How do Public Goods Providers Play Public Goods Games?*

Daniel M. Butler, Yale University
and
Thad Kousser, UC San Diego

Abstract:

We study how policymakers play public goods games, and how their behavior compares to the typical subjects we study, by conducting parallel laboratory experiments on college undergraduates and American state legislators. We find that the legislators play public goods games more cooperatively and more consistently than undergraduates. Legislators are also less responsive to treatments that involve social elements, but are more likely to use information that optimizes their interests. We also find that much of the variation in how legislators play is explained by legislators’ fixed characteristics. Methodologically, our results can facilitate further lab experimentation on the effect of political institutions by explaining when and how we can use student samples to make inferences about elite behavior. Substantively, our results yield insights about how various institutions can affect levels of public goods provisions.

* Earlier versions of this paper were presented at the 2013 Annual State Politics & Policy Conference and the 2013 Annual Meeting of the American Political Science Association. We thank conference participants for thoughtful comments and feedback.
An extensive and fruitful literature has used laboratory experiments to learn about levels of public goods provisions. These studies have produced important theoretical insights and practical solutions to significant problems (e.g., Ostrom, Walker, and Gardner 1992, 1994; Dietz, Ostrom, and Stern 2003; Milinski, Semmann, Krambeck, and Marotzke 2006; Buchan, Grimalda, Wilson, Brewer, Fatas, and Foddy 2009; Hamman, Weber, and Woon 2011). Public goods experiments have yielded powerful insights because these researchers have been able to directly manipulate the institutions in the study and thus assess the causal impact of institutional variations observed in the real world.

In practice, researchers typically recruit undergraduates to participate in these lab experiments; even researchers who want to make inferences about the behavior of elites typically draw participants from undergraduate subject pools. This is not necessarily an unreasonable approach. It is more convenient to recruit students and they provide a baseline for how intelligent and educated but otherwise ordinary human beings will act.

However, these students only provide a baseline. In practice we often want to make inferences, based on lab experiments, about how elites would act. Our review of recent lab experiments in the next section shows that many researchers, especially those studying how institutions affect individuals’ interactions, explicitly draw conclusions about elite behavior from their work. This is not surprising – the ability to vary institutions is a great advantage of lab experimentation and scholars are interested in learning how institutions affect elites’ behavior. Even when authors do not directly draw these conclusions, readers are interested in thinking about such conclusions. For example, are the studies showing how deliberation

---

1 Of course this is not a phenomenon unique to lab experiments. Few experiments directly study the aspects of elite behavior (prominent exceptions include Grose and Russell 2008; Bergan 2009; Butler and Broockman 2011; Butler and Nickerson 2011; Harden 2013; Enemark and Zimmerman 2013; Hafner-Burton, Leveck, Fowler, and Victor 2013). One reason we directly study the behavior of elites less often is the high cost to doing so, whether by experiments or surveys (e.g., Carey, Niemi, and Powell 2000; Maestas 2003; Burden 2004; Maestas, Fulton, Maisel, and Stone 2006; Layman, Carsey, Green, Herrera, and Cooperman 2010).
affects levels of cooperation among students (Ostrom, Walker, and Gardner 1992, 1994) likely to equally apply to elites?

In order to answer questions like this, we must learn whether and how elites’ behavior differs from those of standard subject pools. Are they more or less selfish? Are they more or less strategic? Are they more or less responsive to institutional changes that affect interactions with others? By answering such questions we can learn when and how to draw inferences about elites based on the play of undergraduates and thereby make the most of lab experiments.

We provide new insights into these questions by presenting the results of two parallel public goods experiments: one conducted with 37 elected officials from twelve American state legislatures, and the other conducted with 31 university undergraduates. The rules, experimental manipulations, and incentives for the players were identical in both games. This allows us to compare and contrast how these lawmakers and college students play public goods games and how they respond to rule changes in the game.

The experiment for both groups is modeled after conventional public goods games (see Chaudhuri 2011). In the basic version of the game, players first decide how much to withdraw from a group resource for themselves. The remaining amount is then doubled and shared equally among all players in the group. Like all public goods games, and the prisoner’s dilemma, group welfare is maximized when everyone cooperates by not withdrawing anything from the group fund. This is difficult to achieve, however, because everyone’s best response is to deviate and take as much as they can for themselves.

For our experiments, we had the participants play different versions of the game. These different versions of the game are analogous to institutional design rules that have been implemented or proposed in actual state legislatures. Specifically, in different versions
of the game we vary (1) whether participants are given a chance to deliberate about what should be done, (2) whether participants will play only once together or if they know they will have the chance to play repeatedly with the same players, and (3) whether all participants share the same information or if some are only given part of the information.

How do the elected officials and college students play this game, and how do they react to the experimental treatments? First, we find that the elected officials play the game more cooperatively than the college students. While this is the general pattern, we do see some instances of legislators taking the maximum amount possible, a behavior that we never observe among students.

Second, the legislators and students respond to different types of treatments. Deliberation, for example, causes students to act more cooperatively but does not significantly affect legislators’ behavior. Our results from the student population confirm the seminal findings from Ostrom, Walker, and Gardner (1992, 1994): deliberation can facilitate cooperation among students. However, and to the very point of this paper, deliberation does not cause the policymakers in our study to change their behavior.

On the other hand, when we provided some players with exclusive information about how the payoffs would change, the legislators responded – pressing their informational advantage – but the students did not. Sample selection strongly influences the results we find. As a discipline, the more our samples reflect the population we want to learn about, the more informative our studies will be. The best way to learn about elites is to study elites.

Third, we find that much of the variation in how legislators play is explained by the characteristics of the legislators themselves. Individual fixed effects explain fifty-three percent of the variation in legislators’ contribution levels to the public goods, but only
twenty-eight percent of the variation in students’ contribution levels. Exploring this individual variation further, we find that members of the minority party, Democratic legislators, and those who are married contribute more to the public good in the game. Overall, we find that elected officials are more pro-social, are less reactive to changing institutional structures, and are guided by their personal traits and life experiences.

I. A Review of Recent Published Lab Experiments

Before describing our experimental design in greater depth, we provide a broad overview of how lab experiments are used in political science and lay out expectations about how our student and elite populations are likely to behave differently. Box 1 provides an overview of the lab experiments published in the American Political Science Review, the American Journal of Political Science, and the Journal of Politics during the period 2002-2012. Our review is based on the fifty-one articles using lab experiments that were published in these outlets during this period. Twenty-nine of those articles (about 56 percent of the sample) were lab experiments that did not involve any interaction between participants. The vast majority of these twenty-nine articles reported on experiments that exposed the lab participants to either a putative newspaper article or television segment. The participants read or watched the relevant media, part of which had been experimentally manipulated, and then answered survey questions. The majority of these studies were aimed at understanding the behavior of the mass public. Only ten percent of these articles (three out of twenty-nine) tried to extrapolate to the behavior of elites.

By contrast, the researchers conducting experiments that involved interactions between participants were more likely to extrapolate from their results to make claims about the behavior of public officials. Half (eleven out of twenty-two) of the articles that involved
participants interacting with each other in some way used the results of their study to make causal claims about the behavior of elites. Often these claims involved the likely influence of institutional change on politicians’ behavior.

While many researchers use their experiments to make inferences about the behavior of elites, we were actually surprised that more researchers were not doing research in this vein. One of the great advantages of lab experiments, in particular, is that it allows researchers to directly manipulate the institutional context. In the real world we typically cannot manipulate these institutions, making it difficult to know where the variation in the institution comes from. Researchers who are able to manipulate the institutions in the lab setting so that they mimic the institutional variation observed in the real world have the potential to learn about how institutions affect public officials’ behavior.

II. Why Legislators and College Students Might Distribute Public Goods Differently

Perhaps one reason that more researchers are not using lab experiments to study how institutions affect the behavior of elite actors is that there are questions whether the behavior of students can be used to make inferences about the behavior of elite actors. It is unclear whether undergraduates’ behavior can inform us about the behavior of elite policymakers (McDermott 2002, 2013). Because we do not know how students and elites differ in how they act in the lab, we do not know how we should extrapolate results from lab experiments in making inferences about policymakers.

Theoretically, there are a few reasons that students may behave differently than state legislators. First, legislators are different because they have office-holding experience. Evidence suggests that holding office causes individuals to play a “Trust” game, which can
lead them to be more cooperative (Enemark, Gibson, McCubbins, and Zimmerman 2013). In our experiments we test this prediction by comparing the differences between the legislators and students in how much they contribute to the public goods. Our prediction is that students will take more for themselves.

Holding office also provides legislators with more experience in pressing their strategic advantage. As part of our experiment we test how the students and legislators use information about the payoffs in the game (information that is randomized) to maximize their interest. We predict that public officials will change their play more in response to this information (so as to maximize their interest) than will students.

Second, students and legislators are at different stages in their life cycle. While undergraduates are typically around 20 years of age, the average age of the legislators in our sample was 53. The youngest legislator in our sample was still 30 years old, a decade older than most undergraduates. Younger people are systematically more compliant (Sears 1986). They also have less life experience and individuals with less experience are less stable in their preferences and more susceptible to manipulation (List 2003; Haigh and List 2005; Alevy, Haigh, and List 2007). Because of these dynamics, we expect student samples to change their behavior more frequently across the rounds of the game. We test this prediction by comparing the difference, for students versus legislators, in the variance of individual-level play across the rounds.

Third, peers have a bigger influence on young people (Sears 1986). Consequently we expect students to be more influenced by mechanisms that increase the importance of peer influence on individual play. In our experiments we test whether deliberation affects

---
2 Enemark, Gibson, McCubbins, and Zimmerman (2013, p. 30) are careful not to give a normative frame to their empirical finding that becoming an elected official leads to greater reciprocity in game play, but do conclude that it lends support to those who foresee an “ennobling effect of office.”
individuals’ willingness to contribute to public goods. We predict that deliberation will have a bigger impact on students than it does on legislators.

As part of our study, we also test the impact of changing expectations about playing with each other repeatedly. We expect that the prospect of playing with the same people repeatedly may have a bigger impact on the behavior of students because of the greater concern they have for peer evaluations.

III. Our Public Goods Game: Participants, Structure, and Manipulations

A. Two Participant Samples

We test our predictions by running parallel experiments on a group of state legislators and a group of undergraduate students. We gained access to the state legislators in our sample by working with the Council of State Governments to plan a session for their “Western Legislative Academy.” Founded in 1933, the Council is an organization that works with all three branches of government across the American states to foster “the exchange of insights and ideas to help state officials shape public policy.”

Its Western Legislative Academy is a three-and-a-half day training program that brings together three legislators from each of 13 western states. While all of the legislators attending the Academy are in their first four years of service, states generally send lawmakers who are tabbed for future leadership positions. The Academy has trained 488 legislators since 2000, with 69% of those who still serve in state legislature doing so in caucus or committee leadership positions.

---

Additionally, five alumni now serve in the House of Representatives and two in the US Senate.⁴

Beginning more than a year before the November, 2012 Western Legislative Academy, we worked with the Council of State Governments to develop a public goods game for legislators at the academy. We worked with the Council’s staff to adapt traditional public goods game to the theme of the conference: building cooperation in their state legislatures. In line with the theme of the conference, we drew on the theoretical literature in economics and political science to design experimental manipulations of the rules in the game that are analogous to institutional reforms that have been proposed in the states. In a project aimed at exploring the external validity of abstract laboratory experiments, we view the collaboration through which the Council pushed us to maximize the game’s realism as a major benefit. In the Academy’s program, we summarized the game as follows:

“The Budget Game: Building Trust and Cooperation inside 21st Century Legislatures

In this session legislators will play a fast-moving, interactive budget game designed to highlight the difficulty of building cooperation inside a legislature. The game will challenge you to identify ways to gain each other’s trust. After the game is over, you will be debriefed and learn practical applications for real legislatures as you strive to build trust and cooperation in your statehouses.”

We conducted this game with the 37 attendees at the 2012 Western Legislative Academy, held in Colorado Springs, Colorado from November 13-16. All attendees were elected legislators from one of twelve western states.⁵ Our sample included 20 Republicans and 17 Democrats, with 24 being members of the majority party⁶ in their house and 13 members of the minority party. Because the Academy draws legislators from a broad range of states, our sample includes Republicans who were in the majority in their houses and

---

⁴ These figures come from personal communication with Mary Lou Cooper of the Council of State Governments on March 12, 2013.

⁵ The represented states are Alaska, Arizona, Colorado, Hawaii, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming. No lawmakers from California attended due to last-minute scheduling conflicts.

⁶ This includes two members of Oregon’s House, which had a partisan tie at the time of the Academy.
Republicans who were in the minority, along with Democrats of both types. Thirty of the legislators were men and seven women. Eight of the legislators served in houses where term limits were in effect, and our participants served on an average of 3.83 committees.

We worked with the [name of lab - redacted] at [one of the researcher’s university] to conduct a parallel experiment with participants from their subject pool in a university laboratory setting. In April, 2013, after receiving approval to administer the experiment from the Human Subjects Committee, we conducted the game with 31 participants who were told they would be paid $10 show-up payment for participating in “The Budget Game”.

B. Basic Structure of the Game

Both in [the researcher’s university town] and in Colorado Springs, participants began the game by sitting down at tables with four or five participants and receiving a packet of instructions. In both cases, we described the game’s basic rules and showed how payoffs would be calculated. We then explained key elements of a written consent form and gave participants time to read and sign it. In both settings, participants played for a public university of their choosing. We informed players that every 100 points they earned in the game would translate into a one dollar for the general scholarship fund of their chosen university, with an expected total contribution, including all of the participants in the group, of about $1000. One advantage of using these incentives as opposed to paying participants directly is that it better matches what occurs in the real world context. Public officials are not acting to secure a personal financial payoff, but instead seeking benefits for their respective districts. Although these rewards are not as large as those given in the real world.

---

7 We also considered paying participants based on their winnings in the game, but this does not reflect the environment in which public officials work. Further, even had we wanted to directly pay participants, we could not do so per Council of State Governments policy.
political context, they injected a level of competitiveness into the room in each setting. Significantly, holding the incentives constant across the two experiments diminishes the possibility that differential incentives explain any differences between our legislators’ and undergraduates’ behavior.

The basic version of the game that each set of participants played was modeled on classic public goods games (Chaudhuri 2011) designed to observe how people will behave when they have an individual incentive to take actions that harm a group goal. The instructions for the basic version of the game read as follows:

In each game, each table will start with a “table fund” that has a value equal to the number of players multiplied by 100. Thus if there are 5 players, the table fund will be equal to 500 and if there are 4 players the table fund will be equal to 400. Each of you will be asked to choose an amount between 0 and 100 that you will take from the table fund. You will keep all the money you take from the table fund. We will then double the amount left in the table fund for your table and divide the resulting total equally among all players at your table.

Example with five players:

<table>
<thead>
<tr>
<th>Player</th>
<th>Amount Taken from Table Fund</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player 1</td>
<td>100</td>
</tr>
<tr>
<td>Player 2</td>
<td>50</td>
</tr>
<tr>
<td>Player 3</td>
<td>80</td>
</tr>
<tr>
<td>Player 4</td>
<td>20</td>
</tr>
<tr>
<td>Player 5</td>
<td>0</td>
</tr>
</tbody>
</table>

Total in Table fund initially: 500
Total in Table fund after players make choices = 250
Amount distributed from Table fund = 500 (i.e., 250*2)

Final Payout For Each Player

<table>
<thead>
<tr>
<th>Player</th>
<th>Amount Directly Taken from Fund</th>
<th>Amount from Table Fund</th>
<th>Total Amount</th>
</tr>
</thead>
<tbody>
<tr>
<td>Player 1</td>
<td>0</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Player 2</td>
<td>50</td>
<td>100</td>
<td>150</td>
</tr>
<tr>
<td>Player 3</td>
<td>20</td>
<td>100</td>
<td>120</td>
</tr>
<tr>
<td>Player 4</td>
<td>80</td>
<td>100</td>
<td>180</td>
</tr>
<tr>
<td>Player 5</td>
<td>100</td>
<td>100</td>
<td>200</td>
</tr>
</tbody>
</table>
After participants read the example of how players could behave and what payoffs they would earn – and thus seeing that the player who took the most from the table fund earned the largest final payoff – they played one practice round. After practicing, each participant followed the instructions on her information packet to go to a randomly assigned table. Players then began the first of the six periods, with each period bringing a different version of the game. After finishing each version of the games, players were instructed to move to another randomly assigned table. Before the first period began, we told them that they would be moving to a new table with a different set of players after finishing each game.

These six periods were composed of three sets of paired games, each of which gives us the opportunity to compare how players behaved in a “control” and a “treatment” version of the game that manipulated one key characteristic (described in the next section). We alternated the order of play within each pair across participants. For instance, half of the participants played the “basic” (control) version of the game in the first period and then the “deliberation” (treatment) version in the second period, at a different table with other players who were on the same schedule. The other half of the participants played the “deliberation” version first, and then the “basic” version in the second period at another table. Participants were also randomized into different schedules during periods 1, 3, and 5. In other words, treatment assignment period 1 did not affect treatment assignment order in periods 3-6.

We employed this within-subject design so that we could use participants as their own counterfactuals. Our randomization into two different schedules ensures that treatment, within the paired treatments that we compare, is independent of the order of play and player history.
C. Experimental Manipulations

Our first pair of games was designed to test the impact of a stylized form of deliberation by prompting players to reveal and discuss how much they planned to take from the collective pot before making a final decision. None of the payoffs of the game changed; all of the observed differences in how much players took for themselves resulted from the revelation, the discussion, and the social dynamics that emerged over the course of four minutes. In order to isolate the impact of this form of deliberation, we had participants engage in both the “basic” and the “deliberation” versions of the game with different sets of partners during the first two periods of play. (Again, half of our sample played the control version of the game first, while the other half played the treatment version first).

In the basic version of the game, participants simultaneously wrote down the amount that they wished to withdraw from the table fund on a sheet of paper that each had in their envelope, revealed it to the group, and then moved to another table (that was randomly assigned for each individual). In the deliberation version, the instructions explained that they would begin by writing out and revealing how much they would like to take out of the table fund. “We will then give you 4 minutes to deliberate about what happened and what you should do in the second round. We will then play the second round.” Notably, nothing would be different about the game or the division of the points after the deliberations. Players would reveal how much they wished to withdraw, knowing that what was left in the table fund would be doubled and evenly distributed before everyone moved on to a different table. This intervention is designed to parallel current proposals that have been made in
California requiring that every bill be in print for three full days before a vote on final passage, in order to bring greater sunshine and deliberation to the legislative process.\(^8\)

We study the role of communication because previous studies have shown that such deliberation has a huge effect on contributions to public goods (Bornstein and Rapoport 1988; Orbell, van de Kragt, and Dawes 1988; Dawes, van de Kragt, and Orbell 1988; Ostrom, Walker, and Gardner 1992, 1994). Ostrom and her colleagues (1992, 1994) show that allowing communication has a bigger effect on donations to the public goods than giving players the power to sanction non-cooperative behavior. They find that giving college students ten minutes to discuss a game allows them to play it much more cooperatively, even when the game’s structure and payoffs do not change. This is consistent with Cohen’s (2009) argument that deliberation pushes people away from arguments based on their individual interests and instead “to find reasons that are persuasive to all” and “focused on the common good.” Such deliberation can also change participants’ intrinsic payoffs (Ackerman and Fishkin 2003; Reich 1985) and increases the social pressure that participants feel to act cooperatively. As outlined above, our expectation is that students will respond more to this pressure and thus be more affected by the deliberation treatment.

Our next pair of game periods was designed to test the impact of short-term versus long-term relationships on cooperation. These manipulations were designed to be analogous to the different time horizons that state lawmakers face in term-limited legislatures versus those that place no cap on service (see Carey et al. 2006 and Kousser 2005), or the way that behavior changes when legislators (Jenkins and Nokken 2008) or governors (Kousser and

---

\(^8\) This proposal was part of the unsuccessful Proposition 31 on California’s November, 2012 ballot, and has been revived in bills (Assembly Constitutional Amendment 4 and Senate Constitutional Amendment 10) proposed during the 2013-14 legislative session. Of course, the requirement that bills would be in print for three days would do more than simply force legislators into deeper deliberations over bills; it would also allow interest groups time to study new language and mobilize in support and opposition to the bill.
Phillips 2012, Ch. 3) are lame ducks. In our application, the “short-term” version of the game is structured identically to our basic game, with the added reminder in the instructions that, “We will play just one time before switching to a new table and a new group.” In the “long-term” version, the instructions informed participants that, “You will play the game with the same group for a potentially long number of rounds. After each round, we will flip a coin. If the coin comes up heads your group will play another round. If the coin comes up tails, we will stop and change groups.”

We compare how participants played in the short-term version with how they played in the first round of the long-term version to estimate the impact of this lengthened time horizon. We only use the first round because the tables ended up playing the long-term version for different numbers of rounds, and saw very different patterns of play based on the coin toss and the evolving dynamics between players randomly assigned to sit together. Looking only at the first round of play isolates the impact of expected time horizons from these idiosyncratic factors.

Research dating back to Axelrod (1984) has shown that players are often able to incentivize and enforce cooperation when they play a game repeatedly, but that cooperation breaks down when they know they will not interact again. With repetition, “Legislators do not wholly succumb to granting every constituency a benefit so as to wholly bankrupt the state… people do contribute to charities even though they might each be better off by defecting to a pattern of non-contribution; and subjects in the laboratory when playing the Prisoners’ Dilemma frequently chose to cooperate. (Ordeshook 1972, p. 176)” The reason is that, in repeated play, those who defect from group cooperation early on can be sanctioned by others. By contrast, when players cooperate in the initial round of a repeated game, they develop a good reputation that can engender future cooperation from others and thus higher
future payoffs for all. Anticipating this, players should cooperate more, by withdrawing less from the table fund, in the first round of our long-term game than in the short-term game. Because peer influence has an even stronger influence on the behavior of students, we expect that they will be more responsive to having a longer time horizon.

The last pair of games explored the effects of asymmetric information by changing the payoffs in a meaningful way and informing all players about it in the control version of the game, but informing only some of them in the treatment version. The new information was whether the table fund would triple in value (our “high multiplier”) before being redistributed or whether it would grow only by a factor of 1.5 (our “low multiplier”). In the “all informed” version of the game, every player at the table received instructions about the multiplier, and was told that this was common knowledge. Before the “asymmetric information” version, three of the five players at the table received a special envelope giving them exclusive information about the new multiplier. We gave out these envelopes visibly, so that the remaining two players recognized that they were the only ones lacking this critical piece of information.

We designed this experimental manipulation to parallel the information (a)symmetries that often develop between members of the majority and the minority parties in legislatures. The majority party sometimes seeks to monopolize information about the impact of policies or its plans for the lawmaking process. In American states, the majority may be especially empowered to do so when legislative staff members are hired through partisan, caucus-based staff organizations rather than through a bipartisan model. According to a 2003 survey by the National Conference of State Legislatures, states vary greatly in how they staff their legislatures. More than half of legislative staffers are partisan in 22 states, while 14 states employ fewer than one in ten of their staffers this way (Kurtz 2006).
Our expectation is that when a majority of players at a table hold information that a minority lacks, cooperation will diminish. Players who lack the information may feel resentful of those who became fully informed, with the asymmetry of information creating rivalry. Members of the information minority are thus likely to withdraw more from the table fund, both because of their uncertainty about the multiplier and because their exclusion leads them to feel less invested in the common pool. Members of the informational majority may anticipate this collapse of cooperation, which in turn could cause them to withdraw more from the table fund than they did when everyone was informed (holding constant the type of multiplier that the table was dealt). In the asymmetric information case, players in the majority group should also be less responsive to the magnitude of the new multiplier, knowing that only some of the other players will be informed and thus act on it. Further, state legislators, because of their experience and sophistication, should be more likely to press their informational advantages and thus more responsive to the informational treatments.

IV. Results

A. Baseline Contribution Patterns

Our first prediction was that legislators should play more cooperatively because holding office leads public officials to become more public goods oriented (following Enemark, Gibson, McCubbins, and Zimmerman 2013). Figure 1 compares the distribution of our two samples in terms of the average amount that each participant took. As predicted, the legislators were the more cooperative group, with the mean legislator withdrawing an average of 16.4 points from the table fund over the six periods of play. As the distribution shows, many of the legislators consistently left nearly all of the points in the common pool,
though a handful of them always took nearly the full 100 points for themselves. On average, however, the undergraduates took more for themselves over the course of the session, with an average withdrawal of 25.2 points (the difference between the two group means is statistically significant at the 90% confidence level). From one perspective, this is a comforting finding; the lawmakers who exert authority over public goods left more in the common pool than the undergraduates participating in the study. Yet it also highlights that the students and policymakers in our sample played the public goods game quite differently.

We also expected students to change their behavior more over the rounds of play, consistent Sears’ (1986, p. 515) findings that college students have “less-crystalized attitudes, less-formulated senses of self” than older and more experienced individuals. Figure 2, which reports the distribution of standard deviations of individuals’ contribution levels, confirms this prediction. With a mean standard deviation of 16.6 points, legislators played the game more consistently than students (with a mean standard deviation of 24.7) did. This difference is significant at the 95% confidence level.

B. Responses to Experimental Treatments

We now examine the effects of our experimental treatments on legislators and students. Figures 3 and 4 reports the average amount that participants withdrew from the table fund during our three sets of paired-games. In Table 1, we estimate the treatment effects – with the basic version of the game serving as the comparison category – using multivariate models that include individual fixed effects, with robust standard errors clustered by individuals.\(^9\)

\(^9\) We use robust standard errors because a Breusch-Pagan/Cook Weisberg test revealed heteroskedasticity across rounds (as players learned to play the game the amount of variance changed).
One of our predictions was that deliberation would have a greater effect on students, drawing on Sears’ (1986) reminder that peers have a greater influence on younger people. In this regard, Figures 3 and 4 show that both students and legislators took less for their district when they deliberated about what they should do. Yet the effect of deliberation was much stronger among the students. After deliberating, legislators withdrew an average of 12 points from the table fund, compared with the 16 points that they withdrew when playing the basic version of the game (a difference that is not statistically significant). By contrast, the students took only 9 points from the common pool, compared with the 24 points that they withdrew in the basic game. As Table 1 shows, this difference is significant at the 95% confidence level. Further, the difference-in-differences is statistically significant. In a separate model (not shown here), we estimated a model that was the same as the model reported in the first column of Table 1, except that it added an interaction between the deliberation treatment and a variable indicating that the player was a legislator. This interaction (which had a coefficient of 14.5 and a standard error of 6.1) was significant at the 95% confidence level. Deliberation has a much stronger effect on students’ behavior than it does on legislators’ behavior.

These results both confirm earlier findings about the role of deliberation on levels of public goods provision (e.g., Ostrom, Walker and Gardner 1992, 1994) and challenge the application of these results to the behavior of policymakers. Deliberation has an effect on students (Ostrom, Walker and Gardner 1992, 1994), but not policymakers.

In our second set of paired-versions, we tested whether participants would respond to variations in the expected time-horizon of the game. As noted above, we tested this argument by having the participants play both a single-shot version of the game and one which they knew play could be repeated an indefinite number of times at their table (with a
coin flip determining if play would continue after each round). Legislators withdrew slightly more from the table fund in the first round of the repeated game (acting contrary to our expectations), while student withdrew slightly less. In neither case, though, did these differences approach statistical significance.

Perhaps the impact of the player’s time horizon may have been dampened by the somewhat artificial nature of our laboratory setting. Players may have believed that, in some sense, they would never be playing simply a one-shot game. They might have expected to continue to interact with the others seated around their table formally in later rounds, if they happened to be randomly assigned once again to the same tables. Or they might have anticipated future informal interactions, either on campus for the undergraduates or during the remaining day and a half of the Western Legislative Academy for the lawmakers. Any of these reasons could have made them wary of withdrawing too much from the common pool during the single-shot version that we intended as an analog for short-term interactions.

Finally, we examine the impact of asymmetric information about the multiplier applied to points left in the table fund, which we expected to bring less cooperation. Figures 3 and 4 show that legislators indeed withdrew more when only some of the players seated at each table were told about the new multiplier, while students did not change their behavior. This manipulation, though, deserves a more detailed exploration because we actually assigned participants, over these two versions, into six possible conditions.

Table 2 reports how players behaved in five of those conditions, compared to how they behaved when it was common knowledge that the funds left in the general fund would be multiplied by 3 (the condition that should minimize withdrawals). When information was common knowledge, but all players were told that the multiplier would shrink to 1.5, both groups responded, as expected, by withdrawing more points from the table pot. While this
effect falls just short of significance in each sample by itself, it does reach significance at the 95% confidence level when all players are considered together.

Students and legislators reacted differently, however, to the asymmetric information version of the game. When the multiplier shrunk to 1.5, but only some of the legislators seated around the table were given the benefit of this information, the informationally advantaged lawmakers withdrew 39.4 more points from the table pot. This was the largest of the treatment effects that we observed, and is significant at the 99% confidence level. Students in the same strategic situation responded by changing their behavior in the same direction, but the effect was not nearly as strong and fell well short of significance. Consistent with our prediction, the legislators were more likely to make the most of their informational advantages. None of the other experimental conditions brought a significant impact, though all moved in the directions that we anticipated.

C. Explaining Variation in Contributions Across Individuals

If, as we have argued, legislators and students act differently in our public goods game because legislators have a stronger career and personal history, this limits our ability to learn about the effect of elite-level, governing institutions from experiments conducted on students. One initial piece of evidence that legislators’ actions are guided more strongly by these personal characteristics than students’ actions comes from a comparison of the explanatory power of individual fixed effects in the models that we present in Table 2. All of the models reported in that table use individual fixed effects. In an unreported analysis, we estimate the same models without individual fixed effects. For legislators, the new model explains only one percent of the variation in the number of points taken, down from fifty-three percent in the model with fixed effects. For students, by contrast, the explanatory
power of the model declines only from 38 percent to seven percent when we remove the fixed effects. The individual considerations that legislators bring to public goods decisions are much more important than the characteristics and quirks of students.

We collected individual-level information about the legislators in our sample to explore the impact of these characteristics in greater depth. We conducted this purely observational study by drawing on legislator biographies. From votesmart.org and other sources, we were able to code each legislator’s marital status, gender, partisanship, and age. In addition to these variables we coded whether each legislator was in a term-limited state and whether they were part of their legislature’s minority party. Finally, because the parties in the Oregon House shared power in the last session, we include a separate dummy variable for legislators from this chamber to account for their unique power-sharing arrangement. The suggestion to look at minority party status came from one of the legislators who participated in our study during the debriefing process. The legislator thought that minority party legislators would take more from the common pool.

We estimate the predictive effect of these factors on how much individual legislators take from the common pool each period by estimating a regression that uses the information from all of the games played. We include legislators’ behavior from the entire session because each play of the game provides information about the legislators’ behavior. To control for the fact that we are using the same individuals’ repeatedly, we cluster the standard errors at the individual level in the regression.

Table 3 reports the regression results. As predicted during our debriefing session, minority party lawmakers withdrew an average of 12.3 more points from the table fund each time, a difference that is significant at the 90% confidence level. How should we interpret this effect? It could be that the type of politician drawn to serve in a house where she might
reasonably expect to be in the minority has a different orientation to politics. Who chooses to run for office when they expect to be in the minority? Legislators who want to work on big bills that will influence policy may find service under these conditions less rewarding. By contrast, district-oriented legislators should still be willing to enter the arena during these times (they can still fight for the district). Alternatively, the very experience of serving a session in the minority could cause a lawmaker to feel that they had fewer common interests with other legislators, and thus make them less likely to contribute to the common pool. Either way, the results are interesting and suggest that institutions can shape political outcomes by shaping elites’ behavior (either directly or by means of incentivizing certain types of politicians to seek office).

Table 3 also shows that Republican legislators took an average of 19.6 more points from the table fund. Why might this be? Members of the GOP might systematically come from professional backgrounds, such as business where they focus more on maximizing profit margins. Alternatively it may be something about the current approaches of the two parties that causes Republican legislators to focus more on strategically maximizing their own district’s interests.

By far the strongest effect of the personal characteristics that we observe is that married legislators played the game much more cooperatively than single or divorced legislators. Married legislators took 32 fewer points per round, a difference that is significant at the 99% confidence level. Again this result could reflect a selection process (perhaps people who marry and stay married are more others-oriented), a development process (perhaps being married causes people to become more others-oriented), or something else entirely. Regardless of the explanation, it highlights an important limit of public goods
studies conducted only on college undergraduates. Students are systematically different than legislators in ways that predict how much they cooperate with others.

V. Conclusions

Laboratory experiments, which allow researchers to mimic and manipulate the institutions that actors operate under in the real world, hold great potential for learning about optimal institutional design. In this vein, a vast and fruitful literature in the social sciences uses lab experiments involving university undergraduates to study determinants of individuals' public goods contributions (Ostrom, Walker, and Gardner 1992, 1994; Chaudhuri 2011). Our ability to draw conclusions from these experiments to understand how institutions are likely to affect the behavior of policymakers depends on knowing how students and policymakers act differently in strategic settings.

In this paper we have tested how students and policymakers act by conducting identical public goods experiments on American state legislators and on college undergraduates. Though the rules, manipulations, and incentives were the same for both groups, the two groups played the game very differently. Legislators were systematically more generous and more consistent when they played and were also more responsive to the information they received. Finally, deliberation had almost no effect on legislators’ behavior but had a large effect on how students played the game.

One of the important findings from the literature on common pool resources is that deliberation can be highly effective, sometimes more effective than other policy levers, at increasing donations to public goods (Bornstein and Rapoport 1988; Orbell, van de Kragt, and Dawes 1988; Dawes, van de Kragt, and Orbell 1988; Ostrom, Walker, and Gardner 1992, 1994). These conclusions regarding the effectiveness of deliberation in the lab are
based on experiments conducted with undergraduate students. We replicate these results, showing that deliberation makes a large difference in how undergraduates play. When the students are not allowed to deliberate they take 24 points from the common resource. When they are allowed to deliberate they only take 9 points from the common pool. In other words, they take 150 percent more when they do not deliberate. Just like with the subjects in Ostrom, Walker, and Gardner’s (1992) classic study, “mere jawboning (p. 413)” leads to more cooperation for our students. At the same time, our results suggest that these results are not applicable to the policymakers who exert control over public goods. Talk was cheaper for legislators, who did not significantly shift their play as a result of deliberation.

We do not purport to make perfect predictions about how these elected officials would respond to institutional changes within their legislative bodies, with millions of dollars of funding and their political reputations at stake. Laboratory experiments are necessarily artificial in some ways. Although our treatments were designed to be analogous to institutional differences, playing a one-shot game is not exactly like facing term limits, and receiving hidden information about a multiplier effect is not quite the same as being advised by staff working exclusively for the majority party. While we ensured that players had incentives to play well, and took pains to ensure that lawmakers and students had the same incentives, we could not hope to replicate the stakes involved in real politics. Our approach only makes the sample of players equivalent to the political actors whom we wish to make inferences about; we cannot fully replicate the legislative setting. Still, we can speculate about how this does and does not limit our findings. If mere deliberation brings no significant increase in public goods provision among legislators playing for small contributions to public universities, this gives us little hope that it will increase cooperation when the distribution of millions of dollars of funding is at issue.
More broadly, these results suggest that we cannot directly infer the behavior of elites based on the results of experiments using undergraduate samples. That is not to say we should never do so. The ability to manipulate institutions in the lab setting offers a great opportunity to study institutional design. We should not abandon that pursuit. We believe that such a pursuit is critically important for our discipline going forward. However, our results do suggest some ways of going about that research agenda.

First, scholars have a great opportunity to study how various institutions affect the behavior of their undergraduate student samples. In many cases the undergraduates can be good estimates of voters writ large (Roth 1998) or of the users of common pool resources (Ostrom, Walker, and Gardner 1992). Yet given the recent advances in technology, we do not necessarily have to rely on the lab as our only venue for experimentation. By moving our experiments to be delivered over the web, we are coming even closer to learning about the population of interest. Lab-in-the-field experiments hold great promise for helping us make sharper inferences about populations of interest (even for those interested in the behavior of regular citizens).

Second, to the extent that we continue to use undergraduate (or other mass public samples) to reach conclusions about elites, we need to acknowledge the biases in doing so. Undergraduates are more changing and susceptible to social pressure (Sears 1968), and so treatments that tap those considerations will have a much larger effect among them than they would among public officials. At the same time, the legislators in our study were more responsive to the informational treatment. This appears to reflect the advantage that they gain by having experience working in legislative environments. As a result, using student samples to estimate the effect of informational treatments will result in biased estimates of the true effect these institutions would have on legislators. When researchers use students in
In lab settings to make inferences about how legislators will behave, we need to take into account the differences between these populations.

Finally, researchers should cultivate the relationships to involve the relevant elites in the process of the experiment. The best way to learn about elites is to study elites. While there are many ways study elites, we believe that building relationships with the organizations that service public officials (such as the Council of State Governments) provide the most promising avenue for researchers hoping to conduct experiments as these groups may have longer time horizons and a greater commitment to institutional development than individual elected officials.
Box 1. Review of articles in the APSR, AJPS, and JOP, 2002-2012

<table>
<thead>
<tr>
<th>Type of Lab Experiment</th>
<th>Total Number of Articles</th>
<th>Extrapolated Results to Elite Behavior</th>
<th>Did Not Make Inferences to Elite Behavior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lab experiments</td>
<td>51</td>
<td></td>
<td></td>
</tr>
<tr>
<td>without interactions</td>
<td>29</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>involving interactions</td>
<td>22</td>
<td>11</td>
<td>11</td>
</tr>
</tbody>
</table>

Total number of articles using lab experiments: 51

Lab experiments without interactions between participants: 29

  - Extrapolated results to elite behavior: 3
  - Did not make inferences to elite behavior: 26

Lab experiments involving interactions between participants: 22

  - Extrapolated results to elite behavior: 11
  - Did not make inferences to elite behavior: 11
Figure 1. How Much Do Players Take From the Table Fund, averaged across all periods?

Note: The difference in the mean averages for the two samples is significant at the $p < 0.10$ level.
Figure 2. How Much Do Withdrawals Vary Across Periods for Each Player?

Note: The difference in the mean standard deviations for the two samples is significant at the $p < 0.05$ level.
Figure 3: The Impact of Experimental Manipulations on Legislators
Figure 2: The Impact of Experimental Manipulations on Students
<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>All</td>
<td>Legislators</td>
<td>Students</td>
</tr>
<tr>
<td>Deliberation-Post</td>
<td>-9.2**</td>
<td>-4.1</td>
<td>-15.2**</td>
</tr>
<tr>
<td></td>
<td>(3.8)</td>
<td>(4.7)</td>
<td>(6.2)</td>
</tr>
<tr>
<td>Short-term</td>
<td>0.6</td>
<td>-3.1</td>
<td>5.0</td>
</tr>
<tr>
<td></td>
<td>(3.8)</td>
<td>(4.9)</td>
<td>(5.8)</td>
</tr>
<tr>
<td>Long-term</td>
<td>2.1</td>
<td>2.4</td>
<td>1.8</td>
</tr>
<tr>
<td></td>
<td>(4.1)</td>
<td>(5.7)</td>
<td>(6.1)</td>
</tr>
<tr>
<td>Information-Symmetric</td>
<td>4.8</td>
<td>1.0</td>
<td>9.2</td>
</tr>
<tr>
<td></td>
<td>(4.2)</td>
<td>(5.8)</td>
<td>(6.0)</td>
</tr>
<tr>
<td>Information-Asymmetric</td>
<td>6.7</td>
<td>6.1</td>
<td>7.4</td>
</tr>
<tr>
<td></td>
<td>(4.6)</td>
<td>(6.2)</td>
<td>(7.0)</td>
</tr>
<tr>
<td>Constant</td>
<td>34.2***</td>
<td>4.9</td>
<td>33.6***</td>
</tr>
<tr>
<td></td>
<td>(11.4)</td>
<td>(4.7)</td>
<td>(11.3)</td>
</tr>
<tr>
<td>Individual F.E.s?</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>406</td>
<td>221</td>
<td>185</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.465</td>
<td>0.534</td>
<td>0.382</td>
</tr>
</tbody>
</table>

Notes: The dependent variable is the amount that a player withdrew from the table fund in a given experimental condition. Robust standard errors clustered on individuals (is this true?) in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Table 2: The Impact of Asymmetric Information

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>All</td>
<td>Legislators</td>
<td>Students</td>
</tr>
<tr>
<td>Multiplier=1.5, Common Knowledge</td>
<td>24.3**</td>
<td>19.6</td>
<td>27.9</td>
</tr>
<tr>
<td></td>
<td>(9.4)</td>
<td>(12.0)</td>
<td>(16.5)</td>
</tr>
<tr>
<td>Multiplier=1.5, Informationally Advantaged</td>
<td>33.0***</td>
<td>39.4***</td>
<td>26.6</td>
</tr>
<tr>
<td></td>
<td>(9.1)</td>
<td>(10.9)</td>
<td>(16.6)</td>
</tr>
<tr>
<td>Multiplier=1.5, Informationally Disadvantaged</td>
<td>12.8</td>
<td>9.8</td>
<td>15.7</td>
</tr>
<tr>
<td></td>
<td>(12.9)</td>
<td>(15.5)</td>
<td>(22.9)</td>
</tr>
<tr>
<td>Multiplier=3, Informationally Disadvantaged</td>
<td>2.4</td>
<td>12.1</td>
<td>-18.8</td>
</tr>
<tr>
<td></td>
<td>(13.0)</td>
<td>(15.8)</td>
<td>(22.9)</td>
</tr>
<tr>
<td>Multiplier=3, Informationally Advantaged</td>
<td>-0.3</td>
<td>-4.8</td>
<td>2.6</td>
</tr>
<tr>
<td></td>
<td>(8.9)</td>
<td>(12.3)</td>
<td>(13.7)</td>
</tr>
<tr>
<td>Constant</td>
<td>31.3</td>
<td>-4.9</td>
<td>32.7</td>
</tr>
<tr>
<td></td>
<td>(20.7)</td>
<td>(19.1)</td>
<td>(26.1)</td>
</tr>
</tbody>
</table>

Observations: 135
R-squared: 0.7

Notes: The dependent variable is the amount that a player withdrew from the table fund in a given experimental condition. Robust standard errors clustered on individuals in parentheses. *** p<0.01, ** p<0.05, * p<0.1
Table 3. The Impact of Individual Characteristics on Legislator Behavior

<table>
<thead>
<tr>
<th>VARIABLES</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Minority Party</td>
<td>12.3*</td>
<td>(6.5)</td>
</tr>
<tr>
<td>Republican</td>
<td>19.6**</td>
<td>(8.3)</td>
</tr>
<tr>
<td>Term Limits</td>
<td>10.4</td>
<td>(11.7)</td>
</tr>
<tr>
<td>Oregon</td>
<td>2.0</td>
<td>(12.5)</td>
</tr>
<tr>
<td>Male</td>
<td>7.7</td>
<td>(5.4)</td>
</tr>
<tr>
<td>Married</td>
<td>-32.0***</td>
<td>(10.5)</td>
</tr>
<tr>
<td>Age</td>
<td>3.2</td>
<td>(2.3)</td>
</tr>
<tr>
<td>Age-Squared</td>
<td>-0.0</td>
<td>(0.0)</td>
</tr>
<tr>
<td>Constant</td>
<td>-65.5</td>
<td>(62.8)</td>
</tr>
<tr>
<td>Observations</td>
<td>335</td>
<td></td>
</tr>
<tr>
<td>R-squared</td>
<td>0.2</td>
<td></td>
</tr>
</tbody>
</table>

Notes: The dependent variable is the amount that a player withdrew from the table fund in a given experimental condition. We include a variable indicating the lawmakers who served in the Oregon House, in which the two parties held an equal number of seats at the time of the experiment. Robust standard errors clustered on individuals in parentheses. *** p<0.01, ** p<0.05, * p<0.1
References


