



Project
MUSE[®]
Scholarly journals online

THE ROLE OF LEADERS IN DEMOCRATIC DELIBERATIONS

Results from a Field Experiment in São Tomé and Príncipe

By MACARTAN HUMPHREYS, WILLIAM A. MASTERS, and
MARTIN E. SANDBU*

I. INTRODUCTION

STRONGLY participatory approaches to political decision making have been widely promoted in a bid to deepen the practice of democracy.¹ Supporters of these approaches attribute multiple benefits to participation, on normative, substantive, and instrumental grounds. However, there is to date little or no evidence that these practices in fact return the benefits attributed to them. This is of particular concern, since there are strong arguments from the political and anthropological literatures that suggest that such processes may be subject to various forms of manipulation.

This article reports on a field experiment that provides an unusual opportunity to examine the extent to which participatory processes of

* Macartan Humphreys is the corresponding author (mh2245@columbia.edu). Our thanks go to Carlos Gustavo dos Anjos, Justino Veiga, and other members of the National Forum committee, who welcomed the idea of randomization of leader assignment, and to José Manuel Dias Marques, Tomé Santos, and Carlos Tiny, who played a dynamic role in the planning and running of the National Forum. We also thank Albano Germano de Deus and his staff at the National Institute of Statistics for superb execution of the survey component, as well as Clara Albuquerque, Joshua Chaffin, Paulo Cunha, Jan Hartman, and Christine Weydig for their tireless work to help implement the forum, Camille Naquet-Radiguet (now Mazas) for assistance with survey design, Eduardo Leoni for research support, and the Open Society Institute, the Earth Institute at Columbia University, and the UNDP in São Tomé and Príncipe for financial support. Heartfelt thanks to Paul Richards, Blaise Gonda, Paulo Cunha, Daniel Cordes, Abigail Barr, Pedro Vicente, and participants at the Oxford CSAE microeconomics seminar for conversations and comments. To ensure the transparency and credibility of their process, the organizers of the São Tomé and Príncipe National Forum agreed to a randomized design and publication of the resulting data in a public domain. Portuguese and English versions of the forum design and results have been posted since 2004 on the project's Web site (www.earth.columbia.edu/cgsd/STP), alongside analyses of the data conducted by local Saotomean consultants as well as by Columbia researchers.

¹ Prominent examples include Benjamin Barber, *Strong Democracy: Participatory Politics for a New Age* (Berkeley: University of California Press, 1984); Dan Corry and Gerry Stoker, *New Localism: Refashioning the Centre-Local Relationship* (London: New Local Government Network, 2002); and Bruce Ackerman and James S. Fishkin, "Deliberation Day," *Journal of Political Philosophy* 10 (June 2002). Trends in developing areas are discussed in more detail below.

this form are in practice vulnerable to manipulation by political elites. The occasion for the study was an exercise in deliberative democracy in the small African island state of São Tomé and Príncipe. In 2004 the country held a national forum in which citizens gathered in small groups throughout the country to discuss topics related to economic policy priorities for the country. In many ways the consultations, open to every adult in the country and managed in part by the United Nations Development Program (UNDP), were a model of consultative practices in developing countries.

Unlike other such consultation processes, however, the organizers introduced a random element to the design of their consultation: the leaders who moderated the discussions were assigned randomly to groups throughout the country.

This feature, along with a formal reporting mechanism on outcomes of discussions, affords a unique opportunity to identify the impact of leaders on the preferences expressed by groups. While we cannot identify the difference between outcomes that would obtain with and without leaders, we can exploit random variation in the identity of leaders to establish that leaders matter: we do this by showing how the identities of leaders determine outcomes.² Identifying the impact of discussion leaders is usually rendered difficult by a selection problem: leaders' characteristics could be linked to discussion outcomes because of the criteria used in choosing leaders, rather than because of any independent influence of the leaders as such. But we are able to overcome this identification problem because of the randomization introduced into the Saotomean forum.

The findings are striking. Even though in many ways the deliberations in São Tomé and Príncipe were held in an ideal communication environment, there is robust evidence that the influence of leaders on the outcome of deliberation is extremely strong, with leadership effects accounting for a large share of the variation in views elicited across the country. If similar dynamics are at work in such consultations elsewhere, then our evidence suggests that these participatory practices may not return the benefits so commonly attributed to them by civil society groups, governments, and international organizations.

Our examination proceeds as follows. In Section II we describe the nature and rationale behind participatory decision processes, as well as

² As an analogy, a study showing systematic differences between majoritarian versus PR systems could establish that institutions matter even in the absence of a no-institution benchmark. The converse is not true, however. If there is no variation in outcomes associated with different types of institutions, this does not establish that institutions do *not* matter; for such a claim a no-institution benchmark would be needed. We return to this point below.

why we might expect them to be vulnerable to political manipulation that would thwart their aims. In Section III we describe the political context of the deliberations in São Tomé and Príncipe, and in Section IV we discuss the structure of the deliberations in more detail. In Section V we present the main results of our analysis, estimating the degree of leadership bias and determining systematic correlates of bias. In Section VI we move from identification to explanation and seek to relate the outcomes of the discussions to rival mechanisms that can explain leader influence. Section VII addresses the question of the external validity of our results and Section VIII concludes.

II. PARTICIPATORY POLITICS

Participatory processes have long been advocated by civil society groups in developing countries. Since the 1980s they have also been advocated by many bilateral donors. According to one account, “By the early 1990s every major bilateral development agency emphasized participatory policies.”³ Participatory rural appraisals (PRAs)—intended, according to Chambers’s definition, “to enable local (rural and urban) people to express, enhance, share and analyze their knowledge of life and conditions, to plan and to act”⁴—are now employed globally and have indeed become a criterion for funding by some development agencies.⁵

Since the mid-1980s the World Bank and the IMF have thrown their weight strongly behind participatory approaches, recently structuring much of their day-to-day activities around the generation of national-level Poverty Reduction Strategy Papers (PRSPs). The first core principle of PRSPs is that they be country driven, with broad-based participation, because, these institutions argue, participation “can contribute to higher quality strategies, and strengthen the environment for governance and accountability.”⁶

This position has had knock-on effects for governments of developing countries: the production of PRSPs has been a central, time-consuming activity of ministries of development and ministries of finance

³ Heiko Henkel and Roderick Stirrat, “Participation as Spiritual Duty: Empowerment as Secular Subjection,” in Bill Cooke and Uma Kothari, eds., *Participation: The New Tyranny* (London: Zed Press, 2001), 168.

⁴ Robert Chambers, “The Origins and Practice of Participatory Rural Appraisal,” *World Development* 22, no. 7 (1994), 1253.

⁵ Paul Richards, “Participatory Rural Appraisal: A Quick-and-Dirty Critique,” *PLA Notes* 24 (1995).

⁶ International Monetary Fund, *2005 Review of the Poverty Reduction Strategy Approach—Balancing Accountabilities and Scaling Up Results—Synthesis*, <http://www.imf.org/external/np/pp/eng/2005/091905s.htm> (accessed September 10, 2006).

around the globe. In 2005 some forty-nine developing countries had already produced PRSPs, and another twenty-one were in earlier stages of the PRSP process. Prominent in these strategy papers, typically highlighted in the first chapter, is a section detailing the lengths to which governments have gone to consult with their populations.

What is the rationale behind this push for more participatory decision processes? There are three prominent families of argument, based respectively on the supposed *normative*, *substantive*, and *instrumental* benefits of participation.

The normative benefits of participation have been emphasized especially by Sen:

Such processes as participation in political decisions and social choice cannot be seen as being—at best—among the *means* of development (through, say, their contribution to economic growth), but have to be understood as constitutive parts of the *end* of development in themselves.⁷

In principle, Sen and others argue, even if participatory processes fail to produce “good” decisions, they are still of value in that they provide a space in which the voiceless can make their views heard.

Beyond these normative benefits, participation may result in better substantive outcomes. Habermas, for example, argues that participation, and deliberation in particular, leads to higher-quality decisions: reason replaces power in the determination of outcomes, with the result that outcomes are not simply more just but also more “rational.”⁸ Echoing an argument that dates at least from Condorcet, Stiglitz suggests that participation improves information, which can itself promote development.⁹

Finally, even if normative and substantive benefits do not obtain, there can still be instrumental benefits. Even if participation yields no information of use to (or, at least, used by) politicians, it can still generate political buy-in for economic reform. Policies are more likely to be accepted if they are decided upon jointly: “Because individuals have had a voice in shaping the changes, in making them more acceptable, change is likely to be accepted or even embraced, rather than reversed at the first opportunity.”¹⁰ While this may be of benefit to governments, the instrumental benefits of participation may also be valued by international

⁷ Amartya Sen, *Development as Freedom* (New York: Anchor Books, 1999), 291.

⁸ Jürgen Habermas, *Theory of Communicative Action*, vol. 1, *Reason and the Rationalization of Society* (Boston: Beacon Press, 1984); idem, “Three Normative Models of Democracy,” in Sheyla Benhabib, ed., *Democracy and Difference* (Princeton: Princeton University Press, 1996).

⁹ Joseph E. Stiglitz, “Participation and Development: Perspectives from the Comprehensive Development Paradigm,” *Review of Development Economics* 6, no. 2 (2002)

¹⁰ *Ibid.*, 168.

organizations. Piron and Evans (citing Norton)¹¹ argue that by promoting participation, the Bretton Woods institutions are betting on the idea that “if governments are obliged to discuss poverty and what they are doing about it with their citizens, [then] they are likely to regard these things more seriously, and to be held to account more effectively.”¹²

Among participatory processes, deliberative processes have been given pride of place.¹³ Public deliberation is political participation par excellence. Its purported benefits include an ability both to engage citizens more actively in decision making and to aggregate information more effectively, thereby making collective decisions more genuinely representative of the people’s preferences. As argued by Hicks, deliberative processes hold out a promise of greater equality than do other forms of democracy: “Public deliberation, because it gives equal consideration to all views and affords all persons the equal opportunity of political influence, is capable of transforming both the content of citizens’ preferences and their political conduct so they are compatible with the demands of justice.”¹⁴

The arguments in favor of public deliberation are compelling and, admittedly, much richer and more varied than we have conveyed here. But to varying degrees they all rest on the asserted ability of deliberative processes to generate decisions that ostensibly represent the “will of the people” to a greater extent than do other collective decision mechanisms. The problem, however, is that public deliberation, like other political processes, is, in principle, vulnerable to manipulation by discussion leaders and group members. Indeed, formal models of collective decision making suggest that *any* mechanism that produces a single aggregate preference ordering from individual preferences is vulnerable to manipulation (see especially results due to Gibbard and Satterthwaite;¹⁵ for a contrary position, see the research by List and Dryzek¹⁶). Moreover, we know from other contexts that expressed opinions are often strongly influenced by perceptions of what those collecting them want to hear, as

¹¹ Andrew Norton, “Approach Paper: Politics and the PRSP Approach,” Working Paper, PRSP Monitoring and Synthesis Project (London: Overseas Development Institute, 2002).

¹² Laure-Hélène Piron and Alison Evans, “Politics and the PRSP Approach: Synthesis Paper,” ODI Working Paper 237 (London: Overseas Development Institute, 2004).

¹³ Ackerman and Fishkin (fn. 1)

¹⁴ Darrin Hicks, “The Promise(s) of Deliberative Democracy,” *Rhetoric and Public Affairs* 5 (Summer 2002).

¹⁵ Allan Gibbard, “Manipulation of Voting Schemes: A General Result,” *Econometrica* 4, no. 4 (1973); Mark Satterthwaite, “Strategy-Proofness and Arrow’s Conditions: Existence and Correspondence Theorems for Voting Procedures and Social Welfare Functions,” *Journal of Economic Theory* (April 1975).

¹⁶ Christian List and John S. Dryzek, “Social Choice Theory and Deliberative Democracy: A Reconciliation,” *British Journal of Political Science* 33, no. 1 (2003).

seen, for example, in the well-known interviewer effect found in individual surveys.¹⁷ Deliberative approaches are sometimes recommended in order to get around shortcomings associated with survey approaches; if, however, we find that similar biases are at work in deliberative decision making, then there are considerable policy implications for proponents of these approaches.

The concern that collective decision making is susceptible to undue influence by individuals is not new; the Athenians practiced the institution of ostracism to remove influential individuals from their polity in the fifth century BCE. But even these ostracisms were themselves (in Plutarch's account at least) susceptible to undue influence by individuals who used them, like Themistocles against Aristides, to target political rivals.¹⁸ In modern times, political processes that take place at a higher level of aggregation have used checks and balances to defend minority rights. But large states have also been vulnerable to capture, leading to demands for more decentralized decision making. Clearly, as detailed by Bardhan and Mookherjee, the scale of decision making has ambiguous effects.¹⁹ Studies of micro-decision making processes in developed countries reveal how power relations determine outcomes even when participants are formally equals.²⁰ In the case of deliberative institutions in particular, Pelletier et al. find that in New York State "local deliberative processes may produce outcomes that are neither fair nor efficient and that reflect the values and interests of certain stakeholders more than others, even in the absence of overt conflict."²¹ Some studies suggest that the choice of language itself strongly conditions the outcome of conversations or that the search for consensus results in the subordination of minority voices.²² Elster notes that when underlying

¹⁷ Robert Hanson and Eli Marks, "Influence of the Interviewer on the Accuracy of Survey Results," *Journal of the American Statistical Association* 53 (September 1958); Colm O'Muircheartaigh and Pamela Campanelli, "The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision," *Journal of the Royal Statistical Society Series A (Statistics in Society)* 161, no. 1 (1998).

¹⁸ Plutarch, 75 c.e., *Aristides* <http://classics.mit.edu/Plutarch/aristide.html> (accessed September 10, 2006).

¹⁹ Pranab Bardhan and Dilip Mookherjee, "Capture and Governance at Local and National Levels," *American Economic Review, Papers and Proceedings* 90 (May 2000); idem, "Pro-Poor Targeting and Accountability of Local Governments in West Bengal," *Journal of Development Economics* 79 (April 2006).

²⁰ Steven Lukes, *Power: A Radical View* (New York: Macmillan, 1974); Peter Bachrach and Morton S. Baratz, "The Two Faces of Power," *American Political Science Review* 56 (December 1962).

²¹ David Pelletier, Vivika Kraak, Christine McCullum, Uula Uusitalo, and Robert Rich, "The Shaping of Collective Values through Deliberative Democracy: An Empirical Study from New York's North Country," *Policy Sciences* 32, no. 2 (1999).

²² John S. Dryzek and Douglas Torgerson, "Democracy and the Policy Sciences: A Progress Report," *Policy Sciences* 26, no. 3 (1993); Ilan Kapoor, "The Devil's in the Theory: A Critical Assessment of Robert Chambers' Work on Participatory Development," *Third World Quarterly* 23, no. 1 (2002).

preferences are in conflict, deliberation that turns private knowledge into common knowledge may lead to new disagreements over policies that different people formerly agreed on.²³ A corollary is that if a goal of deliberative processes is consensus, those who manage those processes may have incentives *not* to encourage the free flow of information.

In developing areas there has been some awareness of such concerns, but much of the ensuing impetus has been to limit the influence of “outsiders”—foreign development experts and agencies. In focusing on this insider/outsider dichotomy, however, such responses downplay local social inequalities and power relations.²⁴ In the absence of evidence to the contrary we should expect instead that local power relations will affect democratic processes in developing areas much as they do elsewhere. More anthropological approaches have directed our attention to these local dynamics, emphasizing how the outcomes of discussions can depend on the “cultural biases” of a society.²⁵ Consensual discourse, anthropologists suggest, can be a “strategy for managing power, dominance and divisive interests” and used as a mechanism to avoid public accountability on the part of leaders or to reaffirm a social order.²⁶ In a fascinating study of consensual group meetings in Sierra Leone, Murphy argues that public meetings are used strategically to ensure that disagreement take place in a space that is manageable and in a way that provides greater legitimacy to social arrangements.²⁷ He finds, however, that actual *outcomes* are determined by preexisting, “backstage” hierarchies. In such cases the discussions may provide the instrumental returns of participation without producing normative or substantive benefits.

These concerns raise doubts about whether participatory processes do in fact realize their promise of more genuinely democratic decisions. Defenses of participation—especially those drawing on normative and substantive arguments—presuppose that outcomes are in fact faithful expressions of the preferences of participants (in particular, postdeliberation preferences) and are not determined by those in a position to manipulate the process. In this article we focus on that minimal condition and scrutinize an ambitious exercise in deliberative democracy to test whether the condition was met. Specifically, we examine whether the

²³ Jon Elster, “Deliberation and Constitution-Making,” in Jon Elster, ed., *Deliberative Democracy* (New York: Cambridge University Press, 1998).

²⁴ Giles Mohan and Kristian Stokke, “Participatory Development and Empowerment: The Dangers of Localism,” *Third World Quarterly* 21, no. 2 (2000).

²⁵ Mary Douglas, *Natural Symbols: Explorations in Cosmology* (London: Barrie and Rockliff, 1970).

²⁶ William P. Murphy, “Creating the Appearance of Consensus in Mende Political Discourse,” *American Anthropologist* 92, no. 1 (1990).

²⁷ *Ibid.*

decisions made by deliberative groups were driven by who the group leader was. Our concern is not to assess the relative merit of the different theories of deliberative democracy; instead, we examine whether reality fits the presupposition that they make. At the same time, we recognize that the idea that deliberation could be subject to manipulation does not imply that it is *more* subject to capture than other political processes. The point is rather that if deliberative democracy is to be advocated, it should be on the basis of an understanding of when, how, and to what extent it is open to capture—an understanding that current attempts to evaluate participatory processes do not yield. We return to these concerns in the concluding section.

III. THE SETTING

With an area of just one thousand square kilometers and a population of only 160,000 people, São Tomé and Príncipe is one of the smallest democracies in the world. Following a period of one-party rule after gaining independence from Portugal in 1975, São Tomé and Príncipe held its first multiparty elections in January 1991. Representative democracy in São Tomé and Príncipe has survived intact since then, albeit often with uneasy relations between the presidency and the government, frequent changes in the government, and two abortive coup attempts.

The National Forum deliberations took place in the context of a distinctive kind of political shock, namely, the discovery of possibly large amounts of offshore oil. Geological data suggest that reserves could be as high as fifteen billion barrels, although no commercially viable discoveries have yet been declared. Such prospects have attracted international attention to the islands, not least from São Tomé and Príncipe's neighbor Nigeria. Following the discoveries, Nigeria contested the location of the São Tomé and Príncipe/Nigeria maritime border, arguing that if land mass were taken into account, an appropriate redrawing of the border would place the oil reserves squarely within Nigeria's territory. In an out-of-court settlement in 2001, the two countries agreed to leave the ownership question unresolved and to manage the disputed area jointly in the interim through a joint development authority, 60 percent of which would be controlled by Nigeria.

Even taking account of this agreement, calculations in São Tomé and Príncipe yield some heady numbers. Assuming, for example, a price of \$50 a barrel and an effective tax rate of 50 percent, São Tomé and Príncipe's 40 percent share of those taxes could yield nearly \$1 million for every Saotomean, man, woman, and child ($\$50 \times 15\text{bn} \times .5 \times .4 / 160,000$).

Such projections are very dangerous, however. Actual amounts in the ground could be much lower and the price of oil may fall; moreover, the time it takes to produce oil needs to be factored in.

Nevertheless, even in the absence of actual oil revenues, expectations of oil are already influencing domestic politics on the islands, much as they are influencing the country's external relations. In what seemed like a textbook instance of the logic of the resource curse, disclosure of the possible oil discoveries was soon followed by an attempted coup d'état. On July 16, 2003, while President Menèzes was in Nigeria, a small group of former mercenaries, allied with the armed forces, arrested the prime minister and members of the cabinet.²⁸ Although likely motivated largely by personal interests, the coup makers voiced grievances that centered on fears over impending mismanagement of the oil sector and the future oil economy.

International condemnation of the coup, combined with diplomatic intervention, led to a Memorandum of Understanding between the coup leaders and the government. Nigeria, Portugal, and the United States, among others, served as guarantors for the agreement, which reinstalled the government and Menèzes as president. One of the conditions of the agreement was that the nation would hold a "national forum to listen to political parties and civil society."

IV. DESIGN OF THE NATIONAL FORUM

The Memorandum of Understanding that resolved the political crisis in São Tomé and Príncipe called for a national forum. But beyond calling for dialogue, it did not specify either the form or the subject matter of the forum. President Menèzes, responsible for ensuring that the forum would take place, established a forum committee in São Tomé and Príncipe and asked researchers at Columbia University who were providing advisory services to the government to propose a design for the forum.

The design that was ultimately worked out was intended to offer every adult citizen of São Tomé and Príncipe the opportunity to attend public meetings, at which they would be informed about the nature of potential oil revenues, participate in a discussion about how those oil revenues might be spent, and then have their group's expenditure priorities recorded and brought to the attention of the government by a representative of their group.

²⁸ Gerhard Seibert, "Coup d'Etat in São Tomé and Príncipe: Domestic Causes, the Role of Oil and Former 'Buffalo' Battalion Soldiers," in *African Security Analysis Programme*, Occasional Paper, Institute for Security Studies (October 10, 2003), 2.

The forum in effect had two components—a first stage of popular consultations, in which 148 groups met at 56 sites throughout the country, and a second stage of plenary debate among group representatives and government officials in the capital city. In the first phase, meetings began with the distribution and discussion of an information booklet about the potential oil revenues. After a question-and-answer period, participants broke into smaller groups to engage in deliberations aimed at completing a questionnaire designed by the forum committee. In the second stage, first-stage results were collated by consultants engaged by the government and used as a basis for discussions during three days of meetings in the capital, involving both elites and representatives from the 148 previous meetings.

Our study focuses on the design and results of the first phase, when the results of grassroots deliberations were initially recorded. The major elements of the design of this first phase are the following.

MEETING STRUCTURES

The popular deliberations were daylong events held at public sites across the country, in schools, churches, and plazas. To each site, the forum committee sent a team of three to four discussion leaders, one of whom was identified in advance as the team leader (*moderador*). The meetings in each site began with a detailed presentation by the *moderador* about the prospect of oil revenues and their potential impact on government economic policy. This informational component included the distribution of a “popular information bulletin” written in nontechnical language that, in comparative terms for the region, provided an extraordinary amount of information to individuals about the oil sector and the rights of individuals with respect to the management of oil resources.

The information bulletin explained the possible location and quantities of oil, provided the expected timeline for production, gave an account of the new oil law and its provisions for a permanent fund, and provided information regarding plans for a public information office and an oversight commission for oil-related affairs. It also gave a general accounting of the sources of government revenue and the allocations of government expenditure. The team leader then led a plenary discussion of the contents of the bulletin, gave an introduction to the forum and its aims, and fielded questions.

Following these information sessions, each meeting broke up into smaller deliberation groups (of typically fifteen to twenty people) to discuss and record expenditure priorities. Individuals were grouped and assigned a discussion leader (*facilitador*) at random, with individuals

and leaders being given an ordering and every n^{th} individual being assigned to take part in the group to be led by the n^{th} discussion leader. The deliberation groups then convened to discuss a set of questions, determined in advance and listed in a discussion leader's form (the *Questionário do Fórum Nacional*).²⁹

The form contained three types of questions. One set contained questions of a factual nature at the individual level, for example, about access to electricity and experience with crime and petty corruption; for these a poll was taken within the group. A second set was answered directly by the leader: how many participated in the meeting, what was the age and gender composition, and so on. Finally, the form contained twelve "deliberative questions" to be discussed and answered collectively by the group. These are the focus of our study.

In most cases these deliberative questions asked the group to rank a small number of policy alternatives. One focused on health priorities, three on education, and three on transportation priorities, two of which, (7a) and (7b), are closely related. One question elicited priority sectors for national-level expenditures. Two were intended to estimate a discount rate, specifically, a required interest rate to justify saving wind-fall revenue for one year (the first, a qualitative question, asked whether any deferral is justified; the second, conditional upon a deferral being justified, asked what return would be required). Another question focused on optimal income tax rates. A final question asked communities to report their beliefs about national-level corruption. An abbreviated listing of these questions is provided in Table 1.

For each deliberative question, the group was asked to discuss the merits and demerits of the different possibilities and to decide, collectively, on an answer. Once reached, the leader confirmed to the group the answer that he or she intended to record and then marked the answer. Since, in principle, the goal of achieving consensus may do more harm than good (relative to "agnostic pluralism"),³⁰ the leader could select that option that was supported only by a majority or a plurality. In all cases the leader would mark not only the group's collective answer but also the degree of consensus on the issue within the group, ranging from "complete disagreement" to "most agree" to "all agree." Unlike some deliberative settings (such as "deliberative polls"),³¹ alternative

²⁹ Available on-line at http://www.earthinstitute.columbia.edu/cgsd/STP/documents/Questionnaire_P_final_000.pdf.

³⁰ Chantal Mouffe, "Deliberative Democracy or Agonistic Pluralism?" *Social Research* 66, no. 3 (1999).

³¹ James S. Fishkin, *Democracy and Deliberation: New Directions for Democratic Reform* (New Haven: Yale University Press, 1991).

TABLE 1
DELIBERATIVE QUESTIONS

<i>Issue # on Form^a</i>	<i>Issue-Area</i>	<i>Number of Communities Responding</i>
2	What is the #1 priority developmental sector? Coded 1 for Health and 0 for Other	134
3	Health priorities: Clinics (0) or Hospitals (1)?	133
4a	Education priorities: Primary (0) or Secondary (1)	138
4b	Education priorities: Professional (0) or University (1)?	144
4c	Education priorities: Advanced and Specialized (0) or Basic and Universal (1)?	142
7a	Transportation priorities: improve conditions for Commercial (0) or Passenger (1) Travel?	131
7b	Transportation priorities: improve Quality of Roads (0) or of Public Transportation Services (1)?	132
7c	Transportation priorities: improve Links between Villages (0) or between Major Centers (1)?	128
11a	Discount rate: Windfalls should be Consumed (0) or Invested (1)?	143
11b	Discount rate: required annual return to justify Investment of Windfall (percentage)?	81
12	Taxation: share of private windfalls that should be taxed (percentage)?	126
14c	Prevalence of National-Level Corruption (1–5)	138

^aThe discussion leaders' form is reproduced online at http://www.earthinstitute.columbia.edu/cgsd/STP/documents/QuestionnaireP_final_000.pdf.

arguments for one response rather than another were not provided, nor was there external expert information to guide decisions.³² Instead, as is common practice in PRSP consultations, for example, the input of leaders was kept to a minimum and the communities were to draw on the knowledge and views of their members in answering the questions.

SURVEY OF FORUM IMPACT

In parallel with the forum meetings, we collaborated with the National Statistics Institute of São Tomé and Príncipe to carry out an individual-level survey of attitudes on the same issues discussed in the forum. A total of 266 individuals were interviewed in their locality before the forum meeting; 190 of those individuals were then found and interviewed again after the forum. Another 351 individuals were interviewed only after the forum. Of the total of 541 subjects interviewed after their lo-

³² For an argument against such structures, see Chambers (fn. 4), 1441.

cal forum meeting, 247, or around 45 percent, had attended the forum. The survey respondents were sampled so as to represent the forum-going population in the following way. Each electoral constituency (113 in São Tomé island and 17 on Príncipe) was matched with a forum location (the closest one). For each forum location, we then picked one constituency at random.³³ The enumerators were given a protocol to randomly select one sample of respondents who intended to attend the forum (or had attended, in the post-forum surveys), one sample of respondents who intended not to attend (or had not attended), and one sample of respondents who were not asked about their intentions.³⁴ The questionnaires administered to the respondents were virtually identical to the form used to guide the deliberation in the forum group sessions. Note that we know only *that* a given individual participated in the forum and not the specific deliberative group in which he or she participated. Still, we can use this information to examine whether the differences between prior positions overall of individuals and the reported collective preferences that were recorded in the forum meetings are also apparent between pre-forum and post-forum survey responses from individuals who attended (for this we use only sample averages of the survey responses).

PARTICIPANTS

Although the government endeavored to ensure that the sites were widely distributed throughout the country and that attendance was open to all citizens, ultimately participation was a matter of self-selection. Approximately 3,500 Saotomeans attended the meetings with an average of 24 attendees in a given deliberative session. Half of the meetings had between 16 and 29 people, and one meeting, exceptionally, had 131 attendees. Typically one-third of the attendees were women and about one in ten were elders. The number 3,500 represents about 2.2 percent of the total population of São Tomé and Príncipe and

³³ The Saotomean authorities chose to organize proportionately more forum meetings in the less-populated districts, to give equal access to all citizens. Our sampling method follows this approach and should therefore be seen as representative of the population as it was targeted by the forum meetings, rather than of the unweighted national population. There were two districts for which we could not sample respondents before the forum meetings, since the forum was already starting in those locations at the time the survey interviews were launched. The remaining districts, which we did sample, account for 86 percent of the voting population in the country. The forum meetings in the two missing districts (Caué and Cantagaló), moreover, were demographically similar to the forum-going population nationally.

³⁴ For more details about the survey and a more detailed analysis of the effect of forum participation on privately reported individual preferences, see Martin E. Sandbu, "Does Deliberation Engender Public-Spiritedness? A Study of Deliberative Democracy in São Tomé and Príncipe's National Forum" (Manuscript, Wharton Business School, 2006).

perhaps 5 percent of the adult population. In occupational terms, over one-quarter identified themselves as unemployed; most others classified themselves as fishermen (12 percent), cocoa producers (15 percent) or “other agricultural workers” (16 percent). Approximately one in five reported working in professional activities.

Our approach to identifying the causal impact of leaders does not depend on the representativeness of forum participants in the population. Nevertheless to gauge the external validity of our results, it is useful to know whether there were systematic differences in attitudinal characteristics between those who chose to participate in the forum and those who chose not to. By and large we find that there were surprisingly few. We report in Table 2 the mean responses for those who did and those who did not eventually attend the forum for each of the deliberative questions under study.

With few exceptions, the attitudes of those who elected to take part in the forum are similar to those who did not. The exceptions are that forum goers are more likely (78 percent) than nonparticipants (59 percent) to think that health should be the first priority of the government, and that participants (76 percent) are more likely than nonparticipants (61 percent) to support services for passenger travel over commercial transport. These differences can be observed within each gender group and do not reflect the skewed gender composition of the participant group—indeed men, although they attended in larger numbers, were somewhat *less* likely to consider health the government’s first priority. Forum participants are somewhat more patient and slightly more willing to be taxed than non-forum goers, although these differences are not statistically significant.

DISCUSSION LEADERS

The discussion leaders were drawn primarily from two sources: government services (notably the National Statistics Institute) and civil society organizations. The *moderadores* constituted a subset of the discussion leaders and hence played both an informational role during the plenary sessions and a discussion leader role during the group deliberations. The leaders were selected to ensure a gender and age balance, but beyond this they cannot be considered representative in any statistical sense of any particular *demographic* grouping. Importantly, however, the choice of discussion leaders was not determined by the researchers but resulted rather from the national political and administrative process that shaped the organization and running of the forum. This involved a multiparty organizing committee with substantial input from the UNDP.

TABLE 2
PRE-FORUM ATTITUDES

<i>Issue # on Form</i>	<i>Issue-Area</i>	<i>Participant Median Response</i>	<i>Participant Mean Response</i>	<i>Nonparticipant Mean Response</i>	<i>Difference in Means</i>
2	Health as Priority [0/1]	1 (74)	0.78 (74)	0.59 (114)	-0.2**
3	Clinics (0) or Hospitals (1)	1 (72)	0.58 (72)	0.55 (108)	-0.04
4a	Primary (0) or Secondary Education (1)	0 (76)	0.28 (76)	0.25 (114)	-0.03
4b	Professional (0) or University Education (1)	0 (76)	0.34 (76)	0.26 (111)	-0.08
4c	Advanced (0) or Basic Education (1) ^a				
7a	Commercial (0) or Passenger Travel (1)	1 (75)	0.61 (75)	0.76 (113)	0.15**
7b	Roads (0) or Public Transportation (1)	0 (76)	0.14 (76)	0.12 (113)	-0.03
7c	Villages (0) or Major Centers (1)	0 (75)	0.19 (75)	0.21 (111)	0.02
11a	Consume (0) or Invest (1) Windfalls	0 (72)	0.46 (72)	0.44 (105)	-0.02
11b	Required Return (percentage)?	40% (56)	62% (56)	55% (95)	-7%
12	Taxation on Windfalls (percentage)?	10% (58)	16% (58)	13% (96)	-3%
14c	Prevalence of Corruption (1-5)	3 (73)	2.6 (73)	2.5 (109)	-0.1

* significance at the 90% level; ** significance at the 95% level; *** significance at the 99% level

^aQuestion 4c not asked in survey.

While interviews suggest that the leaders were drawn from all of the major political factions in São Tomé and Príncipe, beyond the fact that prior experience with moderating discussion groups was deemed valuable, the method for selecting the leaders is something of a black box. (In this sense, although the leaders cannot be considered representative of any demographic group, they are representative of the types of leaders that are generated by political processes precisely because they *are* the set of leaders that was generated by this political process.)

We have data on the gender of all leaders, we have age data for most leaders, and we have additional information on the preferences of a subset of nineteen leaders.

TABLE 3
LEADER DEMOGRAPHICS ^a

	<i>Male</i>	<i>Female</i>
Aged 40 or under	6 (5)	5 (5)
Over 40	12 (4)	8 (4)
Unknown	8 (1)	2 (1)

^aAverage number of meetings run by each leader is given in parentheses.

The demographic distribution of the discussion leaders, by age and gender, is given in Table 3. Although there is rough gender parity, men outnumbered women. Most discussion leaders were over forty and there is no correlation between the gender and age distributions of the discussion leaders.

Our information on leader preferences, collected through self-administration of the leader's form, is imperfect on two counts. First, it exists for only nineteen of the forty-one discussion leaders. Data availability is positively correlated with the number of sessions run by discussion leaders, as data are missing especially for those leaders in Príncipe who ran just one session each. These nineteen discussion leaders collectively led 64 percent of the meetings. Second, the data on leader preferences were collected *after* the meetings were conducted (typically within one week after the completion of the decentralized meetings). This means that we cannot discount the possibility that the preferences of the leaders are a result of, rather than a determinant of, the outcomes of the discussions, a point to which we return below.

RANDOMIZATION

The random assignment of leaders to meetings was implemented by the researchers using a random number generator and released typically one day in advance of the meetings. The short lead time was designed to minimize the possibility that leaders could engage in swaps and also to ensure that, insofar as possible, logistical constraints including leader availability were settled before session assignments. Although the sites were visited by teams of typically three to four leaders, the randomization was done at the level of the individual; as a result, the composition of the teams changed from meeting to meeting. The procedure first randomly selected the team leaders (*moderadores*) and then the other leaders (*facilitadores*). While the research team could verify that the assigned randomizations were followed on the island of São Tomé, they could not monitor the process on Príncipe. In two more sites the pro-

ocols were not consistently applied for logistical reasons.³⁵ We report robustness tests that eliminate these cases, which altogether number twenty-four deliberative groups. For the remaining 124 groups whose leaders were verifiably randomly assigned there was an “adherence rate” of 88 percent; that is, in only 12 percent of cases were meetings in fact led by an individual who was not the one assigned to them by the randomization protocol. Leaders on São Tomé island were typically assigned to between five and seven sessions, although eventually fourteen leaders led just one session, one led two sessions, three led three sessions, five led four sessions, seven led five sessions, nine led six sessions, and two led seven sessions.

We check the integrity of the randomization procedure *ex post* by searching for systematic patterns in the allocation of individual leaders to groups of distinct demographic composition. An *F* test that leader fixed effects are all uncorrelated with the gender composition of discussion groups yields an associated *p* value of 31 percent; with the age composition, the *p* value is 67 percent; with the size of groups the *p* value is 69 percent. Hence in none of these cases can we reject the null that the relation between leader assignment and group composition is random. Nor is there any relationship between the age and gender of leaders and the size, age, or gender composition of groups. Finally, since the 12 percent nonadherence rate may be nonrandom, we demonstrate below that our results are robust to an “intention to treat” analysis.

OUTCOMES

Overall, the execution of the National Forum in São Tomé and Príncipe was smooth. Almost all meetings were held on schedule and, to our knowledge, only once (when the leader threatened to summon the police in response to an unruly discussion) were concerns raised that a discussion leader was unduly influencing the content or outcome of a meeting. As designed, the results of the meetings were later aggregated and discussed in a centralized meeting in the capital city that involved the president, ministers of government, and members of the National Assembly. A final forum document was produced and submitted to the National Assembly to serve as an input into development planning. To our knowledge, however, there has been little follow-up to ascertain whether concerns raised during the forum deliberations ultimately affected government policy-making in São Tomé and Príncipe.

³⁵ These two are the first session (the pilot) held in Santa Margarida and an extra final session (the makeup) held in Trindade.

V. RESULTS 1: IDENTIFYING THE EXISTENCE OF LEADER EFFECTS

The random assignment of leaders to discussion groups allows for the identification of leader effects by examining *variation* in the effect of different leaders. Even if it is unknown what would have obtained in the absence of such leader effects, the effects of leaders can be detected though systematic differences in outcomes associated with different leaders. Such variation could result either from variation in the ability or willingness of leaders to influence outcomes or, even if these features are constant (but positive), from variation in the preferences of leaders. If all leaders had the same preferences and if there were no variation in the ability or willingness of leaders to affect outcomes, then we would not be able to identify the impact of leaders (relative to a “no leader” baseline) using the present approach.³⁶

There are two important implications of this reasoning for our results. First, to the extent that as a group the leaders’ views on issues are systematically different from those of the communities, we *underestimate* the impact of leaders on outcomes. Insofar as leaders try to sway discussion outcomes in the direction of their own views, this underestimation is more severe the more uniform are their views and the more uniform are their powers of suasion.

Second, because we rely on variation across leaders, we cannot make statements regarding the degree of bias induced by any *one* leader. It is unclear what a notion of “zero influence” would mean, but insofar as it is meaningful at all, it is always possible that the set of responses reported by a selection of communities by any one leader perfectly reflects the zero influence response and that systematic differences between these and those reported by other leaders reflect the influence of those other leaders only. For this reason we can make more progress in determining, for example, that individuals have different effects from each other and that women produce different outcomes than men, without

³⁶ To illustrate the logic, consider a binary setting and assume that in the absence of manipulation by leaders a given group would select option 1 with probability q . Assume that share α of leaders in fact support option 1. Finally, assume that if a leader does not like the outcome that a group would otherwise select, he has the ability to exercise influence and change the outcome with probability β . This bias parameter, β , is not directly observable. Nonetheless we can show that the R^2 statistic—that is, the share of variance in outcomes explained by the characteristics of leaders—provides rich information on β . In this case the expected R^2 is given by
$$R^2 = \frac{\alpha(1-\alpha)}{\beta\alpha(1-\beta\alpha) + (1-2\beta\alpha)(1-\beta)q - (1-\beta)^2q^2} \beta^2.$$

We can see from this expression that conditional upon some level of heterogeneity across leaders ($\alpha \in (0,1)$), $R^2=0$ if and only if $\beta=0$, and $R^2=1$ if and only if $\beta=1$. In addition, for all values of q , α , β , we can show that $R^2 \leq \beta$. Hence if R^2 of the variation can be explained by leader characteristics, then *at least* share R^2 of the time, leaders can enforce their will on the groups that they lead.

this implying that the men's response is the "true" response and that the women's reflects undue influence or vice versa.

INFLUENCE

In this section we provide the core results of our analysis. We examine whether the outcomes of meetings are determined by the identity of the leaders running them. Our aim is to establish not how or why individuals matter but whether and how much they matter.

We do this by searching for leader fixed effects in the determination of community responses. The dependent variable is the outcome of the discussions; the independent variables capture characteristics of the leaders. Effectively we check for leader-specific patterns in responses: is it the case that whenever leader i leads a discussion, the answer provided by the group on a given issue is more likely to be one way or the other relative to the answers provided by all other groups? If there is such a tendency, is this an effect that is statistically significant or is it something that could have arisen by chance, given the variation that is likely to exist across groups even in the absence of a leadership effect?

We undertake tests to see whether the effects we observe could have arisen by chance for all leaders separately and simultaneously, and we undertake them for each of the twelve deliberative questions discussed during the forum. We emphasize, however, two ways in which these tests for the presence of leader influence are biased toward *underestimating* the presence of leader effects. We have already noted the first: if leader effects work in the same direction, then their common component is not identifiable using this method. Second, our results depend for statistical significance on multiple observations for each leader. In the extreme case of only one observation per leader it would be impossible to identify individual fixed effects, no matter how important such individual effects in fact are. We will return to this issue below.

Table 4 presents the outcomes of the forum discussions for each of the twelve issues under consideration. Column 1 presents the mean response for each issue across all meetings, and column 2 reports the differences between these and the mean responses given to the same questions by individual respondents in the pre-forum survey. In four cases we reject the null that mean survey responses are the same as mean forum outcomes at the 99 percent level. In two of these four cases (issue-areas 4a and 7c), the differences are opposite in direction to what would be expected based on median attitudes. Note that the absence of a difference is not evidence for the absence of leader effects, since individual leaders may exercise influence in opposing directions

TABLE 4
IDENTIFYING INFLUENCE

Issue # on Form	Issue-Areas	Mean Response (All Groups)		Difference between Response and Pre- Forum Attitudes ^c		Baseline ^e			Subsample ^b			Intention to Treat-Response		
		Adj R ²	Prob> F	N	Adj R ²	Prob> F	N	Adj R ²	Prob> F	N	Adj R ²	Prob> F	N	
2	Health as Priority [0/1]	0.70	-0.08	0.14	0.04**	134	0.14	0.04**	114	0.06	0.19	111		
3	Clinics (0) or Hospitals (1)	0.63	0.05	0.36	0.00***	133	0.34	0.00***	109	0.26	0.00***	106		
4a	Primary (0) or Secondary Education (1)	0.63	0.35***	0.18	0.01***	138	0.14	0.04**	115	0.12	0.07*	112		
4b	Professional (0) or University Education (1)	0.10	-0.24***	0.13	0.04**	144	0.22	0.00***	120	0.12	0.06*	117		
4c	Advanced (0) or Basic Education (1)	0.76	.	0.49	0.00***	142	0.5	0.00***	119	0.5	0.00***	116		
7a	Commercial (0) or Passenger Travel (1)	0.63	0.02	0.18	0.02**	131	0.12	0.06*	109	0.05	0.26	106		
7b	Roads (0) or Public Transportation (1)	0.16	0.01	0.02	0.39	132	0.07	0.19	110	0.02	0.36	107		
7c	Villages (0) or Major Centers (1)	0.65	0.46***	0.13	0.07*	128	0.16	0.03**	105	0.17	0.03**	102		
11a	Consume (0) or Invest (1) Windfalls	0.51	0.05	0.22	0.00***	143	0.25	0.00***	120	0.22	0.00***	117		
11b	Required Return (percentage)?	50%	-11%	0.43	0.00***	81	0.47	0.00***	71	0.45	0.00***	70		
12	Taxation on Windfalls (percentage)?	7%	-9%***	0.3	0.00***	126	0.24	0.00***	106	0.15	0.04**	104		
14c	Prevalence of Corruption (1-5)	2.70	0.09	0.38	0.00***	138	0.41	0.00***	115	0.34	0.00***	112		

* significance at the 90% level; ** significance at the 95% level; *** significance at the 99% level; for each model the Adjusted R² is reported from a least squares dummy variable regression using only leader fixed effects, the Prob>F column reports the results of an F test that all leader effects are equal to 0.

^a Full sample of Forum returns

^b Sample restricted to forum meetings for which use of randomization procedure was monitored

^c Sample restricted to forum meetings for which use of randomization procedure was monitored, leader dummies take a value of 1 whenever a leader was assigned to attend a meeting *whether or not the leader in fact attended that meeting*.

that will not produce differences in average outcomes. Similarly, the presence of differences does not provide evidence of leadership bias, as these differences may result from aggregation effects or from impacts of deliberation not due to leader effects.

The results of our tests are provided in columns 3, 4, and 5. These three columns provide our baseline model (column 3) and two robustness checks, one conditioning the sample on the subset of meetings for which we could verify the randomization and one conditioning on this same subset but employing an “intention to treat” analysis designed to address the 12 percent nonadherence rate. In these cases we use information on which leader ought to have led a meeting rather than information on who in fact did lead a meeting. While this approach adds noise to our analysis, it has an advantage over eliminating these cases from analysis if such elimination is nonrandom.³⁷

For each of these three models we provide three statistics. The first is the share of the variance explained by the leader fixed effects. The number reported is the *adjusted* R^2 from a least squares dummy variable regression—a statistic that takes account of the number of leaders when estimating the variance explained. We expect this statistic to take a nonpositive value under the null that the variation in outcomes is not related to the identity of the discussion leaders. Under simple assumptions (see fn. 36), the R^2 can be interpreted as a lower bound on the frequency with which leaders can directly alter outcomes that a group would have otherwise reached in the absence of a leader effect. The second figure we report is the estimated probability that we would observe a distribution of responses across communities as we do under the null that none of the leaders had any systematic effect on answers. The third figure reports the N for the corresponding regression.

The results are powerful and consistent across models. The share of variance explained by leader fixed effects is typically large; in the median case for the baseline model about one-fifth of the variance is explained by leader fixed effects; for one-third of cases more than one-third of all variance can be explained by leader fixed effects. Our statistical tests show that it is extremely unlikely that such outcomes could

³⁷ Our “intention to treat” variable was determined by replacing our leader indicator (in cases where the leader for a meeting did not correspond to one of the assigned leaders for a site) with a leader randomly selected from the pool of leaders that ought to have gone (but did not go) to a particular meeting. In some cases there were no such individuals. This can arise, for example, if due to large turnout more meetings are held at a given venue than originally planned. In such cases in which a leader is moved from one venue to another to lead an unanticipated meeting, the leader is not in fact replacing any other leader. Rather, there is an unanticipated observation in the data.

arise by chance. The estimated probability of a chance occurrence like this is below one-third of one percent for half the questions. For eleven of twelve questions we can reject the null of no true impact at the 90 percent level, and for ten of twelve we can reject the null at the 95 percent level.

The results from the robustness checks are only marginally weaker. In the subsample of monitored meetings we again reject the null of no leader-specific influence for eleven of twelve questions under study at the 90 percent level, and this is the case for half of them at the 99 percent level. For the intention to treat analysis we reject the null of no effect for nine questions at the 90 percent level or above, and for five of these we reject the null at the 99 percent level. In all cases the share of variance explained by the identity of the leaders remains large. Finally, these core results survive the addition of controls both for location and for the identity of the *moderador*.³⁸

These results are unambiguous. Leaders matter profoundly: while 3,500 adults were consulted, the actual outcomes were in fact largely determined by only a handful of individuals. Knowing which member of the country's political elite was randomly selected to lead the discussions provides an extraordinarily powerful indicator of what policies the participants in each group ostensibly supported.

EXOGENOUS CHARACTERISTICS OF LEADERS AND THE DIRECTION OF INFLUENCE

We noted above that it is possible that individuals may have effects on outcomes but, because of the small number of sessions run by each leader, these may not be identified using fixed effects. One way to identify such effects is to examine attributes that are shared across leaders. Examining such features is of interest both for identifying the presence of bias and for gaining a richer understanding of the determinants of bias, which could provide more general lessons for other settings. Can we predict, based on observable characteristics, the *types* of bias leadership effects will cause?

We concentrate on readily observable and exogenous characteristics of leaders—age and gender—and examine the impact of these characteristics on outcomes of group discussions. On these two characteristics we again find compelling evidence of strong leadership effects. Either the age

³⁸ If we add a fixed effect for each location, we find that using an *F* test on the leader dummies we continue to be able to reject the null of no leader effect in nine of twelve cases at the 90 percent level and in seven cases at the 95 percent level. By construction, location variables are independent of leader dummies. *Moderador* dummies are not, however; the introduction of these *moderador* dummies is discussed below.

TABLE 5
GENDER AND AGE EFFECTS

<i>Form</i>	<i>Issue-Area</i>	<i>Effect of Leader's Gender^a</i>		<i>Effect of Leader's Age^b</i>	
		<i>On Group Responses</i>	<i>On Leader's Own Responses</i>	<i>On Group Responses</i>	<i>On Leader's Own Responses</i>
2	Health as Priority [0/1]	-0.06	-0.18	.17**	.29**
3	Clinics (0) or Hospitals (1)	-0.32***	-0.33*	0.07	0.19
4a	Primary (0) or Secondary Education (1)	-0.04	0.07	0.05	0.21
4b	Professional (0) or University Education (1)	0.04	0.16	0	-0.35
4c	Advanced (0) or Basic Education (1)	0.1	0.13	-0.05	-0.11
7a	Commercial (0) or Passenger Travel (1)	-0.02		-.16*	
7b	Roads (0) or Public Transportation (1)	0.13 **	0.08	-.13*	-0.04
7c	Villages (0) or Major Centers (1)	-0.12	-.63***	.17*	0.27
11a	Consume (0) or Invest (1) Windfalls	0.14*	0.19	-0.09	-.42**
11b	Required Return (percentage points)?	-37%**	5%	-35%**	13%
12	Taxation on Windfalls (percentage points)?	2% *	4%*	-1%	-3%
14c	Prevalence of Corruption (1-5)	0.1	-0.7	-0.7***	-0.5

* significance at the 90% level; ** significance at the 95% level; *** significance at the 99% level

^a Gender recorded as 1 = female, 0 = male. Thus a positive score indicates that women leaders' responses, or the responses of groups led by women, were higher on the indicator than those of male leaders.

^b Variable takes value of 1 if leader is aged over 40. Thus a positive score indicates that older leaders' responses, or the responses of groups led by older leaders, were higher on the indicator than those of younger leaders.

or the gender of discussion leaders is a powerful predictor of the outcomes of discussions in all but three of our twelve subjects of deliberation.

These results are reported in Table 5. In five of twelve questions, there was a significant difference in the responses of groups run by women relative to those run by men. Forum groups led by women leaders were more likely to prioritize investment in local health clinics over hospitals. They were more likely to opt for transportation services rather than better roads; they were more likely to opt for investment of windfalls over consumption and they were more likely to accept higher taxation levels of windfall earnings.

As described above, our data on the attitudes of the leaders themselves are imperfect; they are incomplete and may be contaminated by the effects of the forum. Nonetheless we note that the patterns in this data are consistent with the patterns we find in the forum results them-

selves. In all but one of these areas in which there were significant differences between responses from women-led groups, these differences reflected, qualitatively, the recorded differences in attitudes between the male and female leaders. Women leaders prioritize local health clinics over hospitals and support taxes on windfall earnings at rates that are approximately twice as high as those of the male leaders. These differences are reflected in the discussions of the groups they led. The average response on the tax question within groups led by men exactly reflected the male leaders' (weighted) average optimal tax rate (about 6 percent). Groups led by women reported more support for higher taxes but not to the same extent, favoring a rate of 8 percent while the women leaders had a (weighted) average optimal tax rate of close to 11 percent.³⁹ For the other questions in which we found significant differences by gender across groups, the differences in our measures of leader preferences are not statistically significant. Leader discount rates is the one case in which there is an inconsistent sign pattern; this is, however, the question for which we have the poorest information, with only thirteen discussion leaders providing a response to this question. We identify only one difference between the attitudes of female and male leaders that is not reflected in the group outcomes: relative to men, the women leaders placed, on average, a higher priority on local roads as opposed to national routes. While this difference is reflected qualitatively in the community responses (with a 12 percent difference between responses), the difference is not significant at conventional levels (it has an associated *t*-statistic of 1.4).

There were significant differences between the outcomes of discussions run by older leaders on six of the twelve deliberative issues. Groups with older discussion leaders were more likely to emphasize health as a national priority; they favored improving conditions for commercial transport over passenger transport and favored improving road quality over improving public transportation services; and they supported investment in regional links over local links. They favored less investment of windfall revenue and lower allocations to governments in the form of taxation. Although somewhat inconsistent with their relative unwillingness to pay taxes, they reported relatively low levels of central government corruption compared with groups led by younger leaders.

Again, in all but one of these, the effects are qualitatively consistent with our estimated differences in leader attitudes based on age. And

³⁹ The weights placed on leaders' preferences to generate these averages are determined by their relative frequency in leading discussions.

again the inconsistent effect lies with the discount rate, a variable for which we have especially poor information on the preferences of discussion leaders.⁴⁰

Although our study is not designed to test hypotheses regarding *why* different discussion leaders have different effects, a number of features nevertheless stand out. The strongest result on gender—that discussion groups led by women support clinics over hospitals—is consistent with the findings by Chattopadhyay and Duflo⁴¹ that women leaders invest more in infrastructure that responds to local needs. The strongest result on age—that groups led by older men are less likely to be critical of corruption—is subject to a number of interpretations, most obviously that older leaders are more likely to have been—or at least to be perceived to have been—involved in the political system and thereby themselves complicit in corruption. Other interpretations may be that older leaders have different notions of what widespread corruption means, either as a result of rising tolerance over time or as a result of having a different historical benchmark with which to compare the present.⁴²

DETERMINANTS OF CONSENSUS

We noted above that one concern regarding deliberative democracy approaches is that, by emphasizing consensus, they in fact lead to an exclusion of dissenting voices. Since the São Tomé and Príncipe discussion forms recorded not just the outcome of discussions but also the degree of agreement, we can address this question in part. We turn to this next.

While we cannot in the context of this study determine the absolute extent to which dissenting voices are excluded in deliberations of this form, we can learn about the degree to which reporting of consensus within groups is related to the identity of discussion leaders. If some leaders are more likely than others to exclude dissent from discussions,

⁴⁰ The tests on the effects of age and gender effectively assume that the observations we examine are all independent. However insofar as the responses of meetings run by a given discussion leader are correlated, this assumption is inappropriate. It is not possible to include fixed effects in our examination of gender and age effects, since these effects are themselves fixed. In order then to isolate these effects, we also implemented a test that focuses narrowly on the effects of gender and age characteristics independent of individual fixed effects by taking the average response to a question over the set of meetings led by a given leader and regressing average responses on the gender and age of the leader. This reduces the number of observations to between thirty and forty for each regression and so provides a difficult environment to identify these effects. We find that for two issues gender effects remain strong and significant at conventional levels and for another two age effects remain strong, accounting for between 10 and 20 percent of the observed variation. These results are available on request from the authors.

⁴¹ Raghendra Chattopadhyay and Esther Duflo, "Women as Policy Makers: Evidence from an India-Wide Randomized Policy Experiment," *Econometrica* 72 (September 2004).

⁴² Although, we note, there is no relationship between the number of older people *attending* the meetings and responses on this issue.

deliberately or not, or ignore dissent once it is voiced, then this will appear in our data as variation in reported consensus. Similarly if some leaders are more skilled than others at generating consensus by managing discussions effectively, this will appear as leader-specific variation. In either case, such variation would indicate that the degree of consensus that is achieved during deliberative democratic practices does not simply reflect the overall design of the deliberative structure and the underlying attitudes of populations; it would indicate, rather, that leadership effects have a causal effect on the degree of (reported) consensus generated by deliberative practices.

This is in fact what we find. Fully two-thirds of the variation in the reported level of consensus achieved during the discussions can be accounted for on the basis of leader-specific fixed effects alone (based on the adjusted R^2 from a least squares dummy variables regression). The chances that such a distribution of consensus scores could occur if leader effects did not matter is essentially 0.

The variation in reported consensus is also systematically related to demographic attributes of the discussion leaders. Regressing the mean reported consensus score for each leader across meetings that they led (and weighting by the number of sessions they led) on age and gender attributes, we find that both of these have significant effects: meetings run by older leaders are more likely to report high levels of consensus, as are meetings run by women (these relations are significant at the 95 percent level). Strikingly, while the relations between leader age and gender and reported consensus is strong, the aggregate level of consensus is unrelated to the age and gender distribution (or even the raw number) of the discussion participants.

VI. RESULTS 2: TOWARD AN EXPLANATION OF LEADERSHIP EFFECTS

We have shown that leadership effects are substantively large, that they matter for outcomes and the level of reported consensus over those outcomes, and, furthermore, that these effects are systematically related to exogenous characteristics of leaders. While this evidence suggests that the outcomes that obtain largely depend on *who* the leaders are, it says little about *how* or *when* these effects operate.

The experimental method we used was not designed to answer such *how* or *when* questions. Understanding these questions is however of considerable importance for both interpreting and responding to the results we provide. It is possible to exploit some features of the study design, as well as supplementary data drawn from separate surveys of

TABLE 6
LEADER PREFERENCES AND THE DIRECTION OF INFLUENCE

<i>Issue #</i>			<i>Mean</i>	<i>Mean</i>	
<i>on</i>			<i>Group</i>	<i>Leader</i>	
<i>Form</i>	<i>Issue-Area</i>	<i>N</i> ^a	<i>Response</i>	<i>Response</i>	<i>Correlation</i>
2	Health as Priority [0/1]	94	0.70	0.86	0.22
3	Clinics (0) or Hospitals (1)	82	0.63	0.82	0.39
4a	Primary (0) or Secondary Education (1)	82	0.63	0.39	0.35
4b	Professional (0) or University Education (1)	80	0.10	0.31	0.31
4c	Advanced (0) or Basic Education (1)	82	0.76	0.62	0.58
7a	Commercial (0) or Passenger Travel (1)	89	0.63	1.00	(no variation)
7b	Roads (0) or Public Transportation (1)	87	0.16	0.20	0
7c	Villages (0) or Major Centers (1)	78	0.65	0.49	0.2
11a	Consume (0) or Invest (1) Windfalls	94	0.51	0.70	0.08
11b	Required Return (percentage)?	61	50%	32%	0.5
12	Taxation on Windfalls (percentage)?	94	7%	9%	0.46
14c	Prevalence of Corruption (1-5)	87	2.7	2.2	0.43

^aFor this table, *N* is the number of meetings for which discussion leader responses are recorded.

discussion leaders and discussion participants, to provide some answers. Even if they are not as methodologically clean as those presented in Section V, they can nonetheless contribute to a richer understanding of the dynamics of the discussions.

LEADER PREFERENCES AND THE DIRECTION OF INFLUENCE

The first feature we address is whether the influence we observe is consistent with the attitudes of discussion leaders; that is, do discussion leaders produce the outcomes they like? After all, influence does not imply intentionality, and even intentional attempts at influence may backfire and produce perverse outcomes from the point of view of the discussion leader.

The question then is nontrivial. To answer it we use data on the expressed preferences of discussion leaders about the questions under discussion. Table 6 reports summary statistics of the reported preferences of leaders. We also report in the table the correlation between individual leader preferences and the outcomes of the discussions in the groups that they led.

The results in Table 6 show that in all but two cases there is a positive, and typically a very large and positive, correlation between the positions held by discussion leaders and those that resulted from the discussions that they led. In one of the cases where there is no positive

correlation, this is because there is no variation in the recorded preferences of the discussion leaders; in the other there is a zero correlation between leader preferences and group outcomes.

This finding is consistent with the claim that leaders influenced discussions in the direction of their preferred outcomes. Unlike the results in Table 4, however, the identification in Table 6 is imperfect: since the attitudes of leaders were recorded *after* the group deliberations, it is possible that this correlation is the result of the outcomes of the deliberations rather than their cause. To distinguish between these explanations for the correlations observed in Table 6, we employ a test based on the following logic. Recall that our estimates of the influence of leaders presented in Table 4 are independent of the attitudes of individual leaders. If the correlation in Table 6 is a result of the influence of discussions on leaders, rather than the other way around, then we should not expect a positive relationship between the magnitude of leaders' influence and the resulting correlation. If, however, the relationship is in the opposite direction, then we would expect a strong positive correlation between the magnitude of influence and the resulting correlation between leader preferences and discussion of outcomes.

This is in fact what we observe. The relationship between the aggregate degree of influence and the subsequent correlation between a leader's preferences and the group outcome is significant at the 99 percent level and, as shown in Figure 1, is substantively large. While some caveats are in order, we interpret this result as strong support for the hypothesis that leader influence is generally in the direction of leader preferences.⁴³

PARSING MECHANISMS

So far our study has established that leadership effects in deliberative settings can be extremely large and that these effects work in the direction of the preferences expressed after the forum by discussion leaders. In this section we discuss briefly a set of rival explanations and provide some evidence from our data, coupled with data from our post-forum

⁴³ A counterargument is the following: plausibly individuals are more likely to change their opinion on some issues than on other issues. For issue-areas in which opinions are relatively fixed, we should expect that none of the variance in responses is explained by fixed effects and that there will be no correlation between leader preferences and group preferences; in those areas in which opinions are more easily swayed, it is *conceivable* that individuals in a group are influenced by some feature not necessarily related to a leader's attitudes (for example, they may be affected simply by a leader's age or gender) and subsequently adopt a position that the leader in turn adopts, swayed by the position of the group (itself a function of some feature of the leader other than her prior preferences). This argument relies on the notion that leaders are different from participants in the sense of being very open to suasion, whereas participants, though *not* open to suasion based on the attitudes of leaders, are very ready to change their views based on other characteristics of leaders. Convolved as this argument is, we are not yet able to discount it.

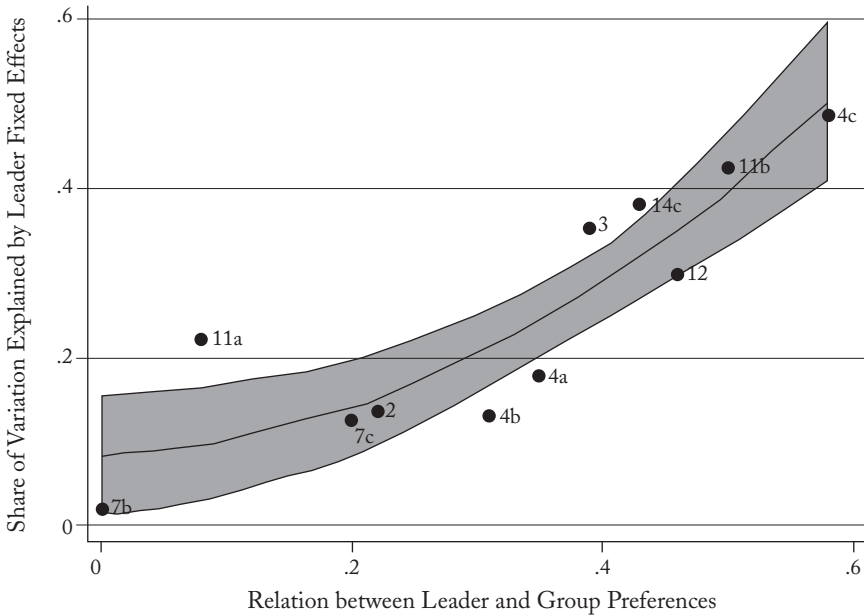


FIGURE 1
CONSISTENCY OF PREFERENCES AND INFLUENCE^a

^a The figure shows the relation between the share of the variation explained by discussion leader fixed effects (from Table 4) and the correlation between leader and group responses (from Table 7). The markers on the datapoints show the corresponding question number on the leader's form (for a mapping, see Table 1). The fitted curve is from a fractional polynomial regression (1 degree); 95 percent confidence intervals are marked with shading.

survey, to support one set of explanations over another.

There are multiple possible explanations for why and how leaders exert influence on outcomes. We highlight four.

—The *false relay* mechanism. One possible explanation is that the leaders of forums simply *misrepresented* the outcomes of the discussions that they led. The São Tomé and Príncipe design is somewhat vulnerable to this charge to the extent that the discussion leaders were also responsible for reporting the outcomes of discussions.

—The *participant self-censorship* mechanism. It is also possible that participants respond differently to different types of leaders—censoring themselves with some more than with others or according a different status to signals received from some rather than others. This mechanism depends on the anticipatory responses of discussion participants.

—The *suasion* mechanism. Plausibly, leader effects work through leader suasion. This explanation and related ones emphasize the active intervention of leaders in discussions that can range from direct attempts at

persuasion to more indirect effects resulting from the ways leaders steer a conversation one way or the other and recognize one set of views over another, or simply emit cues, whether intentionally or not.⁴⁴

—The *information* mechanism. A final mechanism that we consider is that leaders may affect outcomes because of the particular information they bring to bear on the discussion. Insofar as leaders are perceived as elites, information that they bring to each of the discussion groups may be accorded authority by participants (whether or not this information supports any attempts at suasion on the part of leaders) and thereby influence outcomes.

These four mechanisms cover many of the prominent explanations for how leaders plausibly affect outcomes. We emphasize, however, that these four are neither exhaustive nor mutually exclusive; indeed in practice a number of them are likely to operate simultaneously and indeed may even interact.⁴⁵

Although difficult to parse, the differences between these underlying mechanisms are important. Some, such as the suasion or the information mechanism, may produce a change in participant attitudes, albeit possibly one that is unintended, whereas others do not. The reported views may faithfully represent the views of participants, *ex post*, in the former case but not in the latter case.⁴⁶ Even when attitudes change, the normative status of attitude changes due to new information and changes due to suasion by leaders is clearly different. Although our study was not designed specifically to parse these mechanisms, there are features of the design that we can draw on to begin teasing out at least some of them.

Let us begin with the information mechanism. Recall that in the design of the forum not only were the leaders for the individual discussions randomly assigned but so too were the leaders of the plenary sessions, the *moderadores*. The *moderadores* played a key information role, describing the sources of oil wealth and the government's fiscal policies. We can add *moderador* fixed effects to our previous models, since

⁴⁴ For an experimental study on cue taking, see James H. Kuklinsky and Norman L. Hurley, "On Hearing and Interpreting of Political Messages: A Cautionary Tale of Citizen Cue-Taking," *Journal of Politics* 56, no. 3 (1994).

⁴⁵ In a given instance it may not be possible to determine whether a particular conclusion provided by a leader that does not correspond to the opinions of individuals is a result of deliberate misrepresentation or self-censorship; both may be in operation. Persuasion and information provision may in practice be inseparable. And there is much scope for interactions: it is possible that that self-censorship by one section of a discussion group will lead to a preponderance of arguments by a second section that leads to a third section being *persuaded* by the arguments of the second section.

⁴⁶ It is of course possible for leader influence to alter the preferences of participants during the course of discussions but for this influence to be very short lived with a near-immediate reversion to prior attitudes subsequent to the discussions.

TABLE 7
MODERADOR AND LEADER INFLUENCE^a

Issue # on Form	Dependent Variable	Model 1		Model 2		N
		F-Test: All Moderador Effects=0	F-Test: All Moderador Effects=0	F-Test: All Leader Effects=0		
2	Health as priority [0/1]	0.13	0.91	0.22		129
3	Clinics (0) or Hospitals (1)	0.13	1	0.00***		128
4a	Primary (0) or Secondary Education (1)	0.04**	0.46	0.06*		133
4b	Professional (0) or University Education (1)	0.1*	0.16	0.09*		139
4c	Advanced (0) or Basic Education (1)	0.27	0.95	0.00***		137
7a	Commercial (0) or Passenger Travel (1)	0.16	0.41	0.04**		126
7b	Roads (0) or Public Transportation (1)	0.57	0.57	0.3		128
7c	Villages (0) or Major Centers (1)	0.91	0.83	0.13		123
11a	Consume (0) or Invest (1) Windfalls	0.01***	0.64	0.06*		138
11b	Required Return (percentage)?	0.64	0.25	0.00***		79
12	Taxation on Windfalls (percentage)?	0.01***	0.69	0.01***		122
14c	Prevalence of Corruption (1-5)	0.05**	0.48	0.00***		133

^aEach row in this table represents two regressions. The first regression includes fixed effects for *Moderadores* only, the second regression includes fixed effects for both *Moderadores* and for Leaders. The *N* of the regressions (the same in each case) is provided in the final column.

the *moderadores* were also randomly assigned and provided information to all the participants in the location, not just to their own group after the plenary session. If information transmission was an important mechanism for influencing discussion outcomes, then the information presented by the *moderador* should be disseminated to all the groups in the same location even after controlling for leader fixed effects. In addition, any effects that are identified in this way cannot be attributed to the other mechanisms (false relay, self-censorship, or suasion).

In Table 7 we first report results from a test, for each question, of the null hypothesis that there is no correlation between the identity of the *moderador* and the outcome of the discussions. We then present results of the same test but controlling for the identity of the discussion leader in each case. Introducing fixed effects for the discussion leaders is important since, although the *moderadores* were randomly assigned, they also played the role of discussion leaders. The *moderador* fixed effects are therefore not independent of the leader fixed effects. Moreover, examining the leader fixed effects after controlling for *moderador* fixed effects provides a robustness check on our previous results.

The model that ignores leader effects (model 1) gives strong evidence of *moderador* effects for two issue-areas—those two most closely linked to the topics covered by the *moderadores*—and weaker evidence in three further cases. However, these effects disappear once we control for the discussion leader effects (model 2); once we know who the leaders were, no additional information is gleaned from knowledge of who the *moderador* was. The second test reported for model 2 shows that although controlling for *moderador* fixed effects introduces multicollinearity and greatly reduces our degrees of freedom, we nonetheless continue to find strong evidence for discussion leader effects: in nine of twelve cases we cannot reject the null of no leader bias at conventional levels. These results confirm that discussion leaders matter (and do so independently of the information effects resulting from the plenary meetings). They suggest, in addition, that the differential effects of leaders are likely not driven by the differential information that they communicate to citizens.

Next consider the suasion and information mechanisms on the one hand and the false relay and self-censorship mechanisms on the other. To distinguish between these pairs of mechanisms, we need information on the extent to which individual attitudes changed as a result of the forum. Unfortunately because we do not have information that can link individual survey respondents to individual forum meetings, we cannot directly link ex post positions of respondents to the reported conclusions of the forum meetings in which they participated. Further, even if leaders had had strong impacts on outcomes, leader effects will not necessarily produce changes in the *average* preferences of respondents; this is because different leaders plausibly had effects in different directions. Finally, some differences between pre-forum attitudes and forum returns may not be due to influence but may be due more simply to aggregation. If, for example, discussions simply return the median prior opinion of group members, then the average return of group responses will differ from average individual attitudes.

The strategy we employ then is to search for cases in which there are large differences between prior preferences and subsequent outcomes that cannot be explained by aggregation effects. We do this by making use of the fact that, as reported in Table 4, for four of the eleven questions for which we have data the forum produced outcomes that were different *on average* from the attitudes reported by individuals before the forum. For two of these (4a and 7c) the differences were in the opposite direction to what would be expected for median-based aggregation. In each of these cases, before the forum the median individual

level response was 0 and the average response was below one-quarter, while the median forum return was 1 and the average forum return was approximately two-thirds.

The hypothesis we examine is that such “forum effects” work in part though leader suasion or information provision. According to this hypothesis, we expect these effects to have an impact on individuals’ expressed preferences beyond the duration of the meetings.⁴⁷ Our null hypothesis is that these forum effects, when they obtain, are uniquely the result of participant censoring or of misrepresentation by leaders of the outcomes of discussion and not due either to leader suasion or to other ways in which the deliberations affect individual attitudes. Under this null we do not expect a change in post-forum attitudes *even though* the results of the forum differed greatly from pre-forum attitudes.

We report the results of these tests in Table 8. We find in columns 3–6 of the table that despite large differences between the outcome of the forum and pre-forum stated preferences, stated preferences before and after the forum change very little. For one of the two questions under examination the change, though small, was in the *opposite* direction to that implied by the hypothesis.⁴⁸

These results are suggestive. For those issue-areas for which we can be most certain that there was an effect of the forum *in a given direction*, we cannot reject the null that the effect was due entirely to self-censorship or misrepresentation of discussions.⁴⁹ Of the two, the latter—the false-relay mechanism—appears the less likely in this context, since leaders entered results in a public setting after confirming the entry with the discussion groups and, in doing so, were often assisted by members of the groups. False relay would imply a particular malevolence on the part of leaders that seems insupportable. But while it appears unlikely, we cannot rule it out under the present research design.⁵⁰

⁴⁷ Note that leader suasion (direct or indirect) is a sufficient but not necessary condition for changes in opinions.

⁴⁸ Although not reported here, no significant changes in any given direction were observed for any of the questions, with one exception: tax rates. Average preferred tax rates declined from 16 percent to 11 percent. Median rates for this issue-area were constant over time at 10 percent.

⁴⁹ We note, however, a caveat on the interpretation of our results. Our interpretation assumes that individuals do not systematically misrepresent their views when responding to surveys at the individual level. If instead respondents misrepresented their views either because of a *consistency bias*—with respondents wishing to provide the same answer to the question that they had previously provided, even though their positions may have changed, in order to appear to be consistent—or another audience effect on the part of respondents—with respondents not willing to provide a given answer in private even though they feel comfortable expressing the attitude in a public setting—then we would fail to reject the null for these reasons. While neither appears a priori plausible for the questions at hand, we recognize that we cannot rule out these possibilities.

⁵⁰ In further experimental work, however, such an effect could be more satisfactorily identified by distinguishing between the *moderating* and *reporting* functions of discussion leaders.

TABLE 8
LEADER INFLUENCE OR PARTICIPANT SELF-CENSORSHIP?

Issue # on Form	Issue-Area	Mean Individual Response			Median Individual Response		
		Expressed by Attendees before Forum Responses	Expressed by Attendees after Forum	Difference (Before - After)	Expressed by Attendees before Forum	Expressed by Attendees after Forum	Difference (Before - After)
4a	Education Priorities: Primary (0) or Secondary (1)	0.63	0.28	0.33	0	0	0
7c	Transport: Improve Links between Villages (0) or between Major Centers (1)?	0.65	0.19	0.16	0	0	0

THE CONDITIONS FOR INFLUENCE

We have found that when leaders have influence on the expressed results, that influence works in the direction of their own preferences. In addition, mechanisms such as misrepresentation or self-censorship, rather than suasion or learning, appear to underpin this effect. These results provide information on *how* influence occurs but not on *when* influence is likely to arise. As Table 4 shows, the estimated aggregate degree of influence varies enormously across issue-areas; from only 2 percent to as much as 49 percent.⁵¹ This variation may provide a handle for identifying when influence will be more or less significant.

A natural hypothesis is that the degree of influence will vary with the salience of the question for the forum participants. Although we have no direct measures of salience, we can generate a measure of participant ambivalence based on pre-forum responses to the survey questions. In order to elicit marginal preferences, no “don’t know” or “indifferent” option was provided on the pre-forum survey. In cases where respondents did not provide an answer, the enumerators marked the responses as missing. Although we recognize that there are other possible reasons for missing data, for the purposes of this analysis we will use missingness as a proxy for low issue salience. The most abstract questions, on preferred discount and tax rates, exhibit the highest fraction of missing responses (20 percent and 19 percent, respectively). Locally salient questions such as education priorities and public transport have less than 2 percent missing answers. The fact that the most politically sensitive question on national corruption has only 4 percent missing responses suggests that missing answers reflect ambivalence or uncertainty, rather than a refusal to answer.

When we examine the relationship between the proportion of missing responses in the survey and the aggregate degree of influence, we find a very strong positive relationship (see Figure 2). In a univariate regression, the fraction of missing answers is a highly significant determinant of the degree of estimated leader influence and accounts for 47 percent of the variation in influence across issue-areas. The point estimate implies that a 1 percentage point increase in the share of missing answers translates into a 1.2 percentage point increase in this measure of the degree of influence.

That influence appears to be weaker when there is less uncertainty in the population is a good sign for advocates of deliberative democracy, indicating greater resilience on the part of citizens on matters about

⁵¹ Note again that for situations like that described in fn. 36, the R^2 is simply a lower bound on leader influence.

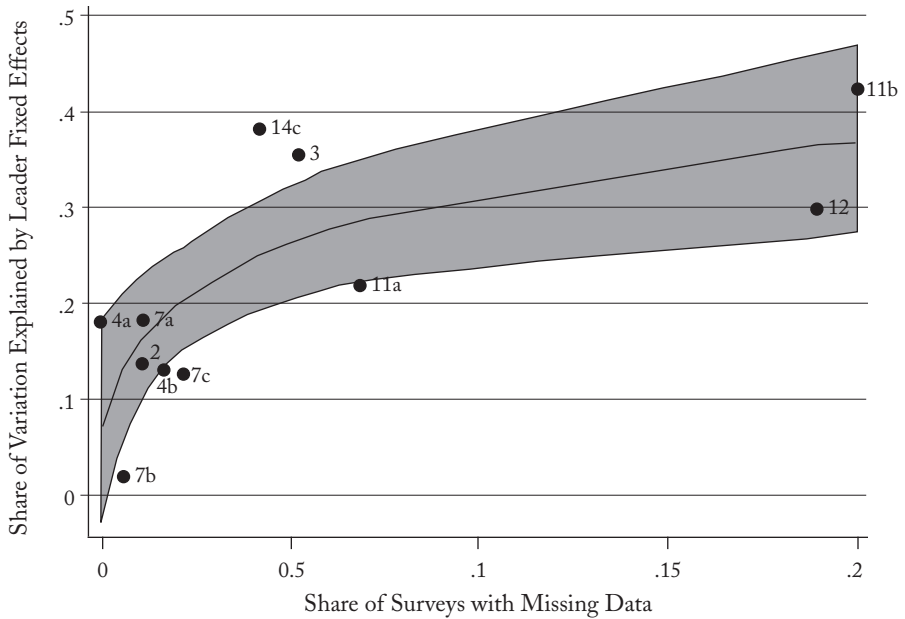


FIGURE 2
AMBIVALENCE AND INFLUENCE^a

^a This figure shows the relation between share of surveys returned with missing data on each question and the share of the variation explained by discussion leader fixed effects (from Table 4). The markers on the datapoints show the corresponding question number on the leader's form (for a mapping, see Table 1). The fitted curve is from a fractional polynomial regression (1 degree); 95 percent confidence intervals are marked with shading.

which they feel more strongly. It also suggests a caution, however. A deliberative approach may be particularly attractive when opinions are not very firm, yet our results suggest that this is precisely when the influence of elites is the greatest.

VII. EXTERNAL VALIDITY

The strength of our approach—a microstudy of a single political event—is that it provides us with a very rich analysis of a particular process and permits us to isolate causality and explore mechanisms. The corresponding weakness concerns the external validity of our results: what do they tell us about such dynamics in other places and in other contexts? In particular, concerns may be raised about the particularity of the *process* (a national exercise in deliberative democracy), the *context* (a very low-income economy possibly on the verge an oil boom),

the *actors* (the forum leaders and the participant populations), and the *location* (a small island state).

Participatory practices can be designed in many ways, a number of which seek to minimize the role of leaders; further, effects that are found in one may not extend to all. The process we examine, however, in many ways offers a best case for participatory practices. The consultation was open, it was domestically driven, there is no evidence that it was captured by any particular political constituency or to pursue any particular agenda, and the role of the leaders in the process was minimal by design (if not in effect). It was conducted in the context of unusually high levels of transparency by a West African government on the management of the economy. These considerations suggest that if such bias is evident in this best of cases, it may plausibly be at least as severe or more so in others.

We think it unlikely that the oil context is determinative of our results. In fact, many of the questions that were asked and for which we found strong effects are questions of very general relevance to oil-producing and non-oil-producing countries alike. Only one of the questions referred to the prospects of future oil windfalls, which suggests that the results are likely to matter more broadly.

With regard to the particular actors, there are two features to consider. First, those who self-selected to become participants in the forum may be precisely those who are more acquiescent to authority; that is, those who turn up may be those least likely to have strong positions and to defend them. In this case the effects found here might not extend to populations that do not participate so willingly in such processes. We suggest two responses to this concern. The first is that although such a selection effect is plausible (as indeed is a selection process with the opposite effects), we have found no evidence to support it: our survey data find few systematic differences between those who did and those who did not elect to attend. The second response is that self-selection is a relatively general, if not universal, feature of political action; hence even if inferences were limited to contexts where individuals self-select, a large range of political processes would still be covered. With respect to the selection of the leaders, we emphasize that the leaders we examine, like those studied in Chattopadhyay and Duflo,⁵² are not in any way a representative sample of their genders or their cohorts. Hence, in interpreting our results on the demographic characteristics of leaders, we make no general claims about the differential impacts of women

⁵² Chattopadhyay and Duflo (fn. 41).

versus men as leaders or about older versus younger leaders. Rather, the leaders represent a population of leaders selected through a particular political process, in this case, a highly inclusive process. The strength of our study lies at a more general level in its ability to identify the presence and extent of leader effects.

Finally, with respect to the location, there may be features specific to the culture of São Tomé and Príncipe that drive our results. This would be the case if it could be established, for example, that the population of the country is more acquiescent than other populations. We have no evidence to support or counter this claim. The appropriate response to such arguments is to replicate our tests broadly in locations that vary in ways that might systematically affect the extent of leadership bias. Such an approach can help move us from establishing *that* influence can be great to understanding, along the lines we pursued in the last section, *why* leaders matter more in some places than in others.

VIII. CONCLUSION

Using a method of random assignment of discussion leaders to discussion groups in a unique nationwide exercise in deliberative democracy, we are able to isolate the impact of leaders on outcomes of political discussions. We find systematic and strong leader effects, related both to the preferences of leaders and to the demographic characteristics of leaders. Participatory forms of political decision making appear to be considerably more vulnerable to manipulation than advocates realize.

We have also provided some results on how and why these effects occur. These results are methodologically less clean, and for each of them we provide caveats that are important for interpreting them. They suggest, however, that influence, when it occurs, does not occur through the ability of leaders to alter the views of participants, whether directly through persuasion or the provision of information or indirectly by shaping the structure of discussions; rather, it occurs through forms of censorship or self-censorship that result in the voicing and recording of opinions that coincide more closely with the leaders' own views. The shortcoming of the deliberative process is twofold: on the one hand, it did not lead to significant *change* in people's preferences (as expressed privately in the post-forum surveys); on the other hand, the preferences recorded in the deliberative meetings to a large degree reflect the preferences of discussion leaders, not participants.

Our results have important implications for advocates of deliberative democracy and of participatory political processes more generally.

We noted at the outset that participatory deliberative approaches are favored for substantive, normative, and instrumental reasons. To some extent the findings of our study provide discouraging news for those favoring deliberative democracy because of the substantive or normative returns. From a substantive point of view the lessons from São Tomé and Príncipe provide an important warning: what you get out may be largely determined by what you put in. Without an understanding of the extent to which outcomes simply reflect the preferences of discussion organizers, it is easy to overestimate the amount of new information generated by participatory approaches. From a normative point of view there is a similar warning: while there may always be an intrinsic benefit to communities coming together to discuss issues, insofar as these discussions are organized centrally, hierarchical relations may continue to be the primary determinants of the course of those discussions. Our results say less about the instrumental benefits of deliberative democracy. Even if discussions are largely influenced by leaders, this influence may pass unobserved, as it did in São Tomé and Príncipe, and may not reduce the extent to which citizens feel ownership over outcomes that follow from such consultations. (Thus, individuals may not feel that their position prevailed but they may nonetheless feel that the views of people in general did.) There is clearly a potential for failed expectations, however, if the publicly reported outcomes of the discussions are at odds with the preferences people actually hold. This risk is perhaps attenuated if the mechanism at work is genuine persuasion by discussion leaders, but even then this attenuation will depend on whether the deliberative process involves a whole population or simply a sample thereof, as advocated in the idea of a deliberative poll.⁵³

Do our results discredit participatory approaches? We argue that they do not, for two reasons. The first is that institutions for participatory politics are still evolving and new designs may be able to prevent or correct for the leader effects identified here. An obvious lesson for proponents of deliberative approaches is that since the idiosyncrasies of organizers can have such very strong effects on discussion outcomes, deliberative practices should be designed to limit these insofar as possible. This could perhaps be achieved by employing multiple leaders with diverse views and backgrounds to moderate sessions in a given community either jointly or sequentially. Another, equally important lesson, in our view, is that designs for deliberation should build in methods for identifying when and to what extent features of the design

⁵³ Ackerman and Fishkin (fn. 1).

of a consultative process determine the outcomes of the process. This should be done through simple randomization procedures with designs that include methods for collecting preferences in private before and after deliberation has taken place. The São Tomé and Príncipe case demonstrates that building in such features is relatively easy and highly informative. This is especially important if discussion outcomes are to be used as inputs into a political process. Such designs could in principle reduce leadership effects *ex ante*. In cases where systematic biases can be identified *ex post* they can in part be accounted for: although statistical techniques cannot correct biases (it is meaningless to conceive of what would have obtained had the leader no characteristics at all), they can be used to rescale outcomes to allow for more meaningful comparisons between results from different meetings.

The second reason is that although our findings directly challenge the claims of participatory approaches, our design does not allow us to compare outcomes under a deliberative procedure with the outcomes that would obtain under other procedures. Even though leaders may play a surprisingly dominant role in formally egalitarian settings, benefits from participation may still obtain relative to other processes. While our study has shown that the absolute merits of deliberation may fall short of expectations, an answer to the larger question of optimal institutional design requires a different approach, one that systematically varies the institutions that can be used to achieve the normative, substantive, and instrumental benefits attributed to participatory approaches.