Transparency & Legislative Behavior

John M. Carey
Department of Government
Dartmouth College
john.carey@dartmouth.edu
http://www.dartmouth.edu/~jcarey/Welcome.html

Abstract
How does transparency affects the behavior of legislators? Specifically, are representatives' proposals and decisions more public-serving when the policy making process is more visible to the public than when it is less so? How does transparency affect constituents' ability and inclination to punish self-serving behavior and reward public-serving behavior by representatives?

Good-government reformers and many (but not all) theorists of representation share an intuition that by facilitating monitoring, transparency mitigates politicians' pursuit of self-interest at the expense of some broader conception of the public good, and so improves the quality of representation. Testing this proposition empirically is difficult, however, because non-transparent legislative procedures are, by definition, unobservable. This project uses experimental methods to compare legislative proposals and decisions, as well as legislators' accountability to the public, in bargaining environments where levels of transparency systematically vary.

The experiment is a repeated game in which Legislators propose and vote on a budget that can be divided among themselves as well as the Public; and the Public, in turn, rewards or punishes the Legislators. Preliminary runs of the experiment in a laboratory setting suggested important effects of varying transparency on how public-serving budgets are, as well as on the mechanics of legislative accountability. However, subsequent runs of a very similar experiment, using a web-based platform, do not replicate many the initial laboratory results, leaving conclusions about the impact of transparency ambiguous. The basic experiment offers a range of opportunities for extensions that would vary the legislator-constituent relationship as well as participants' control over the transparency conditions.

* Thanks to Dartmouth's Neukom Institute for computing support, to Bennet Vance and Yon Soo Park for programming, and to Jeongu Gim, Claire Groden, and Christian Sherrill for research assistance on visible votes.

Prepared for presentation at Rice University, March 12, 2013 and the Brigham Young University, March 14, 2013.

Is transparency a good thing? Empirical, formal, and experimental perspectives

Do legislators behave differently when their actions can and cannot be observed by citizens? Many theorists of representation and good-government reformers share an intuition that transparency facilitates monitoring of politicians by citizens and mitigates the pursuit of representatives' self-interest at the expense of some broader conception of the public good (Bentham 1791; Llanos & Figueroa Schibber 2008; Open.Secrets.org; CongresoVisible.org). The proposition that transparency can affect democratic performance is of practical as well as theoretical interest because although transparency is low in many legislatures, it is relatively easy to do something about it. Machines that record votes can be installed at modest costs (Carey 2009). Attendance and voting records can be disseminated to citizens through simple and cheap media (Humphreys & Weinstein 2008, Hix, Hagemann, and Frantescu 2011). This project aims to test by experiment the impact of transparency on the extent to which legislative decisions serve the public's interest.

High levels of legislative voting transparency have long been taken for granted in the United States, where individual-level voting records on most important motions in both chambers have been made public since shortly after the founding. For just as long, legislators have expected electoral punishment for voting against their constituents' interests (Odegard 1928; Kile 1948; Skeen 1986; Smith 1989; Bianco, Spence and Wilkerson 1996). The prevalent logic in the United States regarding transparency versus anonymity in voting has been that anonymity is necessary for voters, through the secret ballot, in order to free citizens from intimidation in elections, but that transparency in legislative voting enhances democratic accountability. In effect, legislators ought to be subject to pressure on their votes but citizens should not (Lepore 2008; United States Supreme Court 1958).

Not all observers of legislative representation are sanguine about full transparency. Edmund Burke (1774) famously reproached the idea that his constituents' interests, properly conceived, were best served by closely monitoring his behavior in parliament and demanding responsiveness. Schumpeter (1942) advanced a similar point of view nearly two centuries later. Some contemporary empirical accounts make the case that closed-door decision-making produces better policy by freeing legislators from pressures to pursue parochial interests (Birnbaum and Murray 1988). Malesky, Schuler, and Tran (2010) present evidence from an authoritarian setting, the Vietnamese National Assembly, that transparency fosters conformity during deliberations and may therefore discourage the transmission of valuable information among lawmakers. Crisp and Driscoll (2010) present evidence from Argentina and Mexico that legislative decisions on whether or not to make votes transparent are shaped by legislators' desires about which audiences they want to observe their votes.

Formal analyses identify more precisely the conditions under which transparency in legislative deliberations and actions can be either normatively attractive or unattractive. Snyder and Ting (2005) argue that voting transparency ought to be appealing both to citizens and legislators, to the former because transparency exposes potential betrayals of citizens' interests, and to the latter because it makes enforceable commitments to constituent interests possible, and the rewards that might follow from such commitments attainable. Stasavage (2004 and 2007), by contrast, specifies conditions under which transparency can produce worse public policy outcomes than non-transparency, and foster polarization among representatives that prevents citizens from learning through

observation of the policy-making process. Transparency's key liability is that legislators may possess – or be in a position to acquire – a better understanding of policy problems, and of proposed solutions, than do their constituents but fail to deploy that knowledge when doing so in the limelight could damage their reputations as faithful representatives.

The literature on legislative voting transparency, then, is partly historical and partly formal, with relatively little contemporary empirical work. Advocates and skeptics alike concur that transparency should tighten legislator responsiveness to constituent interests, with skeptics warning of the potential for pandering and posturing to offset that advantage. Yet the posited responsiveness bonus has never been directly observed nor have its mechanics been examined. Purely observational studies are constrained partly by the challenge of comparing observable behavior (for example, recorded votes) with unobservable behavior (non-recorded votes), and partly because the votes that are visible are almost certainly not representative of the population of all votes (Carrubba, Gabel, & Hug 2008).[1]

Meanwhile, there is an extensive literature in experimental economics and anthropology on bargaining and cooperation games that bear some resemblance to what goes on in legislatures. This work examines a variety of games in which players are selected either to propose a division of some fixed budget, or to contribute resources to a common pool, and other players must decide whether to accept the proposals. In some variants, responding players may also punish 'selfish' proposers. This scholarship demonstrates that budget proposers are less selfish when they are susceptible to punishment (Fehr and Gachter 2000; Fischbacher et.al. 2001), and also that there is variance across players in different societies about how selfish a proposal must be to warrant punishment (Henrich et.al. 2005, 2006, & 2010; Herrmann et.al. 2008).

So far, so good, but this scholarship has been primarily concerned with identifying norms of cooperation, fairness, and selfishness in interactions among individuals, rather than between citizens and representatives, or within representative institutions, and the staple experiments at the heart of this literature -- ultimatum, dictator, and public goods games -- lack key characteristics that would better approximate legislative environments.

A smaller experimental literature examines the predictions of the Baron and Ferejohn (1989) model of budget division by a legislature with three parties (or participants, in the experimental set-up) that are allocated varying voting weights, although any combination of two of the three is necessary to form a majority. The game is generally played with a finite number of periods, and with the proposal power assigned by some fixed rule (e.g. by probabilities corresponding to voting weights), and reassigned if a proposal is rejected. The focus of investigation in these experiments is the division of spoils among

---

[1] Humphreys & Weinstein (2008) describe research in progress in Uganda in which annual 'parliamentary scorecards' -- audits and reports on the activities of legislators produced by a Kampala-based NGO -- are generated for all legislators during the 2007-2011 term, with active dissemination campaigns to deliver the information in the scorecards are conducted in some legislators' districts, but not in others. This agenda promises a quantum leap in measuring how the transmission of information to citizens affects legislative behavior, but the scorecards themselves focus on indicators of legislators' efforts (attendance, motions introduced, speeches, etc.) rather than specific information about legislative decision-making, precisely because votes are not recorded in Uganda (p.27).

the legislators, and the central result is that proposers exploit their advantage less than predicted by the Baron and Ferejohn model, forming fewer minimim-winning and more universal coalitions, and dividing the budget more equally with coalition partners, than the non-cooperative model would anticipate (Diermeier and Morton 2003; Fréchette, Kagel, and Lehrer 2003; Fréchette, Kagel, and Morelli 2005a & 2005b).

The experiment proposed in this project varies these bargaining games in a couple of simple ways that aim to simulate the monitoring of representatives by citizens. The goal is to shed light on the extent to which transparency affects how public-serving budgets are, and the ability and inclination of the public to reward and punish individual legislators. The experiment also aims to determine whether the public has a preference between minimal versus universalistic coalitions.

How much transparency is there? Counting visible votes

Transparency has been little explored in empirical studies of legislatures outside the United States. One can find general claims that voting against constituent interests risks electoral punishment (Rose-Ackerman 1999). Such an expectation clearly hinges on citizens knowing how their representatives vote. Yet there is huge variation in transparency across legislatures. In most legislatures, the votes of individual representatives are not recorded on most proposals. In many cases, there is no transparency in legislative voting at all.

Almost all national legislatures publish transcripts of their plenary proceedings, including schedules of business, floor speeches, motions and amendments presented, and some information about the results of votes on those motions. Those transcripts are generally published in an official *Record* (or *Hansard, Gazette, Gaceta, Diario Oficial*, etc.). In recent years, these records tend to be published online in the form of PDF documents on legislative websites.

The published records vary substantially in how they present vote results. The two most common formats are aggregate results and individual-level results.[2] For example:

- Aggregate: 75 aye, 10 nay, 5 abstain, 10 not voting/absent.
- Individual-level:
    - 75 aye: Theo Arnold, Linda Bixby, Felix Chalmers, Evelyn Dutton, …
    - 10 nay: Alan Evers, Barney Frank, …
    - 5 abstain: Rafael Garcia, …
    - 10 not voting/absent: Lucia Hernandez, Manuel Irrigoyen, …

Published records in many countries include mixtures of aggregate and individual-level results, although in some countries all votes are published one way or the other.[3]

---

[2] Other formats are rare. El Salvador's record notes only motions that were approved, and only the number of votes in favor. I have not yet encountered instances in which the record provides vote totals disaggregated by party, but not individual legislator.

[3] … or at least they appear to be. The search-ability of the electronic records varies. In most cases, it is feasible to search samples of the records, but not to search every record for every possible vote, so it is not possible to say with 100% certainty that there are no exceptions to the predominant format.

By “visible votes” (VVs), I mean votes on motions taken in plenary sessions of legislatures for which the vote (e.g. aye, nay, abstain, absent) of each representative is recorded and published.

The data on the incidence of VVs presented here come from two distinct waves of data collection. The first extended from about 2000-2006, and involved a combination of:

- field research to eight Latin American countries to visit legislative archives and recover VVs where possible;
- work by research assistants in archives in other countries to recover VVs;
- data exchanges of VVs with scholars in other countries.
- work by research assistants in the United States to recover VVs from legislative websites, where available.

The second is an effort, beginning in 2011, I have undertaken with research assistants in the United States to recover VVs from legislative websites, where available.

In both waves, data from some countries are comprehensive counts of published VVs, whereas in other countries they are estimates. The difference is driven by how hard it is to hunt VVs, and whether comprehensive counts are feasible. In cases where electronic records of votes are available – in particular, where vote records are presented separately from full plenary records– it can be relatively easy to arrive at a comprehensive count. Where it is necessary to download each plenary transcript and search manually (even using word sequence searches) to find votes, comprehensive counts are not feasible, so the approach is to search records from a subset of sessions in a given year and extrapolate an estimate.

Thus, there are gaps in the data collected so far, and there is an asymmetry insofar as the data collected in the recent effort are exclusively from the internet, whereas many of the data from the 1980s and 1990s were originally collected from non-web-based sources, and may not still be available online.

With those caveats on the table, Figure 1 shows the number of VVs recorded and published in the lower or only chamber of 28 national legislatures in recent decades. The general pattern is of low levels of voting transparency a couple of decades ago but increasing variance in recent years as the numbers of VVs in many countries rises.

[Figure 1]

Of course, it would be hard to find any web-based resource that did not increase by an order of magnitude in quantity during the first decade of this century. Yet the increase in web-based VVs is not exclusively a product of blossoming websites. Equally important as a technological factor is the adoption of electronic voting in many chambers, the equipment for which has grown more accessible and affordable. Figure 2 shows the steady increase in VVs for a group of Latin American legislatures that installed electronic voting equipment in the 1990s and adopted it as standard practice. The one exception to this pattern is Peru where, shortly after installation, pro-transparency reformers won a battle, over the objections of the majority party’s leaders, to have votes recorded and full records published online. Fro 1998 until at least 2000, all votes were recorded and published immediately on the Peruvian Congress’s website (Carey 2003). The

electronic equipment is still in place, but at some subsequent point, the practice was terminated, and in recent years Peruvian votes are non-transparent again.

[Figure 2]

Figure 3 shows the even more dramatic increases in voting transparency in a couple of Latin American chambers – Colombia and Honduras – both of which went from nearly no VVs to producing several hundred VVs or more in recent years immediately after adopting the use of electronic voting. Figure 4, by contrast, shows the minimal levels of voting transparency among a number of Latin American chambers that do not employ electronic voting. Recording individual legislators' votes without the use of automated technology is time consuming, procedurally costly, and rare. In short, technology clearly can reduce obstacles to legislative voting transparency, but does not necessarily eliminate political obstacles.

[Figures 3 and 4]

Figure 5 shows VV levels for a series of legislatures outside Latin America.

[Figure 5]

Figure 6 summarizes what is already evident from the previous graphs and narrative – that the production of VVS is strongly correlated with the use of electronic voting technology. But note that there is much higher variance among the electronic than among the manual voters. Some legislatures that installed electronic voting technology either fail to use it, relying instead on traditional methods of voting by show of hands, or use it but fail to publish the individual-level records the machines automatically produce. Venezuela is an example of the former practice, and Nicaragua of the latter (Carey 2009).

[Figure 6]

Figures 7, 8, and 9 show that, among the countries from which data have been collected, lower and upper chambers do not differ markedly in the number of VVs produced, nor to presidential and parliamentary regimes, nor do the Latin American legislatures from those outside the region.

[Figures 7, 8, and 9]

On the whole, transparency has been rising in recent years, but enormous variance remains across countries. For example, the Comparative Political Agendas Project is a vast effort to collect, organize, and disseminate data on policymaking processes in an array of European countries and the United States. Yet although data on the content of legislation, executive orders, and judicial decisions, the substance of party manifestos, parliamentary questions, and government statements, and more are available on the CPAP website, individual-level legislative voting data are available only from the United States and Italy (Baumgartner, Jones, and Wilkerson 2011). Research from other scholars confirms that the widely variant levels of transparency in legislative voting are the norm (Saalfeld 1995; Hug 2010).

The key point, even from this admittedly incomplete exercise in counting VVs, is that individual-level transparency varies enormously across legislatures. Increasing voting transparency dramatically is possible using technologies that are, by now, widely available and not prohibitively expensive. Yet many countries nevertheless fall well short of making votes visible as a matter of course. Questions that follow from this are whether there is reason to think variations in transparency matters and, if so, how? More pointedly, should we think legislators behave differently when a record of their votes will be visible to the public from when it will not? The rest of this paper is an attempt to answer that question, not by comparing transparent and non-transparent behavior from real legislatures – non-transparent behavior being, by definition, unobservable – but through an experiment intended to approximate varying levels of legislative transparency.

The experiment

BASIC structure

The experiment is a game played among participants who are divided between Legislators and member(s) of the Public. It involves a proposal for division of a budget by one legislator, then a vote on whether to approve the proposal by all Legislators (including the Proposer), then budget payouts (if the budget passes), and finally an opportunity for the public to reward or punish each legislator.

The treatments manipulate what information regarding the identity of the proposer, the nature of the proposal, and the legislators' votes are observable by the Public. The 3 different transparency conditions describe what the Public observes:

NT: Non-Transparency

- only its own payout.

ST: Semi-Transparency

- own payout; and
- the identity of the Proposer

FT: Full Transparency

- own payout;
- the identity of the Proposer;
- how much the proposed budget offered to each Legislator; and
- how each Legislator voted (Approve/Reject) on the budget proposal.

The transparency conditions

It is worth saying a few words about what the experimental manipulation of transparency seeks to approximate. Full transparency mirrors the availability of information on most consequential votes in the U.S. Congress, where bill sponsors, party leaders, and floor managers, in amalgamation, are analogous to the Proposer, and where roll call voting records expose every legislator to demands from actors outside the legislative chamber to justify his or her vote.

Semi-transparency is analogous to legislative decision-making without the comprehensive transmission of information that characterizes the U.S Congress, but in the presence of effective legislative parties. That is, even where votes are not recorded and published at the individual level, party leaders generally make their parties' positions known on important proposals before legislatures. Where parties are the main vehicles of policy initiatives, and legislators from the same party vote in concert, then knowing what initiatives party leaders advance or oppose provides citizens with reliable information about how their representatives behave. In the context of this experiment, for the Public to know what s/he got and who the proposer was, as under semi-transparency, is akin to knowing which party championed a policy in political system with strong parties.

Non-transparency is a closer approximation of the legislative process where the full transmission of information is absent and parties are ineffective, either because they are not the main source policy proposals, or because legislative copartisans do not reliably vote in unison, or both. For example, in most Latin American systems, the most important legislative proposals issue directly from the executive branch rather than being formally introduced by specific legislators or parties (Crisp & Driscoll 2010; Morgenstern 2003; Siavelis 2000). Where the president has clear ties to a legislative party, executive initiatives might reasonably be attributed to that party, but in many presidential systems these ties are loose or even non-existent. Presidents' parties are often factionalized precisely by the different demands of competition in executive versus legislative elections (Samuels 2002). Presidents often rely on non-partisan or coalition cabinets, or reject traditional party labels altogether (Linz 1994; Cox & Morgenstern 2001). Under these circumstances, connections between policy proposals and any proposer inside the legislature itself can be obscure. Moreover, when party unity in legislative voting is low, as is often the case in presidential systems, failure to provide a recorded vote can render responsibility for legislative decisions thoroughly opaque (Carey 2009). In short, in many Latin American polities (and I suspect others as well), citizens often find themselves effectively in the experiment's non-transparency mode; they know what they got, but not much else about where it came from.

Versions of the Game

Various versions of a legislative budgeting game, all sharing the basic structure described above, are possible. I discuss a variety of versions briefly in the concluding section, but this paper reports results from two related variants of the game:

- BASIC (web-based)
- BETA (lab-based)

<u>BASIC Version</u>

Players and Preparation

1. Participants are recruited and informed under what transparency conditions the game will be played – NT, ST, or FT, as well as the value of each budget unit (Unit Rate) will be. All periods of each experiment are played under the same transparency condition and Unit Rate.
2. 1 of the 4 participants is drawn at random to act as the Public for all periods of the game. 3 are Legislators (L1, L2, L3).

3. Before the first period, the computer selects 1 of the 3 Legislators at random to be the Proposer for Period 1.

Sequence of Play in Each Period

1. The Proposer is prompted to divide a budget of 24 units among any combination of the 4 players (L1, L2, L3, and Public).
2. The Legislators observe the proposal and vote to approve or reject it.
   - If a majority approves, the budget passes, so all players are awarded their budget shares for that period. For the next period, the Proposer will continue in that role.
   - If a majority rejects, the budget fails, so all players receive zero for that period. For the next period,1 of the non-Proposer Legislators will be selected at random (coin toss) as Proposer.
3. The Public is informed of the outcome, according to transparency condition, and votes "Thumbs Up" or "Thumbs Down" on each Legislator.
4. Each Legislator is informed of the Public's vote (Thumbs Up/Down) for that period, and of her/his running tally of Thumbs Up/Down votes.

Completing the Experiment

- After all 10 periods are complete, the computer determines, on the basis of each Legislator's "Thumbs Up" and "Thumbs Down" tally, whether her/his total budget payoff is doubled.
- Each "Thumbs Up" vote increases by 10% his chance of being doubled. For example:
  - 10 Thumbs Downs means zero chance of doubling;
  - 10 Thumbs Up means 100% chance of doubling;
  - 5 Thumbs Downs and 5 Thumbs Ups mean 50% chance of doubling;
  - ... and so on.

<u>BETA Version</u>

A variant of the experiment – what I refer to here as the BETA Version – was run in a lab at Dartmouth College in August 2010, using networked computers running z-Tree experimental software (Fischbacher 2007), that had been adapted to run the game.

The BETA Version differed from BASIC in the following ways:

- BETA involved 10 participants: the Public, plus 9 potential Legislators, 3 of whom were active for any given period.
- The Proposer was selected at random from among the 3 active Legislators each period, rather than the Proposer continuing in that role if the budget had been approved in the prior period.
- The Public's means of rewarding or punishing Legislators was to reelect or not for the very next period, rather than to cast a vote that affected a Legislator's probability of a subsequent reward (doubling payoffs).
- There were 20 periods of play in each BETA experiment, as opposed to 10 in BASIC.

- After each period except the last, the Public could reelect (or not) each active Legislator. A Legislator not reelected was replaced by one randomly selected from the non-active pool.
- The Unit Rate was $.50 per budget unit, rather than $.05 and $.10 as in the MTurk web-based experiments.

Expectations

The experiment aims to shed light on whether transparency affects two general types of budget distribution outcomes:

- the extent to which budgets serve the Public relative to the Legislators; and
- budget divisions among the Legislators themselves.

The fundamental expectation is that transparency should make the threat of sanction by the Public more effective, so should generate more Public-serving budgets. Specifically:

> ***H1: The higher transparency, the greater the Public's share of budgets.***

There are two ways this might come about, which I refer to as first-order and second-order accountability. The former operates through Proposers' budget offers to the Public, as a result of Proposers' fear of a Thumbs Down vote, and should manifest itself under both ST and FT -- that is, when the Proposer is individually identifiable by the Public -- but less so under NT, when the Proposer is indistinguishable from other Legislators:

> ***H2: Public Offers should be higher under ST and FT than under NT.***

What I call second-order accountability operates through non-Proposer Legislators' desire to be seen as voting for Public-serving budgets and against budgets that ill-serve the Public when votes are visible, so should manifest itself only under FT:

> ***H3: Budget votes by non-Proposer Legislators should be positively correlated with Public Offers (other things equal) under FT, but not under NT or ST.***

Prior expectations regarding whether and how transparency should affect budget divisions among Legislators are more ambiguous. Previous research on budget division games focuses on the magnitude of the Proposer's advantage and, more generally, whether budget divisions are mimimal (paying off the smallest number of legislators necessary to approve a budget) or universalistic. Incorporating a Public and sanctioning by Thumbs Up/Down voting in this experiment, coupled with the transparency manipulations, means that if the Public has preferences about the inclusiveness of legislative coalitions, then the Public's ability to observe the details of budget proposals could affect their inclusiveness.

There are two ways to think about Public preferences over inclusiveness. In principle, minimal coalitions should be less expensive in terms of payments to Legislators, leaving

more resources for the Public, so we might expect the Public to reward minimal coalitions -- perhaps inferring that her own budget share is larger than it might have been had the legislative coalition been universalistic. On the other hand, results from experimental budget division games indicate that players subscribe to norms of universalism to a greater degree than non-cooperative game theoretical models would suggest, in which case the Public may favor universalistic coalitions over minimal ones (Diermeier and Morton 2003; although see also Niou and Ordeshook 1985). In short, the logic of non-cooperative game theory, and previous experimental research, point in different directions regarding what the Public might prefer when it can observe how budget resources are distributed among Legislators. I therefore do not have a strong prior expectation about how transparency ought to affect the incidence of minimal coalitions in the versions of the experiment tested here.[4]

Logistics and considerations: Lab-based versus web-based versions

An advantage of BETA was that the mechanism for the Public to sanction Legislators was directly analogous to that provided by competitive elections. A key disadvantage, however, was that, for the threat of non-reelection to carry weight required a substantial pool of non-active potential replacements. This is both expensive and logistically unwieldy, even in the context of a computer lab in which all participants can be monitored throughout the experiment. Beyond the confines of the lab, a set-up that involves non-active Legislators who must remain engaged and ready to step in – although they might well never be called – is not feasible. Taking the experiment outside the lab with a web-based platform, however, is essential to broadening the participant pool. Moving to "Thumbs Up/Down" as a mechanism for the Public to sanction Legislators allows for this. For a Legislator, the prospect of doubling total payoff is analogous to reelection which, in the context of this game, is an opportunity to amass more budget points, and so a larger payout.

Mechanical Turk

For the web-based experiments, subjects were recruited via Amazon.com's Mechanical Turk (MTurk). MTurk is an online labor market open to anyone with internet access. MTurk employers post discrete jobs – or Human Intelligence Tasks (HITs), in MTurk parlance – and MTurk workers choose HITs according to the job description, the time required, and the payment offered. Amazon takes a 3% commission for operating the market. MTurk is increasingly used by social scientists as a cost-effective way of recruiting subjects for surveys and survey-based experiments. The MTurk worker population, of course, is not a random draw from any population. Yet Berinsky, Huber, and Lenz (2012) replicated a number of seminal experimental studies using samples of US-based MTurk workers and found that it provided pretty good correspondence with in-

[4] Future SMD-versions of the experiment, discussed at the end of the paper, would push the Public away from universalism by dividing the Public and tying each P to a separate Legislator. First, I expect universalism (or minimal-ness) among Legislators' payoffs to correlate with universalism (or minimal-ness) among Public payoffs. That is, I do not expect a budget that distributes payoffs among Legislators evenly to distribute payoffs among P1, P2, and P3 in minimal fashion, or vice-versa. Second, with each P's payoffs directly at stake, notions of universalism that are attributed *Legislators'* payoffs should be weakened, and Public's interest in the value of minimum coalitions should rise. Thus, I expect legislative coalitions to be less universalistic in the SMD versions than in the unitary-Public versions, and for greater transparency to strengthen that pattern

person convenience samples. MTurk workers were slightly younger, more secular, more educated, and with lower incomes (thus, their participation in an on-line labor market) than face-to-face samples. Moreover, the authors replicated results from published survey experiments on public policy preferences and risk acceptance using the MTurk sample (Berinsky, Huber, and Lenz 2012).

Bots

The web-based version of the game cannot guarantee the participation of a full complement of four live players at all times, for two reasons. First, at the outset, fewer than a full complement may "arrive" at the game within a reasonable period, requiring those who are prepared to begin to wait to round out their foursome (potentially producing attrition among waiting players). Second, players who begin the game may abandon it before completing it. The basic problem is that live players have limited attention spans, and will drop out of games if forced to wait too long for other participants to act.

To address the problem of non-response, the web-based version can be populated by robot (or Bot) players who substitute in for absentees. At the start of any game, once the first player completes the instructions and quiz and is prepared to begin, a 60-second timer begins. Any other players, up to four, who arrive at that game within the 60 seconds, are channeled into that game. Once the time elapses, the game begins with those who arrived within the 60-second window. Subsequently, players are required to respond to each prompt in the game (e.g. for the Public to vote Thumbs Up/Down on Legislators, to absorb the information in a screen presenting results from a budget, etc.) within specified periods of time. The specified periods are calibrated to the complexity of the prompt (e.g. Proposers are allowed a bit longer to formulate a budget proposal among four players than the non-Proposer Legislators are allowed to vote on it), and players are allowed a bit more time in earlier periods, when they are learning the game, than in later ones. In all cases, players are made aware of how much time they have to respond to a prompt by a countdown timer on their screen. Players who fail to respond to a prompt within the allotted time are removed from the game, and are seamlessly replaced by a Bot player so the game may continue without disrupting the other live players. Live players are not informed when other participant(s) in the experiment are Bots.

The Bots solve two major potential problems confronting the experiment. First, because the experiment depends on participation in real time among full complements of players, chronic non-response would be devastating. The Bots guarantee responses, and so maintain forward momentum in each game, allowing live players to complete the experiment. The Bots may also serve a more subtle purpose, however, of shedding light on out-of-equilibrium behaviors among live players. That is, in experiments involving all live players, some actions may be rarely observed. The Bots can be scripted to deliver such actions on occasion, allowing us to observe other players' responses to them. Appendix B provides the script, or rules, that governed Bot behavior.

Participants: Lab (BETA) vs. Mechanical Turk (BASIC)

Table 1 summarizes differences between the BETA and BASIC versions of the experiment. I ran BETA three times under each transparency mode, so nine experiments in all, at 20 periods each. Of BETA's 90 participants, over 90% were Dartmouth students, with a couple of faculty and staff who substituted in for no-shows. Participants were almost evenly divided (53%-47%) between men and women. They

were paid a $10 participation fee, then earned $.50 for each budget unit won in the game. Recall that BETA required a pool of non-active Legislators, so many participants in BETA earned no budget units, but those who played – and particularly, those who survived for many rounds – could earn substantial amounts. Many left with $30, $40, or more.

[Table 1]

Under BASIC, the population of participants is more diverse, but it also warrants a bit more attention because it draws from the MTurk online labor market. The gender breakdown among MTurk participants was 62% male, 38% female. The age among MTurk participants runs from 20 to over 60, with most in the 20-40 range, as illustrated in Figure 10.

[Figure 10]

95% of the MTurk participants, reported United States (67%) and India (28%) as their nationalities, with just a handful of participants from other countries. The three panels of Figure 11 compare the US and Indian sub-groups on levels of education and political engagement, as reported in the survey at the end of the experiment. Education levels are fairly similar, although there is a higher concentration of US participants concentrated at 16 years, or equivalent to having completed a bachelor's degree. The Indian participants, by contrast, report slightly higher levels of general attention to politics, and awareness of representatives' votes on important legislation.

[Figure 11]

Of those started the experiment, 80% completed it. Among the 20% who did not finish, 8% did not complete a single period of play, and 12% fell off at some point further along. In both the BETA and BASIC versions of the experiment, players' response times decreased as periods progressed and they learned the game. Figure 12 shows the pattern for budgets proposed by live Proposers in the web-based BASIC experiment. The patterns for voting and rendering Thumbs Up/Down judgments demonstrate the same pattern.

[Figure 12]

Density of live players

Given the limitations of the Bots in their current scripts, "live player density" – the proportion of the 4 players in any given game who are live as opposed to Bots – could be a salient factor in web-based experiments. The challenge is in recruiting via MTurk, because MTurk "workers" select tasks (HITs) independently. The strategy adopted for the February 2013 experiments was to post HITs with short time windows (generally 3 minutes) and with sufficiently attractive base rates of compensation ($1.01 or $2.01) that MTurk workers flock into them during brief windows of opportunity.[5] This approach was

[5] Note that pay rates, especially for HITs that do not require specific skills (e.g. fluency in multiple languages, for translations or transcriptions) tend to be quite low. The cost effectiveness of recruiting experiment participants via MTurk is one of its main attractions for academics (Berinsky, Huber, and Lenz 2012). A HIT that offers a

partially successful, as illustrated by Figure 13, which shows the distribution of live player density across all periods of play by MTurk workers

[Figure 13]

Note that Density is 0 for periods played with no live players (e.g. if a game were triggered but no players survived the instructions and quiz, or if a lone live player dropped out), .25 for a period with one live player and three Bots, and so forth. The histogram shows that the modal period had low Density, but note also that the preponderance of action by live players is in the higher Density periods. The average Density for a period played by any given live player is .53, so most live players were facing at least some other live players. Nevertheless, Density inserts a logistical question into the web-based experiment – Does Live player behavior depend on the proportion of other Live players? To address this, I conducted separate analyses both for the full set of Live player actions and for the subset of Live actions when the Density is either .75 or 1.0 (that is, when either three of four, or all four, players are Live). I found virtually no differences between those conditions.

Payment levels and sufficient incentives

In the lab-based BETA version, players were highly motivated by the payment of US$.50 for each budget unit they earned in the experiment. By contrast, the MTurk players recruited for the web-based BASIC version conducted in February 2013 faced weaker incentives to win each marginal budget unit, with about equal numbers facing an "exchange rate" of $.05 and $.10 per budget unit.[6] The lower Unit Rates were based on the expectation that MTurk workers are motivated at compensation rates lower than those generally used in lab-based experiments (Berinsky, Huber, and Lenz 2012). In light of the overall non-responsiveness to the treatments exhibited in the February 2013 experiments, however, it will be worth considering whether the marginal gain, even to MTurk workers, for winning additional budget units at $.05 or $.10 was insufficient to motivate them to "play for keeps." (I have not detected a measurable difference between the Unit Rates of $.05 and $.10, so I pool the data in reporting results here.)

Summing up

Each of the versions of the game, BETA and BASIC, from which data have been collected to date, was imperfect in important ways. The BETA version had a small number of runs – 3 groups of 10 participants playing a 20-period game under each of the 3 transparency conditions. The BASIC version delivers more iterations of the game, but those are populated by a mixture of live players and Bots. The live players performed well in lots of ways (to be discussed below), but their financial motivations (given the budget unit exchange rate) may have been too weak. The results to date should be viewed with these limitations in mind. I intend to re-run the BASIC experiments with higher Unit Rates of $.25 and $.50 to test whether non-responsiveness to the experimental treatment is due to insufficient financial motivation.

---

guarantee of $2, plus the opportunity to earn more in bonus, for 15-20 minutes participating in an experiment is attractive to MTurk workers.

[6] Note that, in the process of testing the web-based platform and MTurk, I also ran a set of experiments online in August 2012, in which the Unit Rate was a paltry $.01. I do not report the results from that round of experiments here, in part to avoid excessive clutter, and in part because some wrinkles in experimental design and logistics were ironed out between August 2012 and February 2013.

Results

Budget Approval

In both the BETA and BASIC versions, the vast majority of budgets were approved – 97% in BETA, and 96% of budgets offered by Live Proposers in BASIC.[7]

How are budgets divided?

BETA: Overall, budget divisions tended to favor the Public and the Proposer, with the former averaging just over 8 units per period and the latter just under 9, while the high offer to non-Proposer Legislators averaged just over 5 and the low offer just below 2. 47% of budget proposals were minimum-winning, offering some positive amount to the High Legislator and 0 to the Low Legislator. The most frequent Public Offers were at 0, 8, 10, and 12 budget units, with a density in the range around 10, fewer Public Offers in the 2-7 range, and very few Public Offers above 12.

BASIC: Budget divisions were far less variant by player type. Mean offers to Public were right at 7, and to the Proposer were 6.5, whereas mean offers to the High and Low Legislators were 5.5 and 4.9, respectively. Minimum coalitions were also far less frequent in the BASIC version, with only 5% of Live Proposer budgets offering zero units to the Low Legislator. The distribution of Public Offers in the BASIC version was much more heavily concentrated at the modal value of 6 budget units (40%), with another 30% of offers falling in the 7 to 10 range.

What do Legislators approve/reject?

As one would expect, Proposers virtually always (99.5% of cases) vote to approve the budgets they propose in the BASIC version. (The .5% of Nay votes might reflect point-and-click errors.) Non-Proposers, by contrast, are more discriminating, casting 69% Aye and 31% Nay votes. Also as expected, Non-Proposers' likelihood of casting Aye votes corresponded with their budget offers, as illustrated in Figure 14. Live non-Proposers actually voted Aye on budgets about a third of the time even when offered 0 units, but their likelihood of an Aye vote rose to above 80% when offered 6 units, and no non-Proposer voted ever voted Nay on a budget when offered 9 or more units. (The pattern in the BETA version was similar, except rates of Aye votes among non-Proposers jumped above 80% as soon as non-Proposers were offered *anything* greater than zero.)

[Figure 14]

What follows naturally is that budgets that offer at least one non-Proposer Legislator more than 2 units virtually always pass, whereas those that offer both legislators less than 2 units are more vulnerable, as illustrated in Figure 15.

[Figure 15]

Note that Figure 15 aims to shed light on the voting behavior of non-Proposers, so includes budgets offered by Bot Proposers, who occasionally make low-ball offers (see

[7] In reporting from the BASIC version, I report on the behavior of Live players only. Thus, the budget offers reported are only those made by Live Proposers, of budget votes only those of Live Legislators, of Thumbs Up/Down only those by Live Publics, etc. In the BETA version, all players were Live, so this distinction does not apply.

Appendix B). Live Proposers rarely do, so the rate of rejected budgets is higher for Bot Proposers.

What does the Public reward/punish?

As one would expect, the probabilities that the Public rejects Legislators (not reelecting under BETA; rendering a Thumbs Down judgment in BASIC) declined as the Public Offer rose. Public Offer = 6 appeared to be an inflection point for the probability of electoral punishment. For Proposers, Public Offer below 6 triggered rejection 70% of the time, whereas Public Offer at or above 6 triggered rejection only 22% of the time. Non-Proposers were insulated somewhat from retribution for low Public Offers, with rejection at 56% when the Public Offer was below 6, and at 23% when Public Offer>=6. (The pattern in the BETA version was the similar.)

Figure 16 illustrates the rate of Thumbs Up verdicts rendered by the Public on Legislators, by the Public Offer in that period. The left panel shows that, for budgets that failed, the Public's Offer is unrelated to the Thumbs verdict. Recall that the Public receives nothing when budgets fail, regardless of the offer. Among the vast majority of budgets that pass, however, the Public's verdict on Legislators tracks the Public Offer tightly.

[Figure 16]

How does transparency affect how each type of player fares?

The central question the experiment addresses is whether and how transparency affects the representation legislators deliver. Here is where the results from the two versions of the experiment get murky. The central result from BETA was that transparency matters to the relative distribution of the budget among players, and that transparency is the Public's friend. The results from BASIC suggest far less impact for transparency. Figure 17 illustrates the distribution of budget offers to the Public, the Proposer, and to non-Proposer Legislators, under each of the transparency conditions, under both BETA and BASIC

[Figures 17]

In BETA, moving from NT to ST mode almost doubled the mean Public Offer, from 5.2 to 9.3, with an additional boost to 10.3 with the move to FT. Proposers' Offers, by contrast dropped with transparency, but in a non-symmetrical way, falling off only slightly, from 10.0 to 9.7 with a shift from NT to ST, but then dropping further, to 7.5 with the shift to FT. The non-Proposer Legislators, finally, did best under NT (mean high = 6.7, mean low = 2.2), then FT (high=4.5, low=1.8), and worst under ST (4.3 and 0.7). By contrast, the February 2013 results from the web-based BASIC version show little responsiveness to Transparency. The mean Public Offer was 7.6 under NT, 6.4 under ST, and 7.5 under FT. Analysis of variance shows the move from NT all the way to FT – the full spectrum of transparency – to have no measurable impact on Public Offers. Proposers' offers were similarly stable – 6.5, 6.9, and 6.0, respectively, with the shift from invisibility (NT) to visibility (ST) statistically indiscernible. High and Low Legislator offers barely varied, either from each other or across transparency conditions. The mean offers by type of player under each transparency condition, for each version, are shown in Table 2.

[Table 2]

Patterns and learning across periods

Figure 18 shows linear estimations of expected offers to each type of player across all periods, with a separate graph for each transparency condition. The top panel shows the 20 periods played in the BETA version, and the bottom the 10 periods played in the BASIC version.

[Figure 18]

The patterns that are visible are strikingly distinct. First, consider BETA, which fits more consistently initial expectations. Under NT, the Proposer's expected offer is higher than the Public's Offer at the outset, and the expected offers diverge from there, with the Proposer's increasing across periods and the Public Offer decreasing. Under ST, by contrast, the Public's and the Proposer's offers are statistically indistinguishable. Finally, under FT, the Public Offer starts off higher than the Proposer's (by a statistically discernible amount) and the Public maintains that advantage throughout the game. The pattern in BETA over periods suggests that transparency may be an even greater friend to the Public than the aggregate results across all periods suggest, insofar as both the NT and ST graphs suggest that as time horizons stretch, the Proposer's expected share grows relative to the Public's, whereas under FT, the Public's initial advantage appears stable.

Next consider BASIC. Under all conditions, the offers among all four types of players are statistically indiscernible in the initial periods of play. 6-6-6-6, or something close to it, appears to be the default for Proposers in the web-based environment. Some divergence is apparent as players learn the game, but the patterns are not clearly connected to the transparency conditions. In both NT and FT, the Public's Offer increases across periods. The Public's emergence as top dog in FT is consistent with expectations, and not inconsistent with the BETA result, but the parallel pattern under NT is puzzling and inconsistent with BETA.

First-order accountability: Proposers and bad budget offers

The risk that Proposers run for low-balling the Public grows with transparency. On this count, results from BETA and BASIC are consistent.

BETA: Very low Public Offers were a death wish for Proposers under all transparency modes. Proposers who made Public Offers at 2 or below were uniformly rejected, and rates of rejection fell consistently as Public Offers rose. That said, in ST and FT modes, low Public Offers were rarely observed as Proposers, anticipating punishment, offered far more Public-serving budgets than under NT.

BASIC: The participation of the Bots in the BASIC version of the game offers an advantage because the Bots are programmed to offer occasional low-ball budgets to other Legislators and to the Public, even in ST and FT modes. Thus, the BASIC version offers the opportunity to observe Public responses to out of equilibrium behavior by Proposers. Figure 19 show the rate of Thumbs Up judgments by Public Offer on both Non-Proposer and Proposer Legislators, in NT, ST, and FT mode, respectively.

[Figure 19]

Consider NT mode first. The Public knows what she received, but cannot distinguish among the Legislators. The rate of Thumbs Up judgments should rise with the Public Offer, but should not differ from non-Proposers and Proposers. This is, in fact the case. The positive relationship between Public Offer and Thumbs Up verdicts ($p<.00$) is statistically indistinguishable for Proposers versus non-Proposers. (The lines and confidence intervals are superimposed in the top left panel of the figure.) If transparency affects first-order accountability, however, we should expect a divergence between the verdicts rendered on Proposers and non-Proposers once Proposers can be identified, under ST and FT. Specifically, verdicts on *Proposers* should remain closely tied to Public Offers, while those on non-Proposers might vary more. The distinction should be particularly relevant under FT, where the Public can see each Legislator's vote and judge on matters other than the Public Offer. The bottom panel of Figure 21 provides some support for this expectation. Proposers' fates remain closely tied to Public Offers under all three transparency conditions, but thumbs judgments on non-Proposers and Proposers diverge under FT, where non-Proposers' fates appear nearly unrelated to Public Offers. The question, then, is whether second-order accountability explains this.

Second-order accountability: Non-proposers and bad budgets

Proposers catered to the Public more assiduously when they were visible in the BETA version of the experiment, but responded to transparency less predictably in the BASIC version. What about non-Proposer Legislators? Part of the logic of the treatment conditions is to expose the non-Proposers to possible electoral sanction for voting in favor of budgets that serve the Public poorly – or voting against ones that serve the Public well.

Under NT or ST, non-Proposers are anonymous, so we might expect their voting decisions to be driven primarily by how well the budget offer treats them. Recall that if a proposed budget fails, all players receive nothing. Under NT, Proposers and non-Proposers are indistinguishable to the Public. If the Public is expected to punish a failed budget by rejecting the team of legislators that engineered it, there is little reason for non-Proposers to vote against budgets under NT.[8] Under ST, the Public can distinguish the non-Proposers from the Proposer, but does not see individual votes (or the offers, other than the Public Offer), so has limited ability to reward good legislative behavior beyond the proposal.

FT, by contrast, is designed to confront non-Proposers facing budgets that serve the Public poorly (and in doing so, that may serve themselves well) with the dilemma that approving 'bad' budgets may invite electoral punishment. Is there evidence that such budgets put non-Proposers in a tough spot? One can estimate of the effect of transparency on non-Proposer votes by comparing logit regressions of:

$$\Pr(\text{Vote}=1) = a(\text{Constant}) + b1(\text{Legislator's Budget Offer}) + b2(\text{Public Offer}).$$

[8] This was strictly true under the BETA version, where Proposers were selected at random for each period. In the BASIC version, non-Proposers could be motivated to shoot down budgets, thus dislodging the Proposer and possibly replacing her. The rarity of failed budgets suggests this risky strategy was rarely employed by non-Proposers, but further analysis will examine the possibility.

under the various transparency conditions.

Table 3 shows – for both BETA and BASIC versions – the coefficients and standard errors for the variables of interest in these logit regressions, along with the estimated change in the likelihood of a Yes vote from shifting each independent variable from its 20th percentile value to its 80th percentile value, with other variables in the equation held constant at their mean values.[9]

[Table 3]

It comes as no surprise that raising a legislator's budget offer increases her propensity to support that budget under all transparency conditions. The estimated effects of greater interest, however, are the responsiveness of non-Proposers' votes to Public Offers – and particularly, the difference between that responsiveness under FT versus under NT and ST modes. The expectation associated with H3, recall, is that Public Offers should affect non-Proposers' votes under FT, but not under NT or ST.

As it happens, there is consistent evidence in both BETA and BASIC versions of sensitivity to Public Offers in FT. When Legislators' votes were fully transparent, shifting from stingy (20th percentile) to generous (80th percentile) Public Offers increased the likelihood of Aye votes by 12% in BETA, and 10% in BASIC (both significant at .01), even when the Legislators' own budget offer was held constant. But there is also patchy evidence of such sensitivity under NT and ST, where Legislators' votes are not visible to the Public. Non-Proposers' votes were highly responsive to Public Offers in NT in the BETA experiment, and appeared to be as well in the BASIC version (although the latter estimate fell just short of conventional significance, at .06). The reverse pattern held for ST, where non-Proposers in the BETA version showed no responsiveness to Public Offers, whereas they did in the BASIC version.

So the voting behavior of non-Proposers suggests they act *as though* the Public is watching their votes – with a high degree of certainty under FT, which is consistent with expectations – but also possibly under NT and ST, which is not. What is the Public doing? If the Public is monitoring non-Proposer votes when they are visible, and rewarding or punishing accordingly, then we should see:

- opposing patterns of Legislators' Thumbs Up rates across Public Offers, conditional on whether the Legislator voted Aye or Nay under FT;
- no difference in these patterns under NT or ST.

Figure 20 supports precisely this expectation, suggesting that Live Publics are watching non-Proposers when they can. The panels of the figure show, for each transparency mode, the rates of Thumbs Up verdicts passed by Live Publics on non-Proposers, by the Public Offer for the period and conditional on the Legislator's vote on the budget.[10] Under NT and ST, where votes are not visible, non-Proposers who votes Aye and Nay on budgets are statistically indistinguishable. Under FT, however, their fates diverge as

[9] For the BETA version, the votes from the last period of each 20-period experiment were dropped because no threat of electoral punishment existed in last periods of that version.

[10] Thumbs verdicts on Proposers are not graphed to reduce clutter. Proposers virtually always vote Aye, and their Thumbs verdicts always track Public Offers strongly.

Public Offers grow. As Public Offers surpass 5, non-Proposers who vote Nay are measurably more likely to attract Thumbs Down verdicts than are those who vote Aye.

[Figure 20]

The results regarding second-order accountability are suggestive, but not dispositive. Publics show signs of monitoring votes under FT in a way that would produce such accountability, but Legislators do not appear consistently to distinguish FT from other transparency conditions. It may be that the implications of the transparency conditions are too subtle for non-Proposers to appreciate fully the differences in their exposure to punishment – or to expect that Publics will distinguish non-Proposers from Proposers. Alternatively, it might be that the current versions of the game have not sufficiently motivated non-Proposers to scrutinize offers and transparency conditions

Minimum coalitions versus universalism

Does transparency affect the likelihood of minimum coalitions – those that shut out one Legislator? The BETA version of the experiment strongly suggests it does, but here again, the BASIC version shows no clear pattern. Table 4 shows that the percentage of minimum coalitions for both the versions.

[Table 4]

The big difference is the far higher incidence of minimum coalition offers in BETA than in BASIC. Proposers in BETA, where each budget unit was worth US$.50, were aggressive in cutting the third Legislator out the bargain, particularly when the details of budget offers were not visible to the Public, as under NT or ST. When the details of the proposal were visible, under FT, Proposers were more egalitarian in their distributions across coalition partners. By contrast, Proposers in BASIC almost always cut every Legislator in on at least some piece of the budgetary pie. The relative inclusiveness of BASIC Proposers may stem from the fact that the budget unit exchange rate was too low. Alternatively, it could be that the shift in the rule for determining the Proposer (repeated random selection in BETA; continue if the budget passes in BASIC) induced more inclusiveness. Future iterations of the experiment, with higher stakes, may shed some light on whether one of these factors or another accounts for the reduced incidence of minimum coalitions.

Rates of punishment

Rates of overall reward and punishment varied substantially by transparency mode in BETA, but again, not in BASIC. In BETA, rates of reelection were lower (around 50%) under NT than under ST or FT (around 75% for both Proposers and non-Proposers). In BASIC, rates of doubling among live Legislators were 72% in NT, 66% in ST, and 76% in FT.

The pattern in BETA was that the lower the transparency, the more the Public sanctioned Legislators – but to less effect, judging from the Public's budget shares under each mode. Transparency thus appeared to make the Public's sanction of Legislators more effective, in turn minimizing the need for it to be exercised. This is consistent with the general results from formal models of transparency and legislative responsiveness (Snyder & Ting 2005; Humphreys & Weinstein 2008), but contrary to much conventional wisdom in the United States, which often equates high rates of reelection with insulation

of legislators from public sanction, and with low levels of accountability. This conclusion, however, must be regarded as thoroughly tentative, given the absence of a correlation between overall punishment rates and transparency in the BASIC version.

Discussion

U.S. citizens tend to take legislative transparency for granted, but it varies tremendously around the world and is largely absent even in many democracies. Knowing whether and how transparency matters to the sort of representation legislators provide is important – to theory, but also (and moreso) to politics, not least because transparency varies tremendously and is low in many legislatures around the world, but can be modified easily relative to other elements of the relationship between constituents and legislators, such as constitutional structure, the strength of political parties, the demographics of politicians, or the financing of campaigns. Machines that record proposals and votes can instantly produce records of legislative behavior that can be made available on the internet or disseminated by other media to journalists, citizens, interest groups, other politicians, and academics. If transparency improves representation, then to the extent that non-transparent practices are due to the absence of these technologies, the remedy is straightforward. Even where obstacles to transparency are more formidable (read: political), estimating its effects on public welfare, on the distribution of resources among politicians, and on how the accountability mechanism operates, presents a clearer picture of the potential effects of reforms to enhance, or limit, how easily and at what level of detail citizens can peer inside the legislative process.

In terms of logistics, the execution of the experiments discussed in this paper was reasonably successful both in the lab-based BETA and the web-based BASIC version. The two sets of experiments provide plenty of indications that basic strategic problems the experiment presents engaged participants, and some signs that the transparency manipulation affects behavior. But there are also inconsistencies between the BETA and BASIC versions that are stubbornly puzzling. In the BETA version, at least, transparency appears to be good for the Public. The greater the transparency, the higher the Public's budget shares (H1, above) and Public Offers (H2). The direction of the effect is not particularly surprising, but the magnitude of the effect was impressive. That said, the absence of any measurable effect of transparency on Pubic Offers in the BASIC version was, then, all the more mystifying. There is evidence in the results of both first-order accountability and second-order accountability operating as expected. Proposers are rewarded by Publics in accord with the budget offers they make. Non-Proposers are rewarded in accord with their votes on budgets *when their votes are visible*, and not otherwise. At the same time, Proposers (in the BASIC version) did not respond to first-order accountability by treating the Public better when their proposals were identifiable than when they were anonymous. And non-Proposers acted as though they thought the Public could monitor their votes even when the transparency conditions did not allow it.

Looking ahead

Beyond BASIC and BETA, at least six variants of the game I would like to test, pending further development of the web-based platform.

- *SMD Version*: This version involves 6 participants – 3 members of the Public (P1, P,2, P3) and 3 Legislators (L1, L2, L3), such that, during play, P1 may

sanction L1, P2 may sanction L2, and P3 may sanction L3. 10 periods with 1 budget per period.

- *BASIC Multi-Budget*: 4 participants (P, L1, L2, L3), and 5 periods of play, but 3 budgets are divided per period. That is, each period consists of 3 iterations of a Proposer being randomly selected, making a proposal, the proposal being voted on, and all players being informed of the outcomes. But the Public only is offered the opportunity to vote Thumbs Up/Down on Legislators only after each third budget.

- *SMD Multi-Budget*: 6 participants and SMD-style representation, with multiple budgets per period.

- *BASIC Transparency Varying*: 4 participants, 10 1-budget periods, but after the Proposer is randomly selected in each period, the Proposer selects the transparency mode in which that period will be played.

- *BASIC Multi-Budget Transparency Varying*: 4 participants, 3 budgets per period. Proposer is selected for each period (3 budgets), but selects transparency mode separately for each budget.

Each different variant of the game adds at least one element of theoretical interest, although also some additional complexity. The move from 1 Public to 3 allows for Legislators' fiduciary responsibilities, and accountability, to be focused on separate subsets of citizens, so the potential for coalitions that freeze some players out (both Legislators and Publics) increases. The move to multi-budget periods more closely approximates how legislatures operate, making multiple decisions between opportunities for citizens to sanction their representatives. The cognitive load on citizens increases, and the prospects grow for passing less popular policies early in each period, with more 'populist' appeals coming near to the point of reckoning. Allowing Proposers (i.e. legislative leaders) to select their level of transparency is also a step toward external validity, as legislatures generally set their own rules regarding transparency. In this case, we can observe whether Publics are willing to punish Proposers not only for their budget decisions, but for their transparency decisions – and whether Proposers, anticipating such punishment select transparency accordingly.

After quite striking results from the lab-based BETA version, the results from BASIC are muddier. Whether I pursue these future versions depends in part on whether I can resolve the ambiguities surrounding current BASIC version in a way that would justify the considerable effort and expense.

References

Baron, David P. and John A. Ferejohn. 1989. "Bargaining in legislatures." *American Political Science Review* 83(4):1181-1206.

Baumgartner, Frank R., Bryan D. Jones, and John Wilkerson. 2011. "Comparative studies of policy dynamics." *Comparative Political Studies* 44(8):947-972.

Bentham, Jeremy. 1791. *Essay on political tactics containing six of the principal rules proper to be observed by a political assembly.* London: T. Payne.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating online labor markets for experimental research: Amazon.com's Mechanical Turk." *Political Analysis* 20(3):351-368.

Bianco, William T., David B. Spence, and John D. Wilkerson. 1996. "The Electoral Connection in the Early Congress: The Case of the Compensation Act of 1816." *American Journal of Political Science* 40 (1): 145-71.

Birnbaum, Jeffrey and Alan Murray. 1988. *Showdown at Gucci Gulch: Lawmakers, lobbyists, and the unlikely triumph of tax reform.* New York: Vintage.

Burke, Edmond. 1774. *The Works of the Right Honourable Edmund Burke.* London: Henry G. Bohn.

Carey, John M. 2003. "Transparency Versus Collective Action: The Fujimori Legacy and the Peruvian Congress." *Comparative Political Studies* 36(9): 983-1006.

Carey, John M. 2009. *Legislative Voting and Accountability.* Cambridge University Press.

Carrubba Clifford J., Matthew Gabel, and Simon Hug. 2008. "Legislative voting behavior, seen and unseen: A theory of roll-call vote selection." Legislative Studies Quarterly 33(4):543-572.

Cox, Gary W. and Scott Morgenstern. 2001. "Latin America's reactive assemblies and proactive presidents." *Comparative Politics* 33(2):171-190.

Crisp, Brian F. and Amanda Driscoll. 2010. "The importance of legislative voting procedures." Unpublished paper.

Diermeier, Daniel and Rebecca Morton. 2003. " Proportionality versus Perfectness: Experiments in Majoritarian Bargaining. Unpublished paper.

Dumitrescu, Delia, and Andres Blais. 2011. "Increased realism at lower cost: The case for the hybrid experiment." *PS: Political Science & Politics* 44(3):521-523.

Fehr, E. and Gachter, S. 2000. "Cooperation and punishment in public goods experiments." *American Economic Review* 90:980–994.

Fischbacher, Urs, et.al. 2001. "Are people conditionally cooperative? Evidence from a public goods experiment." *Economic Letters* 71:397–404.

Fischbacher, Urs. 2007. "z-Tree: Zurich toolbox for ready-made economic experiments." *Experimental Economics* 10(2):171-178.

Fréchette, Guillaume R., John H. Kagel, and Steven F. Lehrer. 2003. "Bargaining in Legislatures: An Experimental Investigation of Open versus Closed Amendment Rules." *American Political Science Review* 97(2):221-232.

Fréchette, Guillaume R., John H. Kagel, and Massimo Morelli. 2005a. "Nominal bargaining power, selection protocol, and discounting in legislative bargaining ." *Journal of Public Economics* 89(8):1497-1517.

Fréchette, Guillaume R., John H. Kagel, and Massimo Morelli. 2005b. "Gamson's Law versus non-cooperative bargaining theory." *Games & Economic Behavior* 51(2):365-390.

Henrich, J. et. al. 2005. "Economic man" in cross-cultural perspective: Behavioral experiments in 15 small-scale societies." *Behavioral and Brain Sciences* 28:795-855.

Henrich, J. et.al. 2010a. "Costly punishment across societies." *Science* 312:1467-1470.

Henrich, J. et.al. 2010b. "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment." *Science* 327:1480-1484.

Herrmann, Benedikt, Christian Thöni, and Simon Gächter. 2008. "Anti-social punishment across societies." *Science* 319:1362-1367.

Hix, Simon, Sara Hagemann, and Doru Frantescu. 2011. "The Effect of Transparency on Legislative Behaviour: Report on an Experiment in the European Parliament." American Political Science Association conference paper.

Hug, Simon. 2010. "Selection effects in roll call votes." *British Journal of Political Science* 40(1):225-235.

Humphreys, Macartan and Jeremy Weinstein. 2008. "Policing politicians: Citizen empowerment and political accountability in Uganda." Unpublished paper.

Kile, Orville M. 1948. *The Farm Bureau through three decades.* Baltimore: Waverly Press.

Lepore, Jill. 2008. "Rock, paper, scissors: How we used to vote." *The New Yorker*, October 13.

Linz, Juan J. 1994. Presidentialism or Parliamentarism: Does It Make a Difference? In *The Failure of Presidential Democracy: The Case of Latin America*, vol. 2, ed. Juan J. Linz and Arturo Valenzuela. Baltimore: Johns Hopkins University Press. 3-87.

Llanos, Mariana, and Constaza Figueroa Schibber. 2008. "La participacion de la presidencia y el senado en el nombramiento del poder judicial." *Desarrollo Economico* 47(188):607-637.

Malesky, Edmund, Paul Schuler, and Ngoc Anh Tran. 2010. "The adverse effects of sunshine: A field experiment testing the impact of transparency on delegate performance in an authoritarian assembly." American Political Science Association conference paper.

Morgenstern, Scott. 2003. *Patterns of legislative politics : Roll call voting in Latin America and the United States.* New York: Cambridge University Press.

Morton, Rebecca and Kenneth Williams. 2010. Experimental political science and the study of causality: From nature to the lab. New York: Cambridge UP.

Niou, Emerson M.S. and Peter C. Ordeshook. 1985. "Universalism in Congress." *American Journal of Political Science* 29(2):246-258.

Odegard, Peter H. 1928. *Pressure politics: The story of the Anti-Saloon League.* New York: Columbia University Press.

Rose-Ackerman, Susan. 1999. *Corruption and government: Causes, consequences, and reform.* New York: Cambridge University Press.

Saalfeld, Thomas. 1995. "On Dogs and Whips: Recorded Votes." *Parliaments and majority rule in Western Europe.* Herbert Doring, ed. New York: St Martin's Press: 528–565.

Samuels, David J. 2002. "Presidentialized parties." *Comparative Political Studies* 35:461-483.

Schumpeter, Joseph Alois. 1942. *Capitalism, socialism, and democracy*. New York: Harper.

Siavelis, Peter. 2000. *The president and congress in postauthoritarian Chile : institutional constraints to democratic consolidation.* University Park: Pennsylvania State University Press.

Skeen, C. Edward. 1986. "Vox Populi, Vox Dei: The Compensation Act of 1816 and the Rise of Popular Politics." *Journal of the Early Republic* 6 (253).

Smith, Steven S. 1989. *Call to order: Floor politics in the House and Senate.* Washington, DC: Brookings Institution Press.

Snyder, James M., and Michael M. Ting. 2005. "Why Roll Calls? A Model of Position Taking in Legislative Voting and Elections." *Journal of Law, Economics, and Organization* 21 (1): 153-78.

Stasavage, David. 2004. "Open-door or closed-door? Transparency in domestic and international bargaining." *International Organization* 58(4):667-703.

Stasavagae, David. 2007. "Polarization and publicity: Rethinkng the benefits of deliberative democracy." *Journal of Politics* 69(1):59-72.

United States Supreme Court. 1958. *Majority Decision in NAACP v. Alabama Ex Rel. Patterson, Attorney General.* No.91 Supreme Court of the United States 357 U.S. 449; 78 S. Ct. 1163.

Appendix A

Instructions for Online Legislative Budget Experiment

February 2013

BASIC Version

This experiment is part of a study of decision making in legislatures. You will play a game in which players divide a budget. At the end, participants will get paid according to how many budget units they accumulate.

It should take about 5-10 minutes to read and understand the instructions, then 10-15 minutes to play the game. After you read the instructions, you must answer 2 questions about the game correctly before you can play.

The game itself consists of 10 periods of play. Before the first period begins, all participants will be informed whether the experiment will be conducted under conditions of ***Full Transparency, Semi-Transparency***, or ***Non-Transparency*** (explained below). Once the condition is assigned, all periods will be played under that condition.

There are 4 participants. At the outset, 1 will be selected at random to be the Public; the other 3 will be Legislators.

----------[NEXT SCREEN----------

<u>The Budget Proposal</u>
In each period, 1 of the Legislators will be the Proposer. The Proposer will be asked to divide a budget, however she or he wants, among the 4 participants -- the Public and the 3 Legislators.

- The available budget in every period is 24 units.
- Units may not be divided; proposals must be made in full units.
- The Proposer may give some part of the budget to every player, but is not required to do so.

<u>The Legislative Vote</u>
The Legislators then see the budget proposal and vote to approve or reject it. (Note that the Proposer is a Legislator and has a vote, but the Public does *not* vote on whether the budget is approved.)

- If a majority (2 or 3) of the Legislators approves, then the budget passes.
- Otherwise, the budget is rejected, and all players get a zero payoff for that period.

<u>The Proposer</u>
In the first period, one of the Legislators will be chosen at random to be the Proposer. From that point on, if the budget in any period is approved, the Proposer continues in that role in the following period. But if the budget is rejected, then the Proposer loses that role and, at the beginning of the next period, one of the other 2 Legislators is chosen at random to be the Proposer.

----------[NEXT SCREEN----------

The Public's Information

The Public is then informed about the budget.

- If the game is being played under ***Full Transparency***, the Public is informed of:
  - the identity of the Proposer (L1, L2, or L3);
  - how much the proposed budget offered to each of the Legislators, as well as to the Public;
  - how each Legislator voted (Approve or Reject) on the budget proposal.

- If the game is being played under ***Semi-Transparency***, the Public is informed of:
  - the identity of the Proposer;
  - whether the budget was approved or rejected;
  - the Public's own share of the budget.

- If the game is being played under ***Non-Transparency***, the Public is informed of:
  - whether the budget was approved or rejected;
  - the Public's own share of the budget.

----------[NEXT SCREEN----------

The Public's Vote

After being informed, the Public votes "Thumbs Up" or "Thumbs Down" on each Legislator. For each Legislator, a "Thumbs Up" vote increases by 10% his chance of being rewarded at the end of the game. Legislators who are rewarded have their total budget payoffs doubled.

For example:

- If the Public votes "Thumbs Down" to a given Legislator in all 10 periods of the game, the Legislator has zero chance of being rewarded with a double payoff.
- If the Public votes "Thumbs Up" to a given Legislator in all 10 periods, the Legislator has 100% chance of being rewarded.
- If the Public votes "Thumbs Up" in 5 periods and "Thumbs Down" in 5, the Legislator has a 50% chance of being rewarded.
- … and so on.

After Each Period

Each participant is informed of her/his payoff for that period. Each Legislator is informed of whether the Public voted "Thumbs Up" or "Thumbs Down" for her/him.

----------[NEXT SCREEN----------

Completing the Experiment

After all 10 periods are complete, the computer determines, on the basis of each

Legislator's "Thumbs Up" and "Thumbs Down" tally, whether her/his total budget payoff is doubled. Each participant is informed of her/his total payoff, then asked to fill out a short survey. Participants must complete all 10 periods, plus the survey, to be paid.

Note that completing any period requires action by all 4 participants. If one participant gets distracted, others should not be required to wait. Therefore, once the experiment begins, any participant who fails to respond to a prompt within the allowed period of time will be disqualified and immediately replaced with another player so the game may proceed. You are allowed more time in the first couple of periods, as you are learning the game, and a bit less time after that, to keep things moving. A countdown clock will be visible on your screen to let you know how much time you have left.

Payment

If you complete the experiment and the survey, you are guaranteed a payment of $.50, plus your share of the budget payoff pool. Each budget unit is worth $.01. There are 24 budget units available in each period, and 10 periods of play, for a total of 240 budget units. But of course budget units that are awarded to Legislators can be doubled at the end of the game, so the maximum total payoff is 480 units, or $4.80 (plus $.50 guaranteed to each participant). The total payoffs will likely be less, depending how much goes to the Public, how much goes to Legislators whose payoffs are not doubled, and how many budgets simply do not pass.

[BENNET: PLEASE ADVISE ON HOW PARTICIPANTS SHOULD COMPLETE THEIR TASK TO GET PAID.]

----------[NEXT SCREEN----------

QUIZ QUESTIONS

QUESTIONS

2 of the following questions should be drawn at random for each participant. If the participant does not answer *both* questions correctly, she/he should be given the option to return to the first screen of Instructions, click through them again, and then answer 2 different questions drawn from this pool. Once a participant answers a set of questions correctly, she/he may proceed to play the game. If a participant fails 3 times, she/he should be dropped from the experiment.

Questions

How many budget units are available in each budget?

a) 5
b) 24
c) 100
d) 1,000

Of the participants in the game, how many are Legislators?

a) 1
b) 2
c) 3
d) 4

True or False: The Proposer must offer some part of the budget to every player.

a) True
b) False

A player is disqualified if she/he fails to respond to a prompt within:

a) 3 seconds
b) 60 seconds
c) 30 minutes
d) 24 hours

Getting more "Thumbs Up" votes increases a Legislator's chances of:

a) Becoming the Proposer
b) Becoming the Public
c) Being elected President
d) Getting rewarded with double budget units at the end of the game.

After the game ends, in order to get paid, a participant must:

a) Complete a short survey
b) Write a long essay
c) Have won more points than any other player

APPENDIX B: Script for Robot Player in February 2013 Experiments
(in case of need to replace an original participant during experiment)

If Robot is Proposer, offer each of the following budgets with 1/20 probability:

Public	Proposer	Non-Proposer_Leg_1	Non_Proposer_Leg_2
13	9	2	0
14	8	2	0
11	10	3	0
12	9	3	0
9	11	4	0
10	10	4	0
7	12	5	0
8	11	5	0
5	13	6	0
6	12	6	0
3	17	2	2
4	16	2	2
1	18	3	2
2	17	3	2
4	14	4	2
5	11	4	4
6	9	5	4
7	8	5	4
8	6	6	4
9	5	6	4

If Robot is a Legislator, vote on budget proposals by following rules:

- If Legislator is Proposer, always vote Approve.
- If Legislator is non-Proposer, then:
    - Under Non-Transparency or Semi-Transparency, vote Approve if offered >1; else vote Reject.
    - Under Full Transparency, vote Approve if Public is offered > 5; else vote Reject.

If Robot is Public, give Thumbs Up/Down as follows:

Non-Transparency
- If Public pay-off <6, then Thumbs Down for all Legislators.
- If 6 <= Public pay-off <8, then Thumbs Down each Legislator with 50% probability.
- If Public pay-off >=8, then Thumbs Up all Legislators.

Semi-Transparency
- If Public pay-off <6, then Thumbs Down for Proposer and Thumbs Up/Down each at 50% probability for other Legislators.

- If 6 <= Public pay-off <8, then Thumbs Down Proposer with 50% probability. Thumbs Up other Legislators.
- If Public pay-off >=8, then Thumbs Up all Legislators.

Full Transparency

- If Public pay-off <6, then Thumbs Down Proposer and any other Legislator who voted to Approve budget. Thumbs Up to any Legislators who voted to Reject budget.
- If 6 <= Public pay-off <8, then Thumbs Down Proposer and any Legislator who votes to Approve budget with 50% probability. Thumbs Up to any Legislators who voted to Reject budget.
- If Public pay-off >=8, then Thumbs Up all Legislators.

NOTE: When a Robot plays, set response times as follows:

- Make a budget proposal: 10 seconds
- Vote to Approve/Reject a budget proposal: 5 seconds
- Enter Thumbs Up/Down votes: 5 seconds

Figure 1.

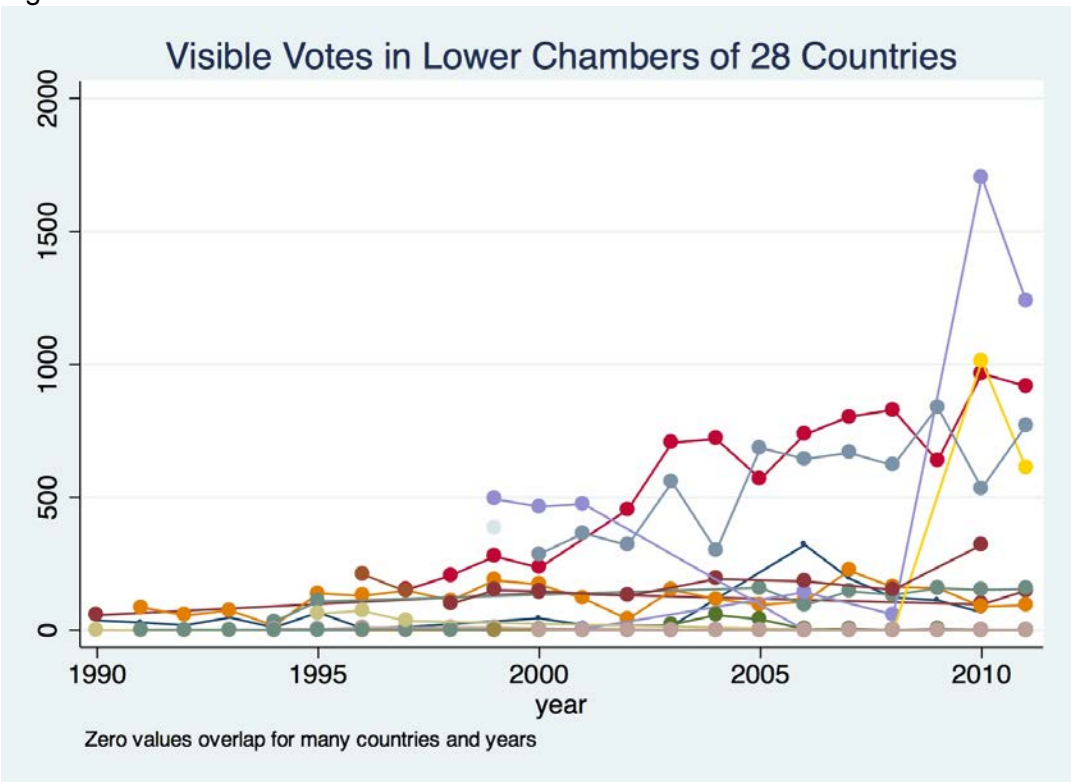


Figure 2.

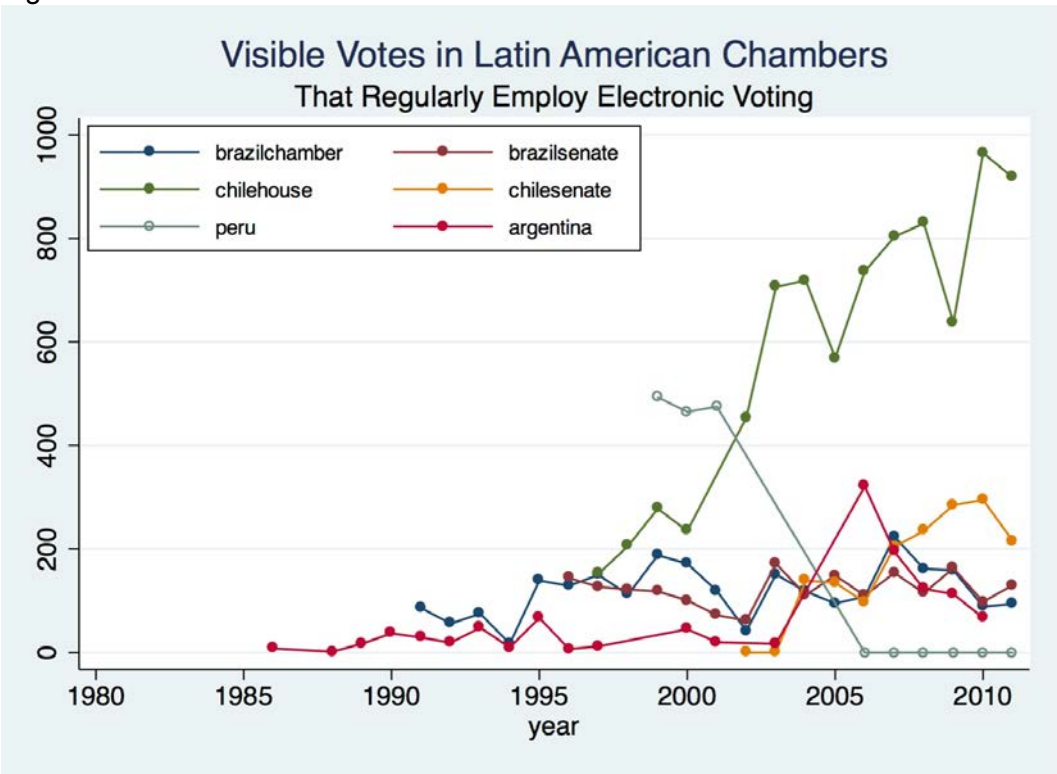


Figure 3.

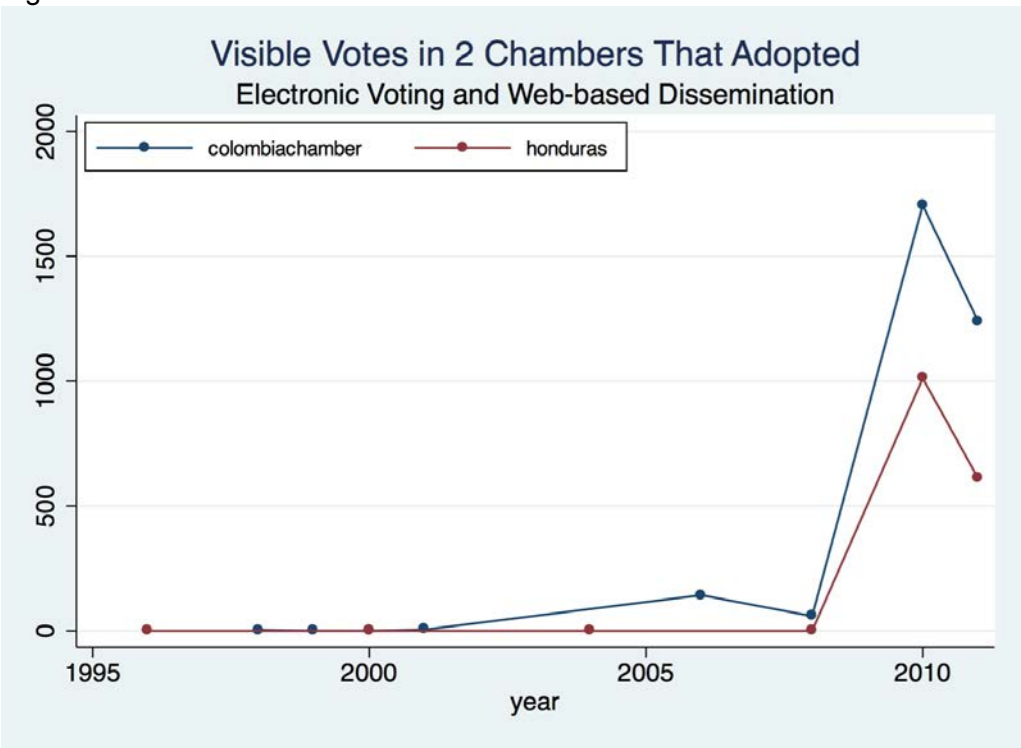


Figure 4.

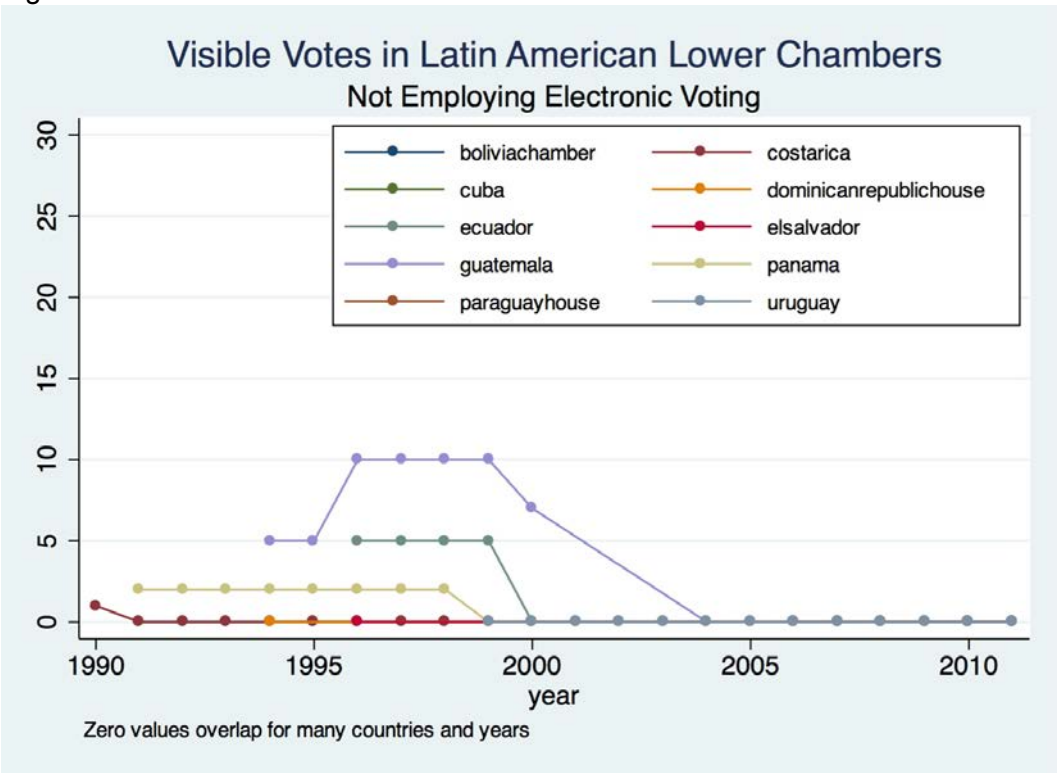


Figure 5.

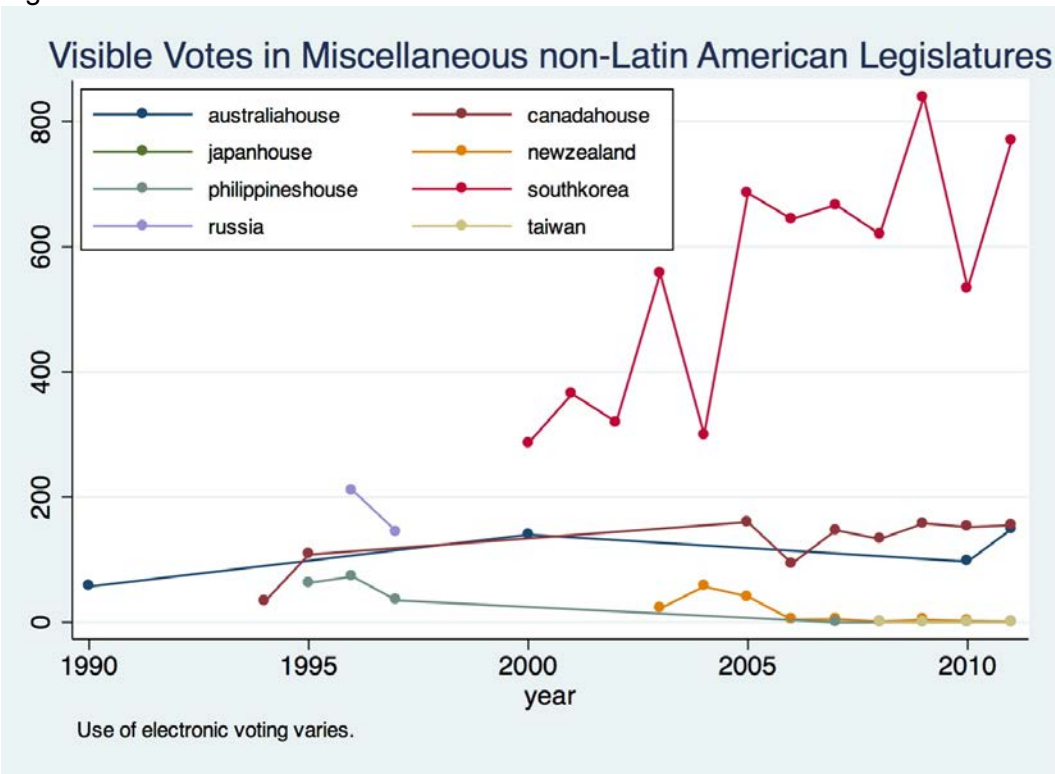


Figure 6.

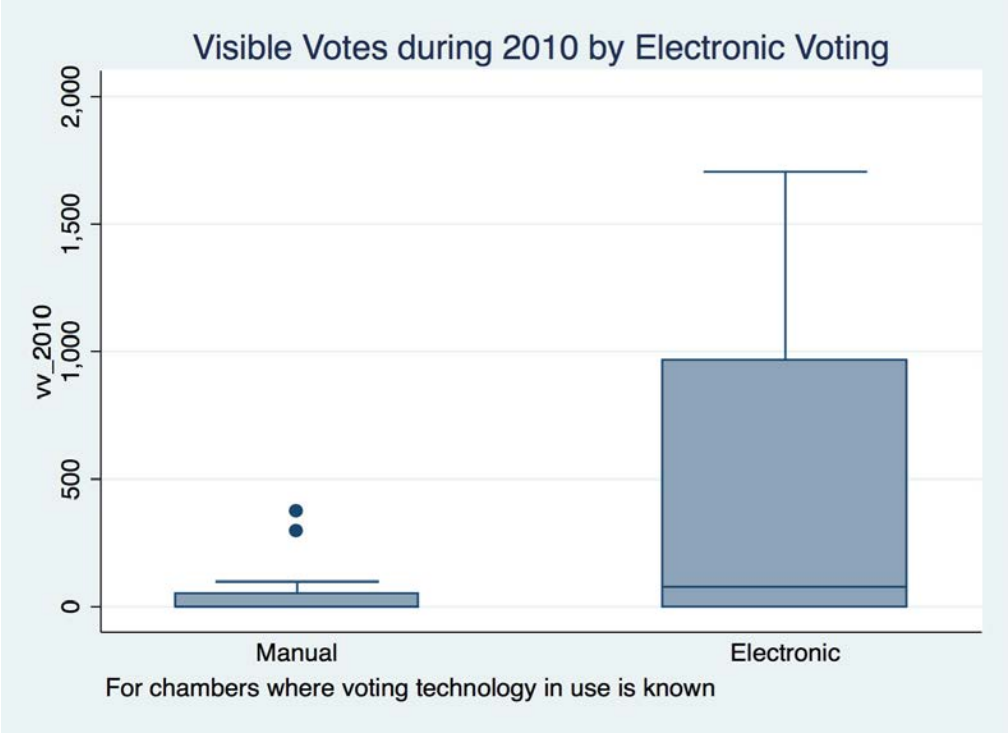


Figure 7

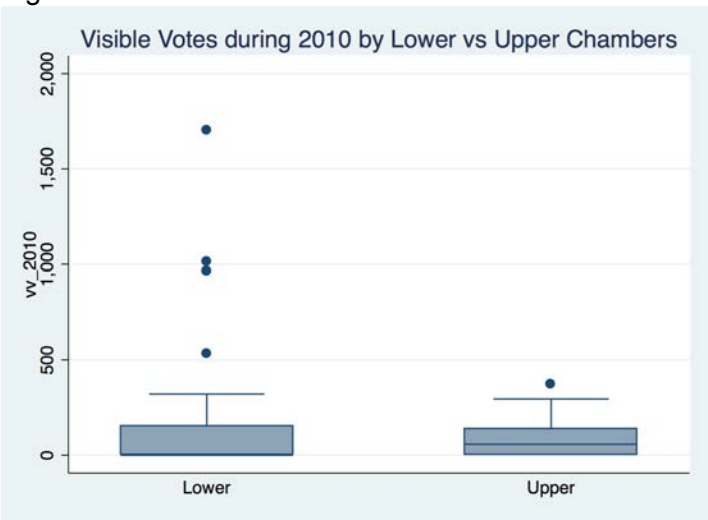


Figure 8

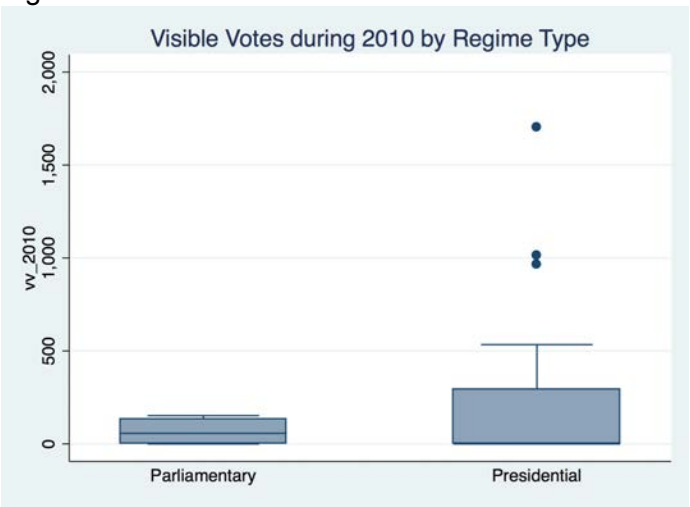


Figure 9

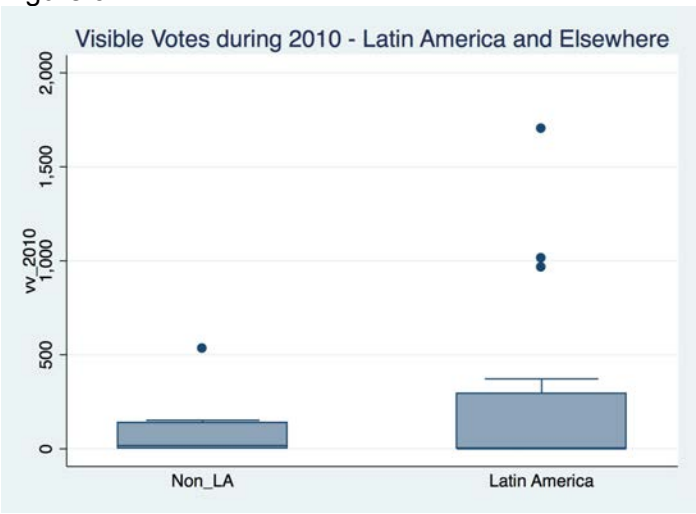


Figure 10

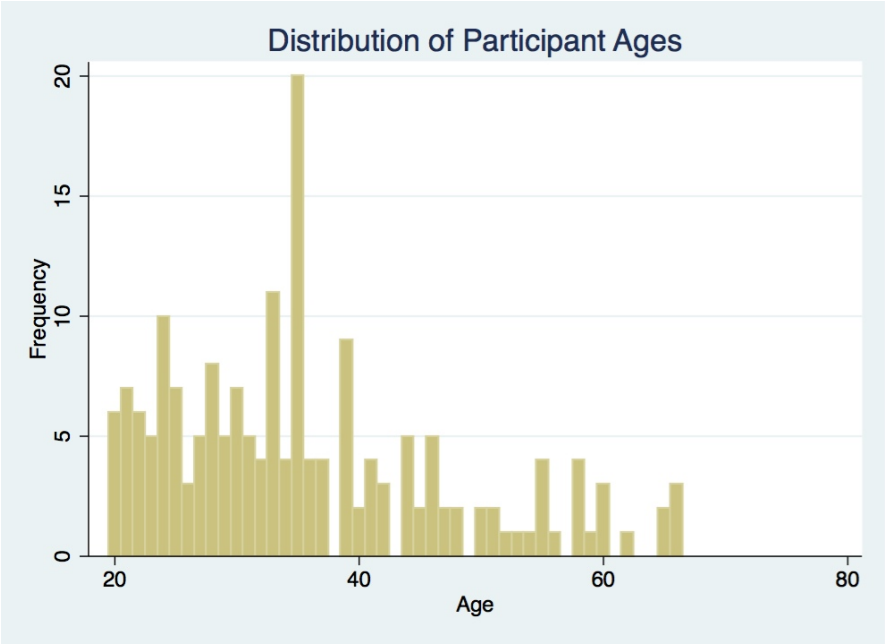

Distribution of Participant Ages
Frequency
0
5
10
15
20
20
40
60
80
Age

Figure 11. US versus Indian MTurk populations

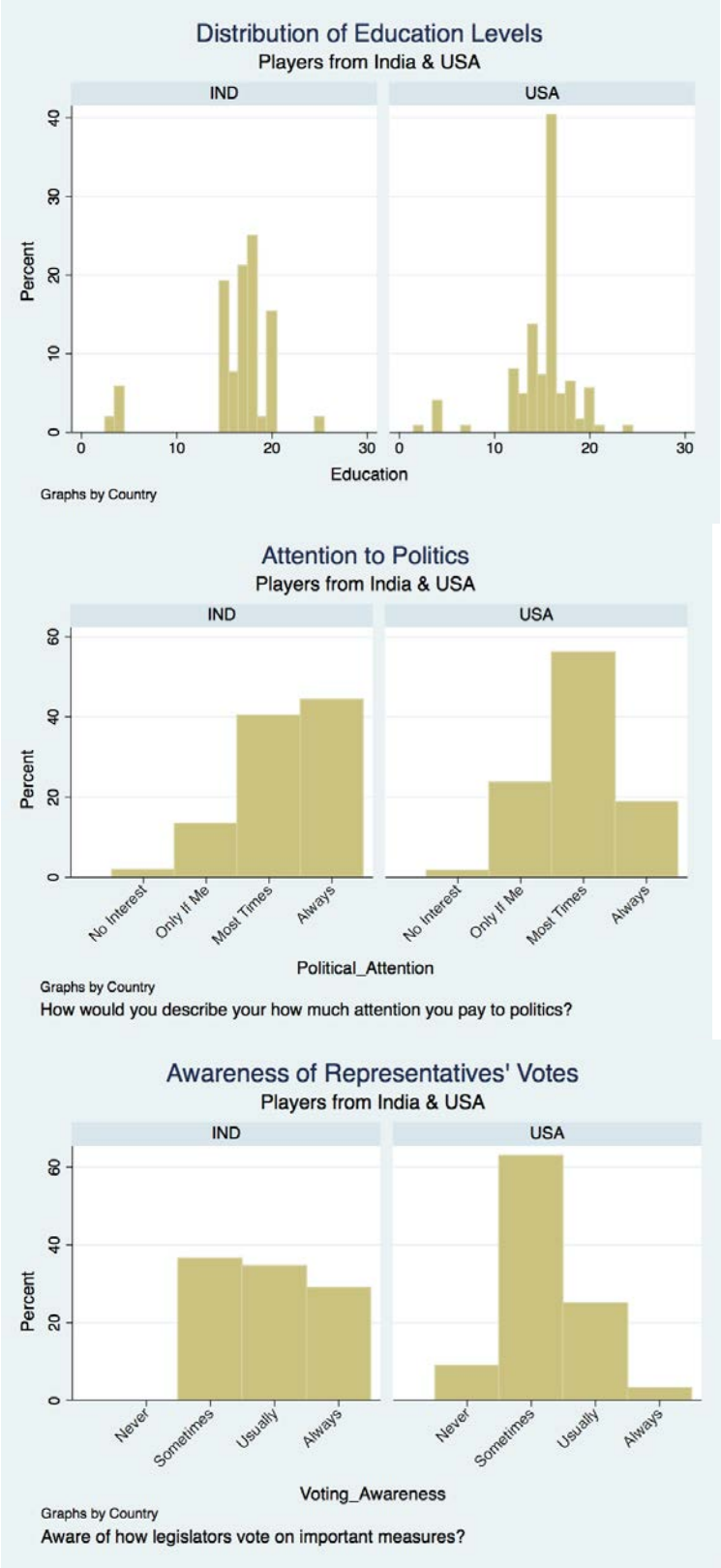


Figure 12. BASIC

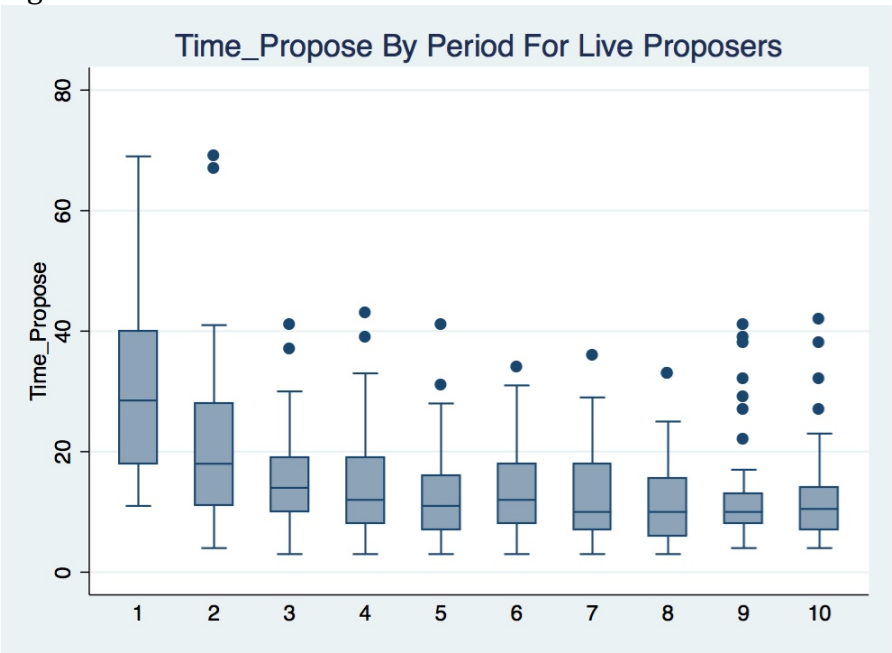


Figure 13. BASIC – Live Player Density

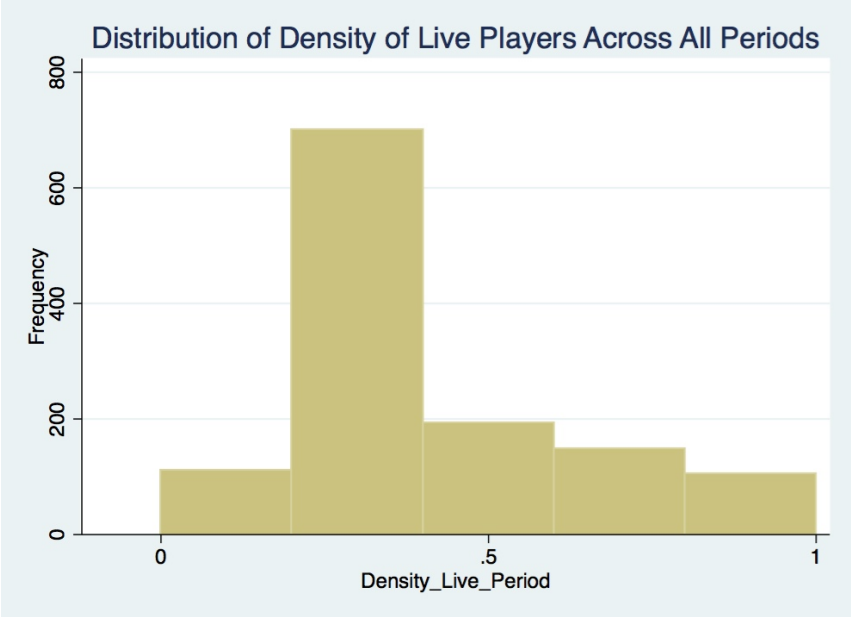


Figure 14: BASIC

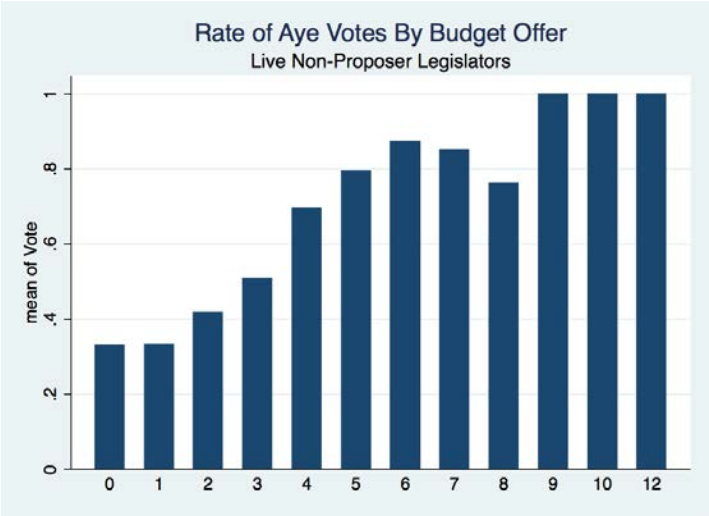


Figure 15. BASIC

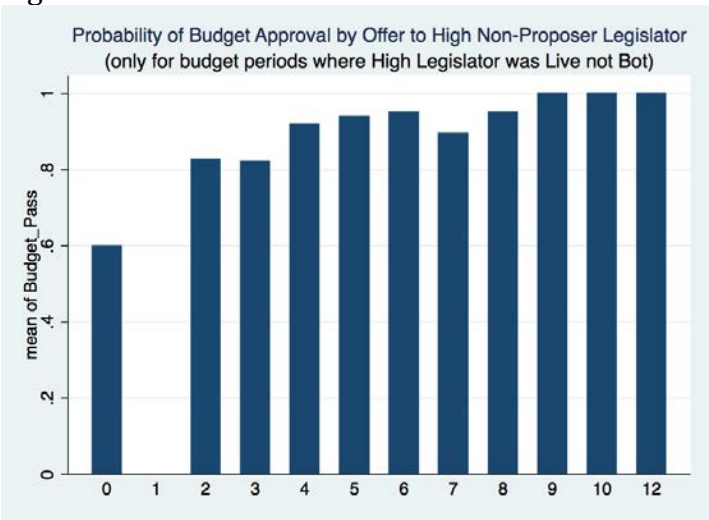


Figure 16: BASIC

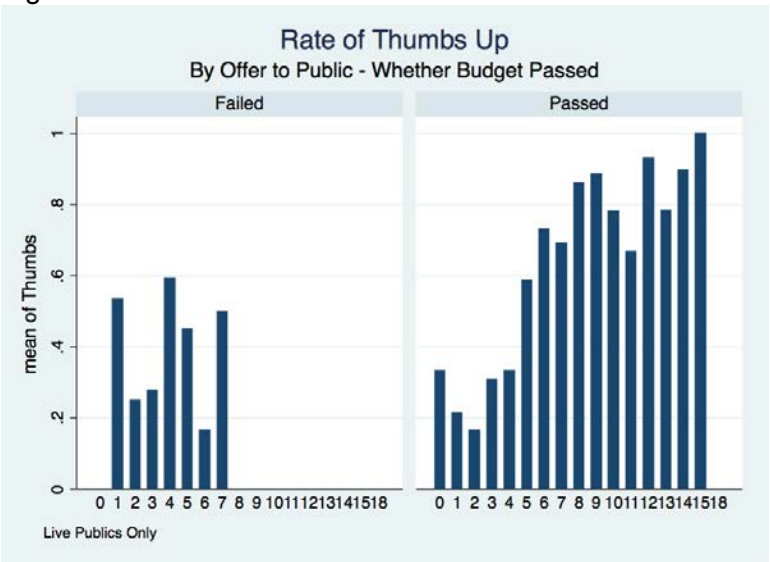
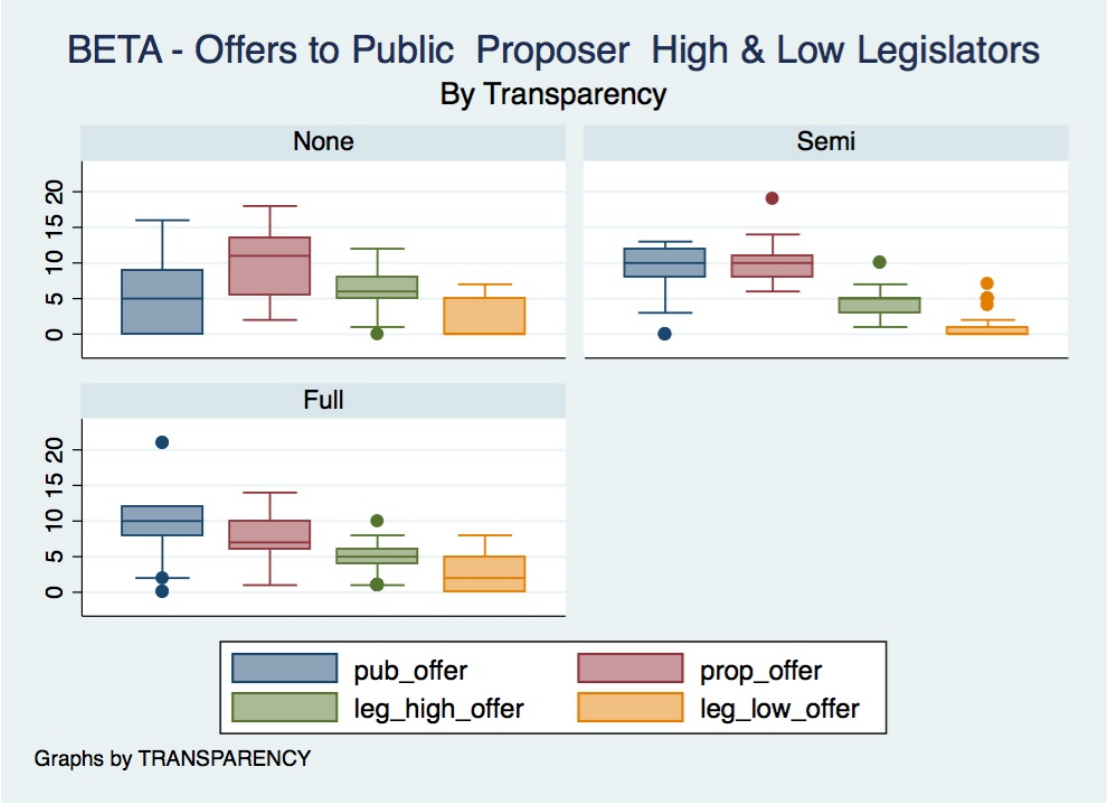


Figure 17. Offers to Public, Proposer, and High and Low Legislators, by Transparency



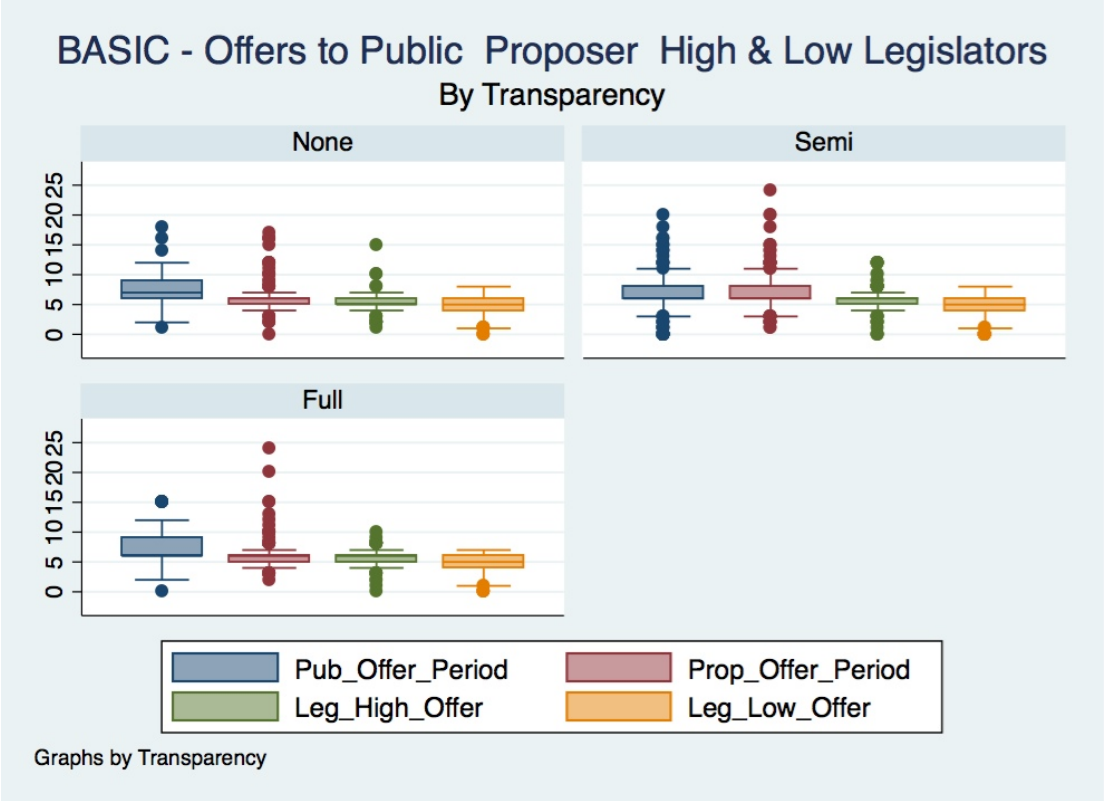


Figure 18. Linear estimation of offers to each type of player by period and transparency.

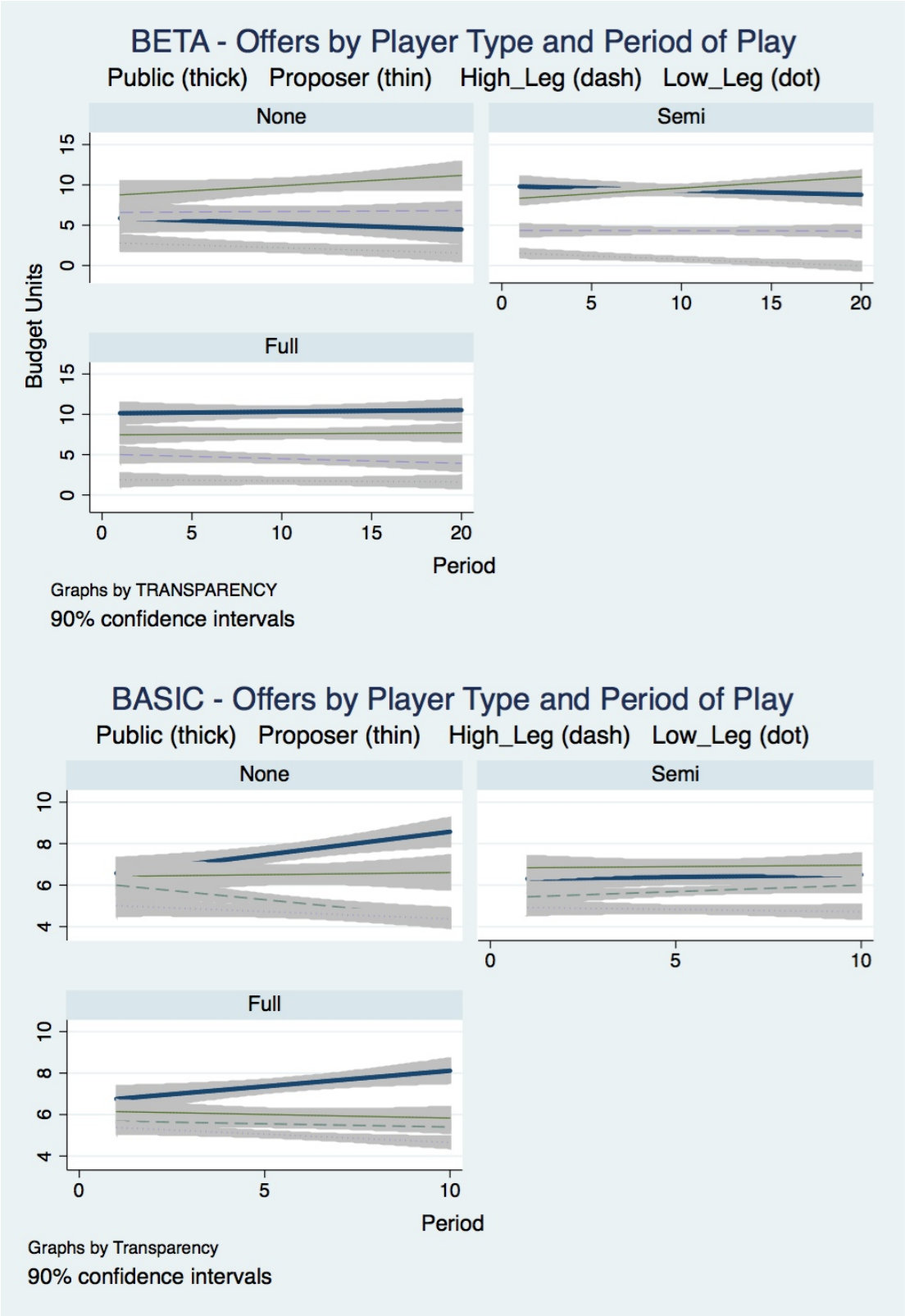


Figure 19. BASIC - First-order accountability by transparency

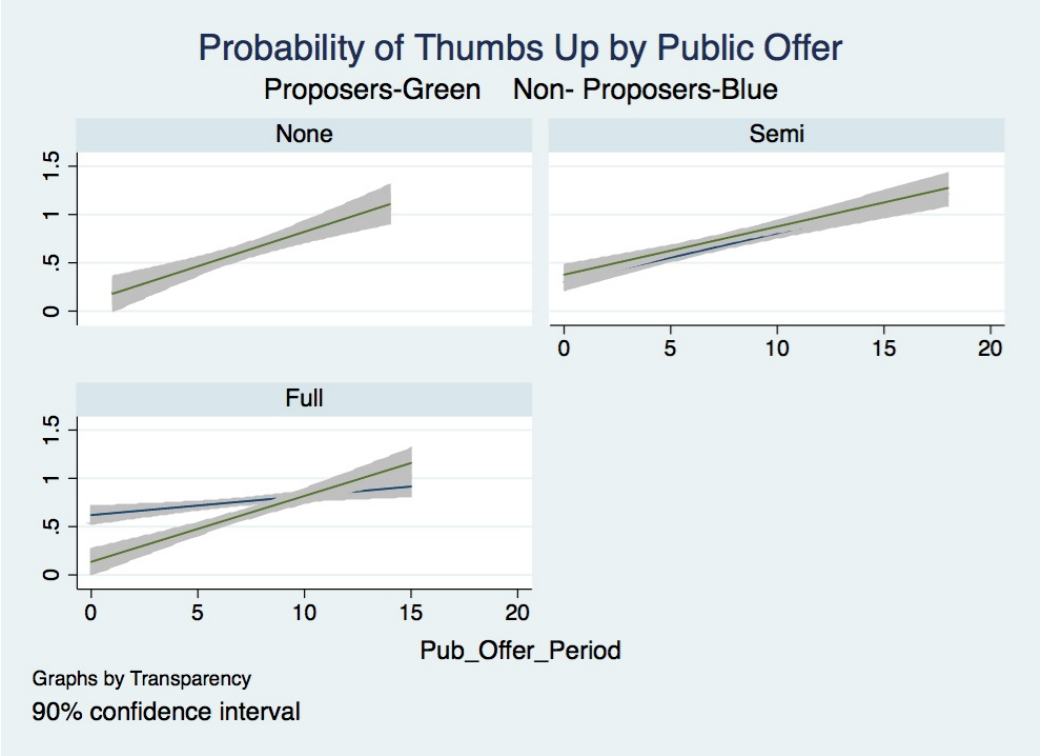


Figure 20. BASIC – Second-order accountability for Non-Proposers under Full Transparency

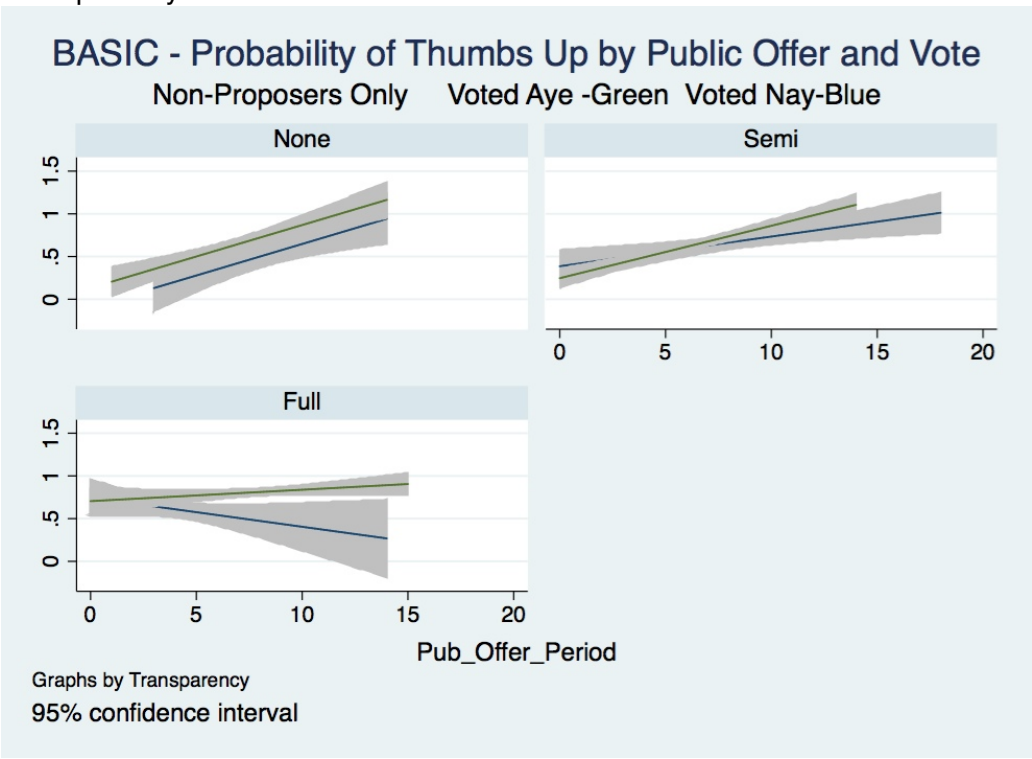


Table 1. Summary of Logistics of BETA and BASIC Experiments.

Version	Platform	Date	Subjects	N Start	N Complete	Unit Value
BETA	Lab	Aug.'10	Dartmouth students	90	90	$.50
BASIC	Web	Feb.'13	MTurk workers	119	95	$.05
BASIC	Web	Feb.'13	MTurk workers	106	87	$.10

Table 2. Mean offers by Live Proposers to each type of player, by Transparency

	Transparency			
	None	Semi	Full	
Public	5.2	9.3	10.3	BETA
Proposer	10.0	9.7	7.5	
High_Legislator	6.7	4.3	4.5	
Low_Legislator	2.2	0.7	1.8	
Public	7.6	6.4	7.5	BASIC
Proposer	6.5	6.9	6.0	
High_Legislator	5.2	5.7	5.5	
Low_Legislator	4.7	4.8	5.0	

Table 3. Logistic regressions of budget approval votes by non-Proposer Legislators on their own budget offers, and Public_Offers, by transparency conditions. (N=114)

	Non-Transparency		**Semi-Transparency**		**Full Transparency**		
	Coefficient (Std.Error)	20th→80th	Coefficient (Std.Error)	20th→80th	Coefficient (Std.Error)	20th→80th	
Legislator's Budget Offer	.72 (.13)	+72% (9%)	1.65 (.40)	+69% (7%)	.41 (.14)	+19% (6%)	**BETA**
Public's Offer	.30 (.09)	+39% (13%)	-.16 (.14)	-2% (3%)	.38 (.14)	+12% (5%)	
Legislator's Budget Offer	.66 (.11)	+52% (8%)	.56 (.06)	+47% (5%)	.35 (.06)	+19% (03%)	**BASIC**
Public's Offer	.13 (.07)	+10% (5%)	.17 (.04)	+14% (5%)	.10 (.04)	+10% (4%)	

Table 4. Percentage of Minimum Coalitions

	BETA	BASIC
Non-Transparency	53	3
Semi-Transparency	67	8
Full Transparency	26	2