0.61
Education	8.52	9.03	9.20	0.04	0.29
Registered to Vote	0.91	0.93	0.91	0.36	0.61
Trust in Member of Parliament	2.14	2.27	2.17	0.14	0.57
Voted in Prior Election	0.77	0.78	0.74	0.30	0.61
Paid Income Tax in Prev 6 months	0.07	0.07	0.07	0.98	0.98
Average Quality of Local Services	2.88	2.88	2.84	0.42	0.61

Table D.6: Balance Test for Successful Randomization of Survey Experiment Treatments. First three columns show the means of each covariate for each treatment condition. F-test p value generated by regressing the treatment variable on each covariate. Multiple-testing correction performed via the Benjamini-Hochberg procedure. High p-values are consistent with successful randomization. Covariates chosen to present range of demographic, socioeconomic and political measures that have maximum overlap with both the lab-in-the-field and survey experiments.

References

Martin, Lucy. 2016. Taxation , Loss Aversion , and Accountability: Theory and Experimental Evidence for Taxation's Effect on Citizen Behavior. Working paper. Available from https://www.poverty-action.org/sites/default/files/publications/Martin_LossAv.pdf.