Can Religion Bridge Ethnic Divides?

A Field Experiment on Leadership and Collective Action

Gabriel Granato[†] Elie Kabue Ngindu [‡]

November 17, 2026

JOB MARKET PAPER.

Please click here for the latest version.

Abstract

Ethnic fragmentation is often associated with worse economic development. In one prominent mechanism, free-riding in heterogeneous societies undermines public good provision. This paper examines whether religious and ethnic leaders can sustain cooperation in a voluntary public-works program in an ethnically diverse city in the D.R. Congo. Residents are randomly assigned to ethnically homogeneous or mixed work teams that either include an ethnic chief, a pastor, or no leader at all. Chiefs increase work by 6.0 percentage points in homogeneous teams but not in mixed ones, while pastors increase work by 5.8 points in mixed teams but not in homogeneous ones. Heterogeneity by the strength of ethnic attachment explains this reversal. Chiefs still mobilize their coethnics on mixed teams, and on homogeneous teams their effect is concentrated among teams of residents who report a strong ethnic identity. Pastors do increase work on homogeneous teams whose members have weak ethnic attachments, and their effect is largest on mixed teams whose members hold strong but competing ethnic identities. A cross-randomized public disclosure intervention provides further evidence that the chief mobilizes labor contributions through social enforcement while the pastor uses internal incentives. The paper provides the first experimental evidence on how religious leaders can sustain voluntary public-goods provision in fragmented societies.

[*]Gabriel is grateful to his advisors, Ernesto Dal Bó, Frederico Finan, Edward Miguel, and Jonathan Weigel, for their guidance and support. We also thank colleagues in the Business and Public Policy group at Berkeley Haas, the Development Economics group at UC Berkeley, and the Network for a New Political Economy (N2PE), as well as participants in the BPP Student Seminar, Oliver E. Williamson Seminar on Institutional Analysis, Development Lunch Series, Development Economics Seminar, and ASREC conferences. We are grateful to the many other colleagues and friends whose comments, questions, and encouragement have helped shape this project. We thank ODEKA ASBL, the Mairie de Kananga, and our field team for their collaboration and dedication, and the residents of Kananga who participated in this research. We gratefully acknowledge financial support from J-PAL's Governance Initiative; the Center for Effective Global Action (CEGA); the Berkeley Center for Workplace Culture and Innovation; the Institute for Business Innovation and the Williamson PhD Fellowship fund; the Berkeley Economy and Society Initiative (BESI) and N2PE; the Haas School of Business through a departmental seed grant; the Clausen Center for International Business and Policy; the Center for Governance and Markets at the University of Pittsburgh; and the Andrew and Mary Thompson Rocca Fellowship at UC Berkeley's Center for African Studies. This research is funded by The Weiss Fund for Research in Development Economics at the University of Chicago. This document is an output from the Centre for Economic Policy Research (CEPR) research initiative "Reducing Conflict and Improving Performance in the Economy" (ReCIPE), a programme funded by the Foreign, Commonwealth & Development Office (FCDO). The views expressed are not necessarily those of FCDO. This study was approved by the University of California, Berkeley Committee for Protection of Human Subjects (2025-05-18574) and was preregistered in the AEA RCT Registry as AEARCTR-0016364.

[†]University of California Berkeley, gabrielgranato@berkeley.edu

[‡]Organisation d'Études Économiques au Kasai (ODEKA), eliekabue@gmail.com

1 Introduction

Ethnically diverse societies have slower economic growth and worse public goods than homogeneous ones (Easterly and Levine, 1997; Alesina and La Ferrara, 2005). One important channel, particularly in Africa, is the failure of voluntary provision: local public goods depend on contributions of labor, money, and materials made outside the formal tax system (Olken and Singhal, 2011). Voluntary contributions are lower in diverse communities because they struggle to sanction free-riders (Miguel and Gugerty, 2005; Habyarimana et al., 2007). In principle, local leaders could solve this enforcement problem (Ostrom, 1990; Kosfeld and Rustagi, 2015). But the authority of chiefs — the leaders best placed to do so — is also ethnically bounded (Acemoglu et al., 2014). Who, then, can bridge ethnic divides?

Religious leaders are natural candidates. Sub-Saharan Africa is now home to more Christians than any other region in the world, and religious leaders play important social roles — from conflict resolution to job training to service provision (Freeman 2012, WCD 2025). Religious leaders also differ from chiefs in two crucial ways. First, their scope of authority extends beyond the ethnic group (McCauley, 2014), and past work has documented a correlation between major religions and cooperation (Henrich et al., 2010). Second, rather than relying on social sanctions, religious leaders can promote internal incentives (e.g., fear of divine judgment), a source of enforcement that does not require an audience of peers (Shariff and Norenzayan, 2007). Yet there is little evidence on whether religious leaders can sustain cooperation across ethnic lines.

We present experimental evidence on the ability of ethnic and religious authorities to promote collective action in a citywide voluntary public-works program in the Democratic Republic of the Congo (DRC). As in many low-income settings, public good provision depends on voluntary labor that residents contribute through *salongo*, a community-work institution, to repair roads and drains, remove waste, and fight erosion. In villages, customary chiefs coordinate this work, while in cities, avenue chiefs play this role (Balán et al., 2022; Hoffmann et al., 2022). In both cases, chiefs' authority depends on ethnic ties, which poses a problem in large multi-ethnic cities like Kananga (pop. 1.8 million), the setting for this study. By contrast, Christian churches are ubiquitous in Kananga, and their congregations typically cut across ethnic cleavages. But can pastors promote costly salongo work in heterogeneous communities?

We partnered with the Kananga city government during a campaign to boost salongo participation. The campaign organized volunteers into work teams and tracked their daily contributions, documenting approximately 70,000 hours of labor and over 800 completed local infrastructure improvements. We embedded an experiment in this campaign, randomly assigning 1,040 residents to 336 work teams across 28 neighborhoods. Each team included either the local avenue chief, a local pastor, or none of these leaders (without any special role for or mention of the status of the leader). We cross-randomized two features of work teams. First, teams were either ethnically homogeneous — and coethnics of the chief — or mixed, a variation we leverage to study the *scope of authority* of leaders. Second, we disclosed to teammates either each member's work contributions or only the team total, which allows us to study the *source of enforcement* on which leaders

can draw. Because the same chief and pastor serve on several teams within a neighborhood, the analysis can hold the leader fixed and vary the team setting in which they operate.

Theory gives a clear prediction for chiefs and an ambiguous one for pastors. Chiefs should be more effective in homogeneous teams, where shared ethnic ties make social sanctions more consequential (Miguel and Gugerty, 2005). For pastors, one view is that their authority crosses ethnic lines, so they should be equally effective in both team types. Another view emphasizes that group composition shapes which identity — ethnic or religious — guides behavior (Akerlof and Kranton, 2000). In homogeneous teams, ethnicity may be so relevant for organizing cooperation that religious appeals fall flat. In mixed teams, religion may be the one identity the group shares, creating scope for the pastor's authority. Our preregistration reflects this: a positive chief effect in homogeneous teams and no signed prediction for pastors.

Our main result is that leaders do raise work contributions, but in different types of teams. Chiefs increase work by 6.0 percentage points in homogeneous teams ($p = 0.032$) and have no detectable effect in mixed teams (0.4 points, $p = 0.893$). Pastors increase work by 5.8 points in mixed teams ($p = 0.039$) and have no detectable effect in homogeneous teams (1.7 points, $p = 0.484$). This reversal is large and significant: the chief's advantage over the pastor is 9.8 points larger in homogeneous teams than in mixed teams ($p = 0.017$). Without a leader, work levels are the same in both team types: 58 percent of possible days in mixed teams and 57 percent in homogeneous teams ($p = 0.536$). The reversal therefore does not reflect baseline differences in cooperation. Ethnic cohesion alone does not raise contributions. Homogeneous teams may hold latent capacity for collective action, but it takes a chief to activate it. Instead, ethnic cohesion changes who can promote cooperation; while it enables chiefs, it notably prevents pastors from organizing collective action.

What explains the reversal in the effectiveness of leaders by group composition? Treatment effects track the *strength of ethnic attachment*. Start with the chief. In homogeneous teams, the chief effect is concentrated among groups with a strong ethnic identity: 9.7 points for them and indistinguishable from zero for others. In mixed teams, chiefs are still able to mobilize their coethnics by roughly as much as in homogeneous teams, though this estimate is imprecise. This suggests that the authority of the chief does not vanish in diverse teams. It is spread over fewer people who recognize him as their leader. Chief effects are also larger among residents who see the chief as able to sanction and reward, and among those who rate him favorably (such as coethnics). Coethnicity delineates where the chief has authority.

By contrast, the effect of pastors does not depend on residents' ethnic ties to the pastor. It depends on the ethnic cohesion of the team. In homogeneous teams, the average null masks offsetting effects. Pastors, in fact, boost work by 6.8 points in homogeneous teams with weak ethnic attachment. They fail only where strong ethnic attachments bind the whole team tightly. In mixed teams, strong ethnic attachments depress work contributions in the absence of a leader: such teams work 53% of possible days compared to 62% in mixed teams with weaker ethnic attachments. Yet these divided teams are where the pastor's effect is largest: 13.2 points, enough to close the gap. In short, where ethnicity divides most, religion has the most to bridge.

One concern is that these patterns reflect individual traits rather than team composition: perhaps residents with strong attachments to the chief ethnicity are simply less religious. They are not. Moreover, the chief's coethnics barely respond to pastors in homogeneous teams but respond strongly in mixed ones. Put simply, the same type of resident responds differently to the pastor depending on who else is on the team. In sum, ethnic attachment is the source of the chief's authority and the main obstacle to the pastor's.

Having shown whom each leader can mobilize, we turn to their source of enforcement using the cross-randomized disclosure intervention. Although residents were part of small work teams, they were free to work at different times and sites, so teammates often did not know who had contributed. Disclosure varied this visibility gap. Chiefs are thought to rely on social sanctions through ethnic networks, so they should benefit when disclosure reveals who is shirking (Kosfeld and Rustagi, 2015). By contrast, pastors are thought to activate internal incentives, such as fear of divine judgment (Shariff and Norenzayan, 2007). Disclosure should not help them, and it could even hurt if social image concerns crowd out internal motivation.

Both predictions hold. Chiefs raise work by 10.6 points in homogeneous teams when individual contributions are disclosed ($p = 0.002$). Without disclosure, they have no significant effect, even in homogeneous teams, and they are ineffective in mixed teams either way. For pastors, by contrast, disclosure does not help. Their effect in mixed teams does not differ significantly with and without disclosure. The point estimate is larger without disclosure (9.0 points, $p = 0.019$), which offers suggestive evidence that social image concerns crowd out internal motivation. Yet, the stronger takeaway is that public disclosure matters more for chiefs, who rely on social sanctions, than for pastors, who appear to rely on internal incentives.

Three additional pieces of evidence support this interpretation. First, a treatment-blind text classifier sorts speeches into three different categories: calls for attendance and better group performance, requests for material support, and appeals linking divine oversight to solidarity and community work. Chiefs devote most of their speaking time to attendance and performance, emphasizing the social enforcement account of their role. Pastors devote theirs to divine oversight and community work, consistent with cultivating internal incentives. Second, the two leaders encourage work from different residents. If chiefs raise the reputational cost of shirking, their effects should be largest among prior nonparticipants when disclosure exposes them. That is what we find. Pastor effects are similar across prior participation, consistent with appeals that do not need peer enforcers. Third, pastor effects are concentrated among residents high in intrinsic religiosity (Allport and Ross, 1967), those for whom religion is a guiding principle and who should be most responsive to internal incentives.

Could leaders simply be trying harder in the teams where they succeed? We find little support for this explanation after triangulating evidence from several sources. Meeting transcripts show that leaders' discourse and effort are similar across team compositions. Detailed work records show that leaders do not devote unequal effort to bringing teammates to participate. Residents' reports show that differential contact outside meetings and workdays does not track leader effects. Leader influence appears to reflect how audiences

respond, not how leaders behave.

Finally, we ask how leader-induced cooperation shaped residents' attitudes toward one another and toward their leaders. Mirroring the rise in voluntary work, residents on leader-led teams became more willing to help one another financially. This suggests that religion may help build informal insurance networks in multi-ethnic settings where the state provides little social protection (Townsend, 1994; Berman, 2000; Chen, 2010; Auriol et al., 2020). Successful pastors also gained standing in the eyes of their teammates. Successful chiefs did not, likely because they already enjoyed favorable evaluations in homogeneous teams. Religious leaders may thus be able to convert cross-ethnic mobilization into broader social and political influence (Chaney, 2013; Bazzi et al., 2020). Did sustained cooperation also shift attitudes beyond the team? We find no change in beliefs about social or divine accountability, the salience of religious or ethnic identity, or views of the state and other authorities. The effects of cooperation appear to be confined to relationships within the team.

To our knowledge, this is the first field experiment to illustrate the ethnic fault lines that determine which local leaders can mobilize voluntary work. Chiefs are effective when cooperation stays within an ethnic community and contributions are visible. Pastors are effective when cooperation crosses ethnic lines without relying on visibility. Religion can bridge ethnic divides, but struggles if ethnic ties already organize the group. Instead of impeding cooperation, we find that ethnic fragmentation changes who can govern collective action.

1.1 Related Literature

This paper makes three contributions. First, a large literature finds that ethnic fragmentation lowers growth and public good provision (Easterly and Levine, 1997; Alesina and La Ferrara, 2005)[1], in part because diverse communities struggle to enforce social norms against free-riding (Ostrom, 1990; Miguel and Gugerty, 2005; Habyarimana et al., 2007).[2] We illustrate that religious leaders can sustain collective action in fragmented societies. This result also complements work showing that cross-cutting identities can soften ethnic divisions (Miguel, 2004; Sambanis and Shayo, 2013; Bazzi et al., 2019; Depetris-Chauvin et al., 2020).[3] In addition to our focus on leaders, we show that religion has more muted effects in groups already united by ethnic ties.

Second, our field experiment examines two motives for cooperation: social image (Bénabou and Tirole, 2006; Andreoni and Bernheim, 2009) and internal motivation rooted in identity and self-image (Akerlof and Kranton, 2000; Bénabou and Tirole, 2016). Religion can draw on both, through community sanctions (Iannaccone, 1992; Berman, 2000) and belief in divine judgment (Norenzayan, 2013; Purzycki et al., 2016).[4]

[1] Beyond public-good provision, another strand of research studies how ethnic and religious divisions shape discrimination and productivity within firms and organizations (Hjort, 2014; Marx et al., 2021; Ghosh, 2025).

[2] Other explanations emphasize divergent preferences (Alesina et al., 1999), aversion to interethnic interaction (Alesina and La Ferrara, 2000), and between-group inequality (Baldwin and Huber, 2010), though diversity do not necessarily undermine public good provision (Glennerster et al., 2013). Group membership shapes cooperation and norm enforcement (Goette et al., 2006), and reputation can sustain repayment within ethnic groups (Sánchez de la Sierra, 2021).

[3] A related literature tests whether intergroup contact can weaken social divisions (Scacco and Warren, 2018; Rao, 2019; Mousa, 2020; Lowe, 2021; Ghosh et al., 2026). Findings in this literature emphasize that the effects of intergroup contact depend on how interaction is structured and often do not extend much beyond those directly encountered (Lowe, 2025).

[4] For reviews of the economics of religion, see Becker et al. (2024) and Lowes et al. (2026).

Past experiments show that religious interventions change behavior (Bryan et al., 2021; Buccione, 2023) but cannot separate these channels. Cross-randomizing disclosure allows us to make progress. Disclosure helps chiefs, who rely on social enforcement, but not pastors, who appear to rely on internal incentives.[5]

Third, we contribute to research on leadership and informal governance. Prior work studies how leaders sanction free-riding (Kosfeld and Rustagi, 2015), lead by example (Hermalin, 1998; Jack and Recalde, 2015), and coordinate collective action (Boudreau et al., 2025). Chiefs and religious leaders are two key leader types in developing countries (Acemoglu et al., 2014; Baldwin, 2016; Casey et al., 2023; Bazzi et al., 2020). We add to this literature by experimentally comparing both leader types in the same setting and showing the different mechanisms through which these leaders operate.[6] Traditional authority works within ethnic groups, through social accountability. Religious authority works across ethnic groups, through internal incentives. Effective leadership thus depends on the match between the leader's source of authority and the group's composition.

The paper proceeds as follows. Section 2 describes salongo and local leadership in Kananga, and Section 3 presents the experimental design and the data. Section 4 asks whom chiefs and pastors can mobilize; Section 5 examines their sources of enforcement. Section 6 studies relationships within teams, leaders' standing, and attitudes beyond the team. Section 7 concludes.

2 Context

The Democratic Republic of Congo is one of the world's most populous countries and also one of the poorest. Kananga, the capital of Kasaï-Central Province and the setting for this study, is a city of roughly two million inhabitants. It is governed in a low-capacity fiscal environment. Roads are degraded, constant tropical storms cause severe erosion problems, and access to basic urban services such as electricity, piped water, and sanitation is low. The local administration is responsible for many of these services, but the lack of institutional capacity prevents it from delivering them. Public goods remain underprovided, and weak fiscal capacity leaves the city heavily reliant on local authorities to mobilize residents directly (Balán et al., 2022).

Salongo, City Governance and Traditional Authority. The principal institution through which this direct mobilization occurs is *salongo*. In the 70s, salongo was formalized as obligatory civic labor and became part of a broader effort to mobilize local labor for public works (Callaghy, 1984; Robinson, 2023). Recent evidence shows that households are still regularly asked to contribute salongo to neighborhood upkeep projects throughout the country, with participation organized by local authorities and sustained through social pressure, civic obligation, and repeated local interaction (van den Boogaard et al., 2026). In Kananga, salongo

[5] If anything, their point estimate is smaller under disclosure. This evidence aligns with prior evidence indicating that external control can crowd out internal motivation (Fehr and Rockenbach, 2003; Falk and Kosfeld, 2006; Gneezy et al., 2012). For a field experiment contrasting social recognition with task significance, see Fracchia et al. (2023).

[6] For background on traditional authority in Africa, see Michalopoulos and Papaioannou (2020). For indirect rule through customary chiefs in eastern Congo, see Henn et al. (2026).

remains part of everyday life, and many times it is the only solution residents have for waste management, road maintenance, public space cleaning, and, most urgently, anti-erosion works.[7]

Salongo operates through a local governance hierarchy that belongs to a broader political structure in which territorial origin shapes representation and expectations of political support (Hoffmann et al., 2022; Gobbers, 2016).[8] The *chef de quartier* is the formal neighborhood authority responsible for the administrative supervision of the population, hygiene, and public sanitation. Below the *chef de quartier*, the implementation of salongo relies on two further layers of authority. The first is the *point focal*, or *chef de localité*, which effectively serves as a focal point for the neighborhood chief.[9] *Points focaux* transmit instructions from the quartier administration, help identify feasible work sites, and organize the maintenance efforts across several avenue chiefs. They are implementation intermediaries. The second is the *chef d'avenue*, or avenue-level notable, who is the household-facing authority. He is closer to residents than the other chiefs.

The traditional-leader treatment focuses on the *chef d'avenue*, the level of chief authority that is closest to the resident population. Their status matters for both whom they can reach and how they can enforce contributions. First, a chief is a respected resident within a local order organized around ethnic and territorial origins. His authority should therefore be strongest among residents who recognize him as speaking for their own origin community. The second is their social source of enforcement. Given chiefs' standing and embeddedness in the local context, we expect their mobilization capacity to operate through reputational channels.[10] The treatment tests whether chiefly authority can make salongo participation socially consequential, and whether its force declines when the mobilized group falls outside their ethnic scope.

Fragmentation. An initial assessment of Kananga's ethnic landscape may wrongly infer a homogeneous environment. The city is largely Tshiluba-speaking, and most residents are linked to the Lulua umbrella identity. However, this broad classification contains important internal divisions. It consolidated from smaller kin, territorial, and chiefly units in the Kasai during the colonial period. Belgian recognition of the Kalamba lineage as Lulua sovereign led to power struggles between central and peripheral clans that remain unresolved (Ngalamulume, 1995; Stearns, 2018). Because this umbrella identity would flatten most residents into a single group, we define ethno-territorial identity (henceforth, ethnicity) using the five administrative territories of

[7] The relevance of these tasks is visible in the World Bank's Kananga Emergency Urban Resilience Project (World Bank, 2022), which focuses on gully erosion, drainage, and climate-resilient urban planning. Reports emphasize that the constraint is partly organizational since neighborhood erosion-prevention committees have been established across Kananga; yet, community participation remains below target.

[8] Within Kasai, Ngalamulume (1995) traces the formation of geographically different identities among the Lulua around Kananga.

[9] The *localité* appeared as a lower subdivision under the 1982 territorial organization (République du Zaïre, 1982), but is absent from the current 2008–2010 framework, which organizes urban communes into quartiers and/or incorporated groupements and lists territoire, quartier, groupement, and village as deconcentrated entities (République Démocratique du Congo, 2008, 2010). In Kananga, the category remains a relevant de facto administrative scale.

[10] Balán et al. (2022) characterize Kananga city chiefs as local notables nominated by neighborhood elders, typically because they are longstanding and respected residents, and then ratified by the state. They report indefinite and often lifelong tenure, rare deposition, some family transmission of the position, and status as its main return. They also show that chiefs possess household-level information and that this information helps explain their effects on tax compliance.

Kasaï-Central (Demba, Dibaya, Dimbelenge, Kazumba, and Luiza), together with Kananga city[11]. The administrative territories help us identify residents' ties to different kin networks that structure political representation and expectations of mutual obligation. These ties matter for local governance because they shape the reach of the authority of chiefs. Given the informal nature of salongo mobilization by local chiefs, variations in neighborhood ethnic fragmentation should help explain why salongo may have weak support in certain places while succeeding in others.

Descriptive evidence collected in this study corroborates the hypothesis that fragmentation complicates local governance around salongo activities. Using participants' stated ethnicity, we build a fractionalization index for each neighborhood and investigate how it relates to the perceived importance of salongo among local residents. Figure A-1 reveals that fragmentation weakens perceived salongo importance in the neighborhood through two margins: (i) residents from the majority ethnicity in a given neighborhood see salongo as less important the more fragmented the neighborhood is; and (ii) at the average level of neighborhood fragmentation, residents from less represented ethnicities see salongo as less important than residents from the majority ethnicity. A key part of our account is that this governance problem is tied to the ethnic bounds of *the chef d'avenue's* authority. Consistent with this account, Appendix Table A-2 shows that chiefs' personal networks are disproportionately composed of people from their own ethnicity, while Appendix Table A-3 shows that residents who are coethnics with the chief rate the chief more favorably in several aspects than non-coethnics.

Conflict. Beyond historical fragmentation related to the consolidation of several groups under the broader Lulua identity, the Kamuina Nsapu conflict made these internal divisions especially salient. Beginning in 2016 as a dispute over customary authority in Dibaya, the conflict became a regional war after the state's refusal to recognize Jean-Prince Pandi as the successor to the Kamuina Nsapu chiefship.[12] The war not only hardened those divisions due to clashes occurring in rural areas but also created chaos in the city of Kananga, which became a stage for hostilities. Several killings took place in the city, while schools, health centers, homes, and public buildings were destroyed. It also became a new home for hundreds of thousands of displaced villagers who were unable or unwilling to return to their villages of origin.[13] Therefore, beyond the deaths, destruction, and tensions left by the war, Kananga recently absorbed a violent re-sorting of people from different ethnicity and customary ties (Mercy Corps, 2019). Many neighbors who now share a communal task to maintain a

[11] Appendix Table A-1 locates the leaders' origins and religious denominations within the population from which experimental participants were recruited. Chiefs and pastors come from broadly the same origin territories as ordinary residents. In Kananga, implicit association tests find bias in favor of one's own ethnic group (Lowes et al., 2015).

[12] Stearns (2018) interprets this episode as an instance of *dédoublement*, a broader strategy through which Congolese authorities weaken customary leaders viewed as sympathetic to the opposition by withholding recognition and supporting rival claimants, while locating this occurrence within the historical divisions of the Lulua umbrella ethnicity. The Kamuina Nsapu lineage heads the Bajila Kasanga, the dominant group within the peripheral Lulua clans, which has long contested the Kalamba lineage and the political center around Kananga.

[13] Mercy Corps reports more than 3,300 direct conflict-related deaths and 1.4 million displaced people in 2017; by early 2019, several hundred thousand displaced residents had still not returned because of livelihood collapse, insecurity, and ethnic tensions. The same report notes that more than 400 schools and 200 health centers, along with many government buildings, were damaged or destroyed (Mercy Corps, 2019).

drainage channel or to address erosion risk are more likely to have arrived from different places of origin. The experiment focuses on the resulting mismatch between the chief's administrative neighborhood and the social networks through which his authority becomes relevant.

Religious Environment and Religious Authority. The DRC is overwhelmingly Christian, comprising 95% of the population, and is roughly divided between Catholics and broadly defined Protestants (Pew Research Center, 2011). The religious environment in Kananga roughly corresponds to these numbers, with the caveat that a broad Protestant classification hides the rapidly shifting religious landscape where autonomous *Églises du Réveil* are taking over the space previously filled by Catholics and traditional Protestant churches (Kapepula et al., 2022). The gap left by a weak state creates a wide space for churches to mobilize welfare and collective goods. Institutional churches, Catholic and historic Protestant networks, occupy the organizationally dense pole of this field through schools, clinics, development projects, and social assistance.

Entrant denominations perform a more decentralized and entrepreneurial role. They provide support through pastoral counseling, healing, deliverance, mutual aid, and informal insurance against illness, unemployment, and other risks. This decentralized model relies on local mobilization with few large bureaucratic assets.[14] Appendix Table A-1 locates leaders' religious denominations within the population from which experimental participants were recruited. Religious leaders are concentrated in Pentecostal/born-again churches (12 of 28), and the experiment also includes pastors from Kimbanguist (7), Neo-Apostolic (4), and Protestant/Presbyterian (2) denominations, while chiefs and residents are distributed more broadly across the Christian spectrum. The selection is informative in itself. Pastors whose authority depends on building a local following are exactly those most engaged in local mobilization for public goods. Catholic priests and other mission churches, by contrast, derive authority from their office within a larger hierarchy and need not build it through the same local involvement.[15]

Importantly, churches are spaces of ethnic mixing, where individuals with different kin networks meet and build lives unrelated to their origins. Baseline data show that more than 90 percent of identifiable congregations include members from multiple ethnicities (Appendix Table A-4). Congregants are no more likely to share an ethnicity within their churches than would be expected under random allocation within their neighborhoods. Within-church pairs are co-ethnics in 48.2 percent of cases, compared with 48.8 percent under the no-sorting benchmark. Religious leaders are therefore already leading organizational constituencies that cross the ethnic boundaries on which chiefly authority rests.

Thus, the religious-leader treatment arm focuses on pastors from autonomous churches in a context where religion is an important institution for ethnic mixing. These leaders provide a contrast to avenue chiefs on the two dimensions central to the design. First, a pastor can invoke a religious obligation among residents

[14] On institutional church provision in Congo, especially health and education under weak state capacity, see Seay (2013); Titeca et al. (2013). On healing churches and everyday social support in urban Congo, see Devisch (1996).

[15] Recruitment was open to leaders from all denominations.

from different origins. The same religious framework can travel across the ethnic boundaries that constrain traditional authority.[16] Second, pastors can invoke God's judgment of participation. The relevant audience for such an appeal is God; under social enforcement, it is the neighbors and local authorities who see a resident's contribution.[17] We use the pastor treatment to test whether these appeals can sustain cooperation among residents over whom the chief has little ethnic authority.

3 Experimental Design and Data

In this context of post-war recovery, where neighborhoods have recently integrated a massive wave of displaced residents, Kananga's city administration organized a city-led *salongo* campaign to coordinate basic neighborhood maintenance and incentivize citizens to join community activities. The campaign brought residents together to contribute labor and materials to neighborhood upkeeping tasks and used city evaluation and public recognition to reward performance, a common practice in the region.[18] We partnered with the administration in these efforts and embedded an experiment in the campaign.

Neighborhood selection. Before rollout, we identified neighborhoods with active *salongo*, identifiable boundaries, and sufficient scale for team implementation. Fifty neighborhoods met these criteria. We stratified them by commune and by whether the median predicted property value lay above or below the sample median, and randomly assigned 40 neighborhoods to receive the intervention rollout. The main estimands compare randomized teams within these intervention neighborhoods. Our main analysis covers 28 neighborhoods, 336 teams, and 1,040 non-leader participants due to a public health emergency that caused the program to stop.[19]

Recruitment and short baseline. Before each wave, enumerators visited households in participating neighborhoods to introduce the campaign and register adults interested in taking part. This short baseline recorded residents' demographic characteristics, territorial origins (ethnicity), and willingness to contribute, and asked them to refer traditional and religious leaders who were especially committed to making *salongo* succeed. The resulting data defined the resident sampling frame, identified candidate leaders, and supplied the variables used in randomization.

[16] On religious identity as an axis of mobilization that differs from ethno-linguistic identity in Africa, see McCauley (2014). On Pentecostal and charismatic Christianity as a portable religious form, see Robbins (2004).

[17] In Kananga's revival churches, concerns with witchcraft, protection, prosperity, and a worldview of hidden supernatural causality are translated into a Christian vocabulary of sin, blessing, and divine judgment. For an account of Congolese revival Christianity, see Pype (2012); Tonda (2005). In eastern Congo, access to protective rituals changed theft beliefs and inventories among sellers who believed the rituals worked (Dunia et al., 2023).

[18] Comparable civic-work and cleanliness contests elsewhere in the region include Kinshasa's Kin Bopeto, which promised prizes to the cleanest quartier in each commune (Actualite.cd, 2019); Rwanda's Umuganda, which links regular community work to infrastructure, environmental protection, and cleanliness (Rwanda Governance Board, 2017); Lomé's "salubrité challenge des quartiers," which had quartiers submit streets for jury evaluation (Vert Togo, 2019); and Cameroon's "Villes propres" competition, which began in 2019 as a Yaoundé "Quartiers propres" contest (Ministère de l'Habitat et du Développement Urbain, 2022).

[19] The study was planned as ten waves of four neighborhoods. Seven waves were completed under the standard protocol before the Ebola emergency disrupted the eighth wave and prevented the remaining rollout (World Health Organization, 2026). Appendix B.1 documents the stopping decision and reports robustness estimates incorporating the disrupted eighth wave.

3.1 Salongo Campaign

During each implementation wave, local officials organized daily maintenance activities and recorded residents' contributions for five weeks. In parallel, in the first week of the campaign, we invited volunteers to participate in coordination teams that would help advance salongo in the neighborhood. Each team met twice during the campaign in coordination meetings held separately from the daily work: first in the second week of activities and again in the third week. The meetings gave residents the opportunity to review their team's recent contributions, coordinate future work, and encourage each other to participate.

At the first meeting, participants got to know their assigned teammates, were exposed to their set of experimental conditions, and learned that the city would publicly recognize highly engaged teams at the end of the campaign. Because this was the first moment participants encountered their team and disclosure condition, the first meeting marked the onset of the randomized treatments. By the end of the fourth week of activities, the three teams with the highest engagement received public recognition from the city in a neighborhood ceremony attended by city officials. A fifth week of activities was organized after the recognition ceremony. All teams faced the same maintenance activities and recognition rules. The upper panel of Figure 1 summarizes the timeline of activities.

3.2 Randomization

The assignment proceeded in two stages. First, a screening stage selected participants from the registration sampling frame, oversampling those from the neighborhood's ethnic majority to facilitate the construction of homogeneous teams. Second, an assignment stage organized the selected residents into twelve teams of four in each neighborhood, composing a factorial design that crossed three leader conditions, two territorial-origin compositions, and two disclosure conditions. The design varies these conditions while keeping the neighborhood, campaign, team task, and recognition rules fixed. Figure 2 summarizes these assignments. Appendix B.2 presents the complete algorithm.

3.2.1 Leader Treatment

Each team counted on a fourth participant who would define the leader condition. Enumerators used the recruitment stage leader nominations to identify leaders who were available to join the campaign. They were instructed to select the leader with the most mentions who was willing to participate. In the traditional leader condition (T), the same ethnic-majority traditional leader joined four teams in the neighborhood. The nominated religious leader did the same in the religious leader condition (R). In the no-leader condition (C), a placebo participant joined the fourth position.[20] We keep all non-repeated residents in the analysis sample

[20]Beginning in Wave 3, the same randomly selected placebo participant joined all four no-leader teams, making team size and repeated participation symmetric across the three leader conditions. In Waves 1–2, each no-leader team contained a different fourth ordinary resident.

while we exclude repeated leaders and placebo participants. All fourth participants followed the same meeting protocol. Leaders received no script or special instructions.[21]

Traditional leaders. The traditional leaders in the study are chiefs who are part of Kananga's customary system of territorial governance. We selected these leaders through residents' referrals collected during volunteer recruitment. These leaders are local notables whose main responsibilities include organizing communal work, mediating disputes, and representing residents before the state (Balán et al., 2022). Appendix Table A-2 shows that, relative to ordinary residents of the same neighborhood, the chiefs recruited into the study are twenty-two years older, have about 1.5 additional years of schooling, and have lived on their avenue for roughly thirteen years longer; they are also overwhelmingly male. Their institutional position is visible given the higher chances of holding a traditional head role, knowing local authorities, planning *salongo* activities, and evaluating higher chiefs' contributions to infrastructure, mediation, and representation favorably. Their personal networks are disproportionately composed of people from the same ethnicity, and they score higher than ordinary residents on an aggregate index of standard leadership traits, particularly conscientiousness. Appendix Table A-3 uses surveys administered before the first meeting to show that chiefs are known and recognized as authorities across ethnic lines, but their broader standing is concentrated among co-ethnics. Residents who share a chief's ethnic origin rate the chief as more motivating, morally exemplary, trustworthy, and hardworking, whereas residents from other origins recognize the chief's formal authority but do not evaluate the chief more favorably on these other dimensions. Coethnic and non-coethnic residents, nevertheless, expect chiefs to participate more in *salongo*, with especially strong expectations of their work effort. Therefore, chiefs enter the experiment as incumbent organizers of salongo whose authority, relationships, and reputations are centered on their ethnic origin.

Religious leaders. The religious leaders are local pastors drawn from Kananga's expanding landscape of Pentecostal, charismatic, and independent churches, institutions whose activities frequently extend beyond worship into community organization and public life (McClendon and Riedl, 2015, 2019; Kapepula et al., 2022; Ngoma Mavungu et al., 2025). We recruited religious leaders through the same resident-referral procedure we used to recruit chiefs. Appendix Table A-2 shows that pastors have a similar demographic profile to chiefs. Relative to ordinary residents, pastors are thirteen years older, have approximately one and a half additional years of education, and have resided in the neighborhood for roughly thirteen years longer; they are even less likely to be female than chiefs. Their most significant difference rests on their participation in congregational life. Religious leaders are much more likely to head a congregation, attend Sunday services frequently, contribute to their church, and depend on networks centered on fellow church members. They are similar to ordinary residents on standard leadership-personality measures but report more positive and less

[21]The fixed placebo participant was selected from the majority ethnicity. Religious leaders were recruited through religious authority and did not always belong to the neighborhoods' majority. Eleven of the 28 assigned religious leaders differed from the traditional leader. We report a robustness check restricted to the 17 neighborhoods in which both leaders were from the majority.

negative religious coping, with weaker evidence of greater intrinsic religiosity. The institutional boundary is permeable: roughly 40 percent of pastors report holding a traditional role, and 29 percent of chiefs report a religious role. Nevertheless, the two leader types remain sufficiently different in their local roles, where chiefs are representatives of Kananga's territorial governance system, while pastors are more firmly embedded in congregational organization. Appendix Table A-3 shows how this distinction is reflected in residents' evaluation of leaders. Pastors are better known across ethnic lines and are rated as more motivating, morally exemplary, trustworthy, and hardworking, with similarly favorable assessments among co-ethnics *and* residents of a different ethnic origin. Residents also expect pastors to participate in *salongo* and contribute financially, although the expectations of their efforts are weaker than those for chiefs. Pastors, thus, enter the experiment with less incumbent responsibility for *salongo* but with a favorable standing that extends across the territorial boundaries organizing customary authority.

3.2.2 Cross-randomizations

Team composition. We filled each of the three remaining seats of the neighborhoods' twelve teams in random order. For the six homogeneous teams, all three residents were drawn from the pool that shared the traditional leader's territorial origin. For the six mixed teams, we assigned residents from territories not yet represented in the team, repeating an origin only when no additional territory was available. This procedure ensured that the three ordinary residents in mixed teams came from different territories of origin. Within each leader condition, two teams were homogeneous and two were mixed.

Disclosure. Residents could contribute at different times and work sites; therefore, they could not fully observe their teammates' contributions. Leveraging the incomplete information setup, we varied the information shown at the meeting. Under group disclosure, participants saw only their team's aggregate prior *salongo* contribution (private condition). Under individual disclosure, they also saw each member's identifiable contribution record, increasing external reputational incentives (public condition). Within every leader-by-composition cell, one team was assigned to each disclosure condition.

3.2.3 Rerandomization and Replacements

For each wave, the assignment program generated 48 complete candidate assignments[22] and retained the one that minimized an imbalance score across 15 pre-assignment characteristics in the cumulative implemented sample. During the invitation period, some selected residents reported that they would not be able to attend their meeting slot. These residents were replaced before treatment using randomly ordered replacement lists of respondents from the same territory of origin. Replacements inherited the original resident's team and

[22]Following Banerjee et al. (2020), we set the rerandomization budget to $K = \min\{N, 100\}$, where N is the number of experimental units. Each wave covered four neighborhoods, each with 12 teams, for a total of $N = 4 \times 12 = 48$ units.

experimental conditions, and all availability decisions occurred before residents met their teammates or the disclosure condition.

3.3 Data

Beyond the short survey used to construct the experimental assignments, the analysis combines administrative records, survey data collected before and after treatment, and recordings of the team meetings. The bottom of Figure 1 shows when each data collection instrument was administered.

Administrative data. The administrative records contain individual daily measures of labor, in-kind contributions such as food or materials, and monetary contributions. Local staff recorded contributions using a digital tracking system created for the campaign. The registry contains the neighborhood, date, and activity site, along with the geolocation of each salongo location; classifies the type of maintenance activity; and collects photographs of the sites before and after work. Arrival and departure time stamps measured contributions of labor for each participant, along with participant ratings of effort. In-kind and monetary contributions were recorded with quantities and estimated values. The records were merged with registered participants and dates, producing high-frequency, individual-level data on actual contributions to neighborhood maintenance.

Survey data. We collected survey data at baseline, two midlines, and endline. The baseline survey was administered before meeting invitations and measured participants' demographic and socioeconomic background, religious practices and beliefs, local social embeddedness, perceived salongo norms, trust in local authorities, and expectations of social and divine reciprocity. The first midline was conducted around the first team meeting. Right before the team meeting, it measured participants' proximity to each other and shared attributes such as church affiliation or kinship. Immediately after the meeting, it measured meeting impressions and immediate expectations of future peer participation. The second midline took place around the follow-up meeting and measured interactions during the inter-meeting period, including communication among team members, encouragement to contribute, praise for contributors, and sanctioning of free-riders. The endline questionnaire was administered in the week after the recognition ceremony and repeated the main baseline modules.

Meeting recordings. We recorded team meeting discussions with individual consent and transcribed the content to conduct a textual analysis of the leader's mobilization strategy. Because leaders received no script, these recordings show what they endogenously contributed to the discussion. We specifically investigate how chiefs and pastors differed in their contributions to the discussion.

3.4 Implementation Checks

Appendix Table B-4 displays covariate balance across experimental arms. The characteristics measured before assignment are similar across experimental arms. Participation during the six days after assignment and before the first meeting is imbalanced. At that point, residents had not learned their assigned leader or reporting rule, so the difference is unlikely to reflect a response to treatment. We nevertheless examine whether anticipation effects can account for the results in the alternative-explanations subsection. The remaining implementation checks are reassuring. Although 22.2 percent of residents entered through replacement queues, nearly 95 percent of groups still had the territorial composition assigned by the design. Homogeneous groups remained overwhelmingly composed of residents from the chief's territory, while mixed groups drew residents from several territories (Appendix Table B-5). Treatment also reached the intended groups. Residents attended first meetings 95.0 percent of the time, assigned leaders attended 99.6 percent of treated meetings, and delivery of the meeting instructions and reporting rules was verified in at least 99.4 percent of groups (Appendix Table B-6). Finally, administrative participation is observed for every resident over the same six-day outcome window, and survey follow-up remains high (Appendix Table B-7). Group formation, meeting delivery, and measurement of the main outcome therefore proceeded as planned.

3.5 First Stage: What did leaders bring to meetings?

Leaders attended their assigned meetings, but what did they contribute to the discussion? Because chiefs and pastors received the same instructions and no script, the recordings allow us to examine whether assigning different leaders exposed residents to different appeals. We study the first group meetings, before the subsequent participation outcomes are realized.

To compare their contributions, we use a topic model to analyze transcripts from the first group meetings (Blei et al., 2003). The model groups words that often appear together into latent topics and is blind to every other aspect of the intervention, including group structure, treatment condition, and leader role. Because our passages are short, we follow the approach of Quinn et al. (2010) and assign each passage to one of three categories. We name the topics only after estimation. We present this simplest specification and compare it with more flexible topic models and embedding-based methods in Appendix Section B.4.[23] The conclusions are unchanged across models. Appendix Table B-8 reports translated passages that help interpret each topic.

Latent topics. The first passages in Appendix Table B-8 highlight appeals for more regular participation. We labeled these as attendance-and-performance speeches. The excerpts display leaders setting clear targets for hours and regular participation, calling out those residents who did not show up and praising groups that

[23]For the use of text as data and the measurement and inference issues it raises, see Gentzkow et al. (2019) and Dell and Rambachan (2026).

meet expectations. In these passages, residents are asked to contribute because they owe something to the group, and the leaders assess what they have done against what the group expects.

The second block of the table turns to religious appeals. The examples tell volunteers to look to God rather than wages for reward, state that "God rewards each person according to their deeds", and connect collective work to patience, peace, and mutual aid. Residents are being asked to contribute both because God judges their actions and because they should support one another. Collective work therefore has a moral value beyond its immediate material return: it meets obligations to God and to other residents.

The last topic concerns operational constraints. A recurrent topic of discussion was the coordination of who would provide the materials for neighborhood improvement. Some examples concern within-group coordination, where peers would discuss and agree on who would provide the necessary tools for the work to take place. Yet, the passages also show substantial reliance on the centralized provision of these materials, with many examples praising the municipality for providing the tools for the work. Although it may be useful material for program improvement and future research, it is less directly concerned with appeals to participate. The main difference between leaders' appeals comes from the relative emphasis on attendance and performance or on divine reward and solidarity.

Who talks about what, and when? After assigning each passage to its most likely topic, we sum the passage duration to measure how many minutes each speaker spends on attendance and performance, material support, and divine rewards and solidarity. We aggregate the classified passages to the leader-by-meeting level. Figure 3 shows the average share of speaking time devoted to each topic, pooling the four composition-by-disclosure environments. Blue bars report traditional leaders and gold bars religious leaders. Appendix Table B-9 reports total speaking time and topic shares separately by environment.

The figure shows that leaders' discourse differs across leader types. Chiefs devote 44.1 percent of their speech to attendance and performance, compared with 26.0 percent among pastors. In contrast, pastors devote 48.3 percent to divine reward and solidarity, compared with 27.4 percent among chiefs. Both differences are statistically significant after adjusting for multiple testing ($q = 0.009$). Leaders devote similar shares to material support (28.5 vs. 25.7 percent). Appendix Figure B-2 shows the same distinction across the full set of models. Appendix Table B-9 shows that these differences persist across experimental conditions. Pastors speak for 3.43 minutes on average, compared with 2.55 minutes for chiefs ($p = 0.067$). Within each leader type, speaking time and topic shares are fairly stable across composition and disclosure conditions. All tests of equality in topic shares across environments have $p > 0.30$, while the speaking-time tests have $p = 0.701$ for chiefs and $p = 0.139$ for pastors. Thus, what participants hear differs mainly by leader type, not by the experimental environment. Appendix Figure B-3 reports the corresponding levels across models in each environment.

Although what leaders say is similar across experimental conditions, their influence on the discussion

varies. Appendix Figure B-4 reports adjusted participant speaking levels, including the no-leader control, across all text models. Chiefs do not consistently shift participants' speech toward their characteristic appeals, while pastors tilt the discussion toward divine reward and solidarity. Pooling disclosure conditions, this shift accompanies additional speaking time in homogeneous groups. In mixed groups, divine speech instead rises as attendance-and-performance talk falls, with little change in total speaking time. We next ask whether these different appeals translate into participation in homogeneous and mixed groups.

4 Scope of Authority Results

Traditional and religious leaders are notably different from each other and from ordinary citizens, but how do the different profiles and appeals translate into the capacity to lead different groups toward a collective goal? We call a leader's scope of authority the range of people and social settings in which the leader can influence behavior toward such a goal. A leader's scope may depend both on whether individuals recognize the leader's authority and on whether the relationships among group members give relevance to the obligations that authority invokes. Group composition can change both.

For chiefs, sharing the same origin provides one channel for individual recognition of authority, as residents who share the chief's ethnic origin should be more responsive to an ingroup authority (Miguel and Gugerty, 2005; Roy and Houser, 2024). This account predicts greater effectiveness of chiefs in homogeneous groups, where all participants share the chief's origin. But it does not necessarily imply that chiefs lose influence over coethnics when they are placed in mixed groups, nor that their authority is necessarily bounded by ethnicity. Ethnic proximity is only one reason for recognizing the authority of the chief. Participants from other ethnic groups may also regard the chief as a legitimate authority and therefore respond to his appeals, even if they are, on average, less closely connected to him than the coethnics who compose homogeneous groups. Hence, chiefs' scope should follow their relationships with particular residents, even when ethnic identity is not shared by the group as a whole.

For pastors, an open question is whether religious motives can broadly influence believers across and within ethnic divides, or whether the social setting can interfere with how much individuals respond to religious leaders' appeals. If belief in divine judgment is sufficient for religious leaders' speeches to influence behavior, pastors may be effective among believers regardless of group composition. An alternative account builds on the distinction between systems of enforcement that operate through kin and community sanctions versus those that rely on internalized moral obligations (Enke, 2019). These systems may coexist within individuals, with social context shaping which identities and associated obligations become salient and guide behavior (Akerlof and Kranton, 2000). In homogeneous groups, shared ethnic identity may give socially enforced ethnic obligations precedence in organizing collective action. In mixed groups, where ethnicity provides no common basis across all participants, it is when religion becomes the common identity denominator that

pastors can leverage as an alternative basis for encouraging cooperation. Under this interpretation, even residents who recognize religious authority may respond differently to a pastor depending on whom they are asked to cooperate with.

We next use our design to examine how group composition shapes each leader's scope of authority.

4.1 Empirical Strategy

Within each neighborhood, we compare groups assigned a traditional or religious leader with groups assigned no leader. Let i index participants, $g(i)$ their accountability group, and $n(i)$ their neighborhood. The variables $T_{g(i)}$ and $R_{g(i)} \in \{0,1\}$ indicate whether participant i's group was assigned a traditional or religious leader, respectively; no leader is the omitted category. Group composition is $c(i) \in \{H, M\}$, where H denotes groups with shared territorial origins and M denotes groups with participants from mixed origins. Finally, $q(i) \in \{0,1\}$ indicates whether individual participation was disclosed.

The estimation sample contains 1,040 participants in 336 groups across 28 neighborhoods.[24] The outcome Y_i is average salongo participation, defined as contributing labor, money, or materials to communal public works, measured in percentage points over the six-day administrative window immediately following the first accountability meeting. All specifications control for baseline participation, Y_{i0}, defined as average participation over the six-day window preceding any accountability meeting.[25] For the scope analysis, let $e(i) = c(i)$ and $\mathcal{E} = \{H, M\}$, pooling leader effects across disclosure conditions. We estimate:

$$Y_i = \lambda_{n(i),q(i),c(i)} + \rho Y_{i0} + \sum_{e \in \mathcal{E}} \left(\tau_e^T T_{g(i)} + \tau_e^R R_{g(i)}\right) \mathbf{1}\{e(i) = e\} + X_i'\beta + \varepsilon_i. \tag{1}$$

The vector X_i contains the 25-meter spillover indicator and 11 LASSO-selected baseline controls used in the main figures and the third table specification. Appendix Section C.2 describes the selection. The neighborhood-by-disclosure-by-composition fixed effects $\lambda_{n(i),q(i),c(i)}$ ensure that each leader group is compared with no-leader groups in the same neighborhood and experimental environment. Hence, for each environment e, τ_e^T, and τ_e^R are the intention-to-treat effects of assignment to a group with a traditional or religious leader, respectively, relative to no leader under the same conditions.

4.2 Main Results

The lower panel of Figure 4 plots the average predicted levels of groups without a leader and shows similar participation across no-leader groups ($p = 0.536$), establishing our first finding on participation: in the absence of a leader, ethnic cohesion alone does not improve participation. This finding aligns with previous work on

[24] The sample includes the three ordinary residents assigned to each group and 32 additional ordinary residents who filled the fourth position in no-leader groups during the first two waves. Repeated fourth participants are excluded.

[25] We use a common pre-meeting window to ensure that the baseline measure is not affected by potential spillovers from earlier meetings. The main specification controls for potential spillover exposure, the construction of which is detailed in Appendix Section C.1.

how informal institutions govern the commons, which highlights the role of leaders in establishing sanctions to prevent free-riding (Ostrom, 2000; Miguel and Gugerty, 2005; Kosfeld and Rustagi, 2015).

Figure 4 upper panel then reports the paper's main finding: ethnic fragmentation changes which leader can mobilize collective action. Assigning an urban chief to a group of participants who share the leader's territorial origin increases participation by 6.0 percentage points (SE 2.8; $p = 0.032$), a 10.5 percent increase over the 56.8 percent no-leader mean. In contrast, the chief's effect in mixed groups is only 0.4 points (SE 2.7; $p = 0.893$). Perhaps surprisingly, pastors display the *opposite* pattern. The presence of a religious leader increases participation by 5.8 points in mixed groups (SE 2.8; $p = 0.039$), or 9.9 percent relative to the 58.3 percent no-leader mean, and by *just* 1.7 points in homogeneous groups (SE 2.4; $p = 0.484$). The traditional-minus-religious contrast is 4.3 points in homogeneous groups (SE 2.7; $p = 0.113$) and -5.4 points in mixed groups (SE 3.0; $p = 0.067$); the difference between these gaps is $\widehat{\Delta}_C = 9.8$ percentage points (SE 4.1; $p = 0.017$).

We interpret these results as consistent with chiefs and pastors having different scopes of authority. Leaders are effective when participants recognize their right to ask them to contribute. Chiefs derive this standing from kinship, ethnic ties, and their customary responsibility for *salongo*. Thus, when chiefs share origins with participants in the group, their presence strengthens the pressure to meet a demand that participants already recognize as legitimate. In contrast, chiefs may lack the same standing to demand contributions from residents of other origins in mixed teams. Pastors derive their authority from a shared Christian worldview that crosses territorial boundaries. Because mixed teams lack a common customary basis for participation, pastors can frame community work as an obligation before God and provide a shared reason to contribute. The same appeal has little impact in homogeneous teams, where participation is already recognized as an ethnically enforced obligation. The force of this social obligation, however, appears to depend on the presence of the chief who is entitled to invoke and enforce it. Pastors' religious appeals to God's judgment seem to matter only where ethnic obligations do not organize behavior.

This interpretation echoes other studies on how religious commitments coexist with ethnic obligations and how their relative importance depends on the social context. Gluckman (1940) describes how kinship obligations remain binding across religious divisions in certain contexts in Zululand, while Laitin (1986) shows how ancestral-town affiliations take precedence over Christian–Muslim divisions among the Yoruba. In the Congolese context, Placide Tempels' *Bantu Philosophy* (1959) links clan membership to communal duties and chiefs' enforcement authority while emphasizing the persistence of inherited understandings among Christian converts. These accounts distinguish religious affiliation from the obligations that govern a particular interaction. They complement our interpretation by establishing a precedent where socially enforced ethnic obligations coexist with internally enforced religious obligations (Enke, 2019), and where the relevance of those obligations depends on the setting an individual faces (Akerlof and Kranton, 2000). In our setting, group composition shapes the obligations that govern communal work; and where territorial ties organize

social relationships, ethnic duties take precedence and reduce the relevance of religious appeals, even among committed Christians.

Composition changes which leader can sustain collective action. A shared territorial origin places *salongo* within the chief's recognized authority, while mixed origins create room for authority that crosses territorial boundaries. This is our *scope of authority* result. But before moving into complementary evidence for the proposed account, we discuss the results' robustness and unlikely alternative explanations.

The response is concentrated in labor. Across alternative outcome definitions, the effects primarily appear in labor participation. We estimate Equation (1) using alternative measures of labor participation: total days, total hours in levels, inverse-hyperbolic sine transformation of total hours, and ratings of effort by city staff. Across the same 1,040 participants and main outcome window, the Δ_C pattern corresponds to 0.56 additional days of participation. Total hours contributed and rated effort also move in the same direction. Adjusted p-values across the four labor outcomes are at most 0.047 for Δ_C (Appendix Table C-2), and the asinh-hours specification yields the same conclusion (Appendix Table C-2). Contrastingly, cash and material contributions show no comparable pattern (Appendix Table C-3).

Robustness. The main pattern is stable after accounting for exposure to potential spillovers from neighbors that were treated and other baseline controls (Appendix Table C-1 and Appendix C.1). Randomization inference, which repeats the original group-formation procedure across 5,000 assignments while holding the realized sample fixed, reinforces our conclusions (Appendix Table C-4). The main contrasts also remain significant when allowing errors to be correlated across groups assigned to the same leader.

Alternative explanation I: endogenous leader behavior. The main alternative explanation for the composition pattern is that leaders changed their behavior in the groups where they succeeded. Section 3.5 already addresses part of this concern through the speech evidence. We revisit that evidence here, then turn to co-arrival at *salongo* and group participants' reports of follow-up contact.

i. *Meeting effort.* We return to the comparison of speaking time and topic shares of leaders across compositions, using the three topics introduced in Section 3.5 and leader fixed effects (Panels A–B of Appendix Table B-9).[26] Chiefs could have put in more effort in mobilization discourse among their own coethnics, or pastors could have changed their speeches depending on the composition of the group. We find little evidence of leaders behaving differently across meeting environments, but we find significant differences between the behaviors of pastors and chiefs. Pastors speak more and place emphasis on religiously related speeches, while chiefs speak for less time and emphasize speeches about social accountability. Leaders are better understood as delivering consistent treatments across environments while differing

[26] For each leader type, we estimate $y_{\ell g} = \alpha_\ell + \beta H_g + \varepsilon_{\ell g}$, where $y_{\ell g}$ is leader ℓ's speaking time or topic share in group g's first meeting, H_g indicates a homogeneous group, and α_ℓ is a leader fixed effect. We pool disclosure conditions and cluster standard errors by neighborhood.

substantially in what is delivered.

ii. *Salongo coordination.* Leaders could, nevertheless, differ in how they coordinate participation after the meeting. Chiefs could actively target coethnics more often because of their social proximity, or pastors could exert more effort to bring mixed teams' participants to the activity because they know they would be constrained by the chief's jurisdiction when with members of homogeneous groups. We approach this potential explanation with behavioral and reported measures. If mobilization efforts occur outside of the meeting and leaders bring participants to the activity with them, we should observe this pattern arising in arrival times at the worksite. Leaders and other team members would arrive at the same time. We measure co-arrival in two ways: (i) whether the leader's and participant's recorded arrival times are no more than one or two minutes apart; and (ii) whether they belong to the same arrival group, defined as a sequence of check-ins with no more than one minute between consecutive check-ins (thirty seconds in the alternative definition).

   The pattern we find mirrors the speech evidence. Behavior differs across leader types, while differences across compositions are imprecisely estimated. Chiefs arrive with participants more often; participants of a group with a chief arrive with the leader on 8.2% of days in mixed groups and 10.5% in homogeneous groups (Panels A–B of Appendix Table C-5). With pastors, the same levels are around 5–6% of days. These rates use all six post-meeting days, counting nonattendance as zero. Pooling compositions, the chief co-arrival rate is 9.4%, about 14% higher than the 8.2% rate per ordinary groupmate; the pastor rate is 5.5%, about 40% lower than the corresponding 9.2% rate. The pastor rate is lower than both the chief rate (difference -3.8 percentage points; $p = 0.009$) and the ordinary-groupmate rate (-3.7 points; $p < 0.001$).

iii. *Effort outside meetings and salongo.* Finally, follow-up contact may have occurred and may not necessarily be related to arrival times. Leaders could be devoting time outside of activities to exert sensitization efforts. We use the group members' reports collected at the start of the second meeting about contact with the assigned leader during the preceding week. Panels A–B of Appendix Table C-6 compare the rate of these responses between homogeneous and mixed groups. We ask whether participants had interacted with the group's leader, whether they discussed salongo, whether the leader explicitly asked the respondent to participate in salongo during this conversation, and if they mentioned social or divine consequences associated with showing up. About 41–50% of participants report contact with the leader during the work week, and 11–14% report being explicitly asked to participate. Explicit mention of social or divine consequences is far rarer, occurring in fewer than 4% of cases. There is no statistically significant evidence that leaders adjust these behaviors based on the composition of the group.

Taken together, these checks provide no consistent evidence that leaders intensify their activity or tailor their appeals in the groups where they succeed; they leave room for differences in effort and in the quality of interactions.

Alternative explanation II: belonging to the majority ethnicity. The composition pattern could instead reflect a belonging to the neighborhood's majority ethnicity. One could argue for an account in which pastors do not work with homogeneous groups because coethnics with the chief would only respond to the chief. In this account, it is not the group context that matters (context-dependent) but the trait of belonging to the ethnicity that composes homogeneous groups, the ethnic majority of the neighborhood (trait-dependent). As we discuss in Section 4.3, the regression in Appendix Table C-7 shows a positive pastor effect among majority residents assigned to mixed groups, compared with a small effect that is statistically indistinguishable from zero in homogeneous groups. Although the direct difference is imprecise, this pattern suggests that belonging to majority ethnicity alone does not explain pastors' limited reach in homogeneous teams.

Alternative explanation III: pastors' null is not due to ethnic mismatch. Pastors may fail in homogeneous groups because they do not always share the territorial origins that link chiefs to those participants. In this account, pastors would fail to sustain participation in homogeneous groups not because of ethnic homogeneity, but because they are from a different territory. Out of the twenty-eight localities and twenty-eight recruited religious leaders, seventeen share their origins with the chief and the homogeneous groups in the neighborhood, while eleven do not. We leverage these seventeen religious leaders to verify whether our conclusion holds in a restricted sample where the religious leader also shares their origin with the homogeneous group.

Appendix Table C-8 restricts pastors' comparisons to neighborhoods where the assigned chief and pastor have the same territorial relationship with homogeneous-group participants. Pastors with the relevant territorial connection do not become more effective in homogeneous groups, and their effect remains concentrated among groups composed of diverse origins. The pastor result, therefore, reflects the limited reach of religious leadership within ethnically cohesive groups, rather than a simpler account in which an outsider leader would have a smaller impact.

Alternative explanation IV: leadership by example. Rather than projecting authority onto different residents, one could argue that leaders would incentivize contributions by serving as an example for the group, and the different effects by composition would be explained by whether residents consider that specific leader example to be followed. Leadership by example would predict larger effects under leaders who contributed more before the experiment (Jack and Recalde, 2015). Appendix Table C-9 compares leaders above and below their type-specific median participation in the common pre-meeting week. The higher-minus-lower difference is -5.8 percentage points for chiefs in homogeneous groups (SE 5.2; $p = 0.262$) and 3.2 points for pastors in mixed groups (SE 5.2; $p = 0.542$). These estimates provide little support for differences in leaders' earlier participation as an explanation of the composition pattern.

Alternative explanation V: No evidence of anticipation or mean reversion. The pre-assignment characteristics used to form groups are balanced. However, administrative participation, measured after assignment and before the first accountability meeting, differs across the realized cells (Appendix Table B-4). Participants received their invitation to participate in volunteer teams during this week, but had not yet learned about their teammates or the disclosure condition during this period. Nevertheless, two concerns arise: (i) participants may have learned that a leader was assigned to their team before the first meeting, creating anticipation effects, and (ii) even without anticipation, the estimates could be driven by mean reversion in the imbalanced cells.

We address the first concern in Appendix Table C-10, which examines the six daily pretreatment observations. The daily data show no systematic increase in participation before the first meeting, although some pretreatment paths differ. We address the mean reversion concern in Appendix Table C-11, allowing the relationship between pretreatment and subsequent participation to differ across all 12 assignment cells. All estimated coefficients remain essentially the same. Together, these checks provide little support for the imbalance as an explanation of the leader effects that emerge after the meeting.

4.3 Heterogeneous Treatment Effects: Who Can Mobilize Whom?

Our key result shows that leaders have distinct scopes of authority. Chiefs are effective when placed in teams where they share origins with residents. Pastors, on the other hand, are effective precisely in mixed teams, where chiefs have little influence. These results point to an account where ethnic cohesion has opposing impacts on the ability of leaders to organize collective action. While it enables chiefs by placing them in a context where their standing in an origin-based governance system[27] matters for enforcing contributions, it limits pastors by grounding cooperation in socially enforced ethnic obligations that reduce the relevance of a religious framing of participation. But what determines the boundaries of each leader's authority? Do broader patterns in who responds to each leader support this interpretation?

We examine this question using groups' average ethnic attachment, residents' relationships with their assigned leader, and their religious orientation. We also ask whether recognition of the chief's authority and personal receptiveness to religion predict participation. We estimate heterogeneity across these characteristics using an extension of the main specification. Let Z_{ig} indicate the higher category of the characteristic under study. Retaining the neighborhood-by-composition-by-disclosure fixed effects from Equation (1), we estimate

$$\begin{aligned} Y_i = \lambda_{n(i),q(i),c(i)} + \rho Y_{i0} + \sum_{c\in\{H,M\}} \eta_c Z_{ig} \mathbf{1}\{c(i)=c\} \\ + \sum_{L\in\{T,R\}} \sum_{c\in\{H,M\}} \left(\tau_c^L + \gamma_c^L Z_{ig}\right) \mathbf{1}\{L_{g(i)}=L, c(i)=c\} + X_i'\beta + \varepsilon_i. \end{aligned} \tag{2}$$

[27] See Section 2.

Here L_g denotes the assigned leader type. Both leader effects and the higher-minus-lower differences, γ_c^L, are pooled across disclosure. In some specifications, Z_{ig} is defined only for participants assigned a leader (e.g., leader ratings). In those cases, we omit the η_c terms, and the comparison is with the average participant in a no-leader group within the corresponding experimental stratum defined by $\lambda_{n(i),q(i),c(i)}$. We also estimate heterogeneous effects according to whether residents belong to the neighborhood's ethnic majority. Since homogeneous groups contain only majority residents by design, homogeneous-group effects remain pooled.

Figures 5 and 6 show heterogeneous effects for both leader types. The main figures include the pretreatment outcome, the design fixed effects, the 25-meter spillover indicator, and 11 LASSO-selected baseline controls. Appendix Tables C-12–C-16 and C-19–C-20 report both leader types, with and without spillover and LASSO controls. Appendix Section C.5 describes the variables. Standard errors are clustered by accountability group.

Group ethnic attachment and the channels of authority. Group composition determines whether all residents share an ethnic origin, while ethnic attachment captures how much importance they attach to that origin. Before being introduced to their groups, respondents ranked six identities by their importance. We score attachment to one's own ethnicity from 1 (least important) to 6 (most important) and average the available scores of ordinary group members. We divide groups at the median of group ethnic attachment, retaining all 1,040 residents in 336 groups. Our main comparison therefore examines an intensity dimension of the composition treatment (Figure 5). In homogeneous groups, the more residents care about the origin they all share with the chief, the more consequential their ethnic obligations should be for participation. This would give the chief a stronger basis for enforcing contributions, but leave pastors less room to present the work as a religious undertaking. Mixed groups pose a different problem: residents who care more about their origins are attached to different ethnic communities. If this makes it harder for them to cooperate across ethnic lines, participation should be lower without a leader and pastors should have more room to bring residents together through a shared religious obligation.

The comparisons in Figure 6 ask what makes residents responsive to each leader. For chiefs, we examine whether the residents who recognize their standing and view them favorably are also those whose participation they can influence, including in mixed groups (Figure 6A). This should be the case if the chief's authority comes from his ties with residents. With pastors, the relevant belief concerns God: if residents feel an internal religious obligation to contribute, the pastor should have more influence when their faith guides personal conduct (Figure 6B).

Chiefs influence residents with strong ethnic attachment. Figure 5A and Appendix Table C-12 compare groups with higher and lower ethnic attachment. In homogeneous groups, chiefs raise participation by 9.7 points in groups with higher attachment (SE 4.0), compared with 3.4 points in those with lower attachment (SE 3.9). The estimated difference is positive, at 6.3 points, but imprecise (SE 5.7; $p = 0.275$). In mixed groups, chiefs raise participation by 5.2 points (SE 5.3) in groups with higher attachment and -2.5

points (SE 3.5) in those with lower attachment. Here too, the higher-minus-lower difference is positive but imprecise, at 7.7 points (SE 6.8; $p = 0.260$). Among these four chief effects, only the effect in homogeneous groups with higher ethnic attachment is statistically significant at the 5 percent level. The direction in homogeneous groups is what we would expect if residents' attachment to the origin they share with the chief helps him enforce contributions. Appendix Table C-13 shows that the positive chief estimate in mixed groups with high ethnic attachment is concentrated among his coethnics (15.1 points, SE 10.5), compared with 1.2 points among other residents (SE 6.5), although the 13.9-point difference is imprecise ($p = 0.284$). Stronger attachment to the origin shared with the chief may thus help him mobilize residents.

Chiefs rely on ties with residents. Residents' recognition of chiefs' standing provides complementary evidence on their ability to enforce contributions, including in mixed groups. We asked each resident to rate the leader in their team in terms of trustworthiness, hard work, moral stance, and how much the leader inspires them. Coethnicity is associated with more favorable ratings across these dimensions (Appendix Table A-3), and chiefs' personal networks are disproportionately composed of coethnics (Appendix Table A-2). Figure 6A shows effects in mixed and homogeneous groups, with pooled no-leader references below, since the assessment is undefined without an assigned leader. Appendix Table C-14 shows a 6.8-point chief effect in mixed groups among residents who evaluate him more favorably (SE 3.6), compared with −3.5 points among those who do not (SE 3.4). The difference is 10.3 points (SE 4.5; $p = 0.023$). We find no clear evidence that pastor effects vary with these ratings in either homogeneous or mixed groups.

A more specific assessment points in the same direction. Appendix Table C-15 compares residents above and below the median answer to a Likert question on whether the leader of their group is able to sanction and reward residents. Estimates show that chiefs are able to influence the participation of residents in mixed groups who recognize their authority. If the resident is above the cutpoint, chiefs lead to 6.8 percentage points (SE 4.0) higher participation, compared with −2.1 points among those below it (SE 3.3). The 8.9-point difference is suggestive (SE 5.0; $p = 0.075$). These comparisons suggest that a chief can mobilize outside the homogeneous group when residents recognize his standing or view him favorably.

Coethnic ties point in the same direction, although the comparison is less precise. Appendix Table C-16 compares coethnics and non-coethnics of each leader, using coethnicity with the chief for chief effects and with the pastor for pastor effects. Chief effects in homogeneous groups remain pooled. Estimates show that the chief effect in mixed groups is 5.3 points among his coethnics (SE 6.7) and −1.3 points among non-coethnics (SE 3.3). The 6.5-point difference is imprecise (SE 8.0; $p = 0.414$). We interpret these results with caution since they draw on relatively few coethnics in mixed groups (34 groups/44 participants in chiefs' groups and 35 groups/42 respondents in the no leader condition). For pastors, the estimate is also larger among coethnics in mixed groups (9.7 points, SE 6.8, versus 4.5 points, SE 3.4), but smaller in homogeneous groups (0.9 points, SE 3.1, versus 2.6 points, SE 3.7). These opposite signs suggest no consistent advantage among

pastors' coethnics, although both differences are imprecise. Together with the ratings evidence, this suggests that the authority of the chief does not vanish in diverse teams. It is spread over fewer people who recognize him as their leader. Coethnicity delineates where the chief has authority.

Pastors' reach depends on the social context. Figure 5B and Appendix Table C-12 show pastor effects using the same group ethnic attachment split as for chiefs.

Our findings show that the null average effect of pastors in homogeneous groups masks opposing effects across the distribution of group ethnic attachment. In homogeneous groups at or below the median, pastors raise participation by 6.8 points (SE 3.3), compared with −4.4 points above it (SE 4.6). The higher-minus-lower difference is −11.1 points (SE 6.4; $p = 0.081$); the negative effect above the median is imprecise. In mixed groups, where greater ethnic attachment means attachment to different origins and therefore more division, we observe the opposite pattern: the pastor effect increases with ethnic attachment, from 0.1 points (SE 3.8) to 13.2 points (SE 4.2), mainly because it recovers cooperation in highly divided groups, which, without a leader, have considerably lower participation compared to mixed groups that are less attached to their divisive ethnic identities (52.4 versus 61.9 percent). The 13.0-point difference is positive (SE 5.8; $p = 0.025$). The homogeneous-minus-mixed difference in these gradients is −24.2 points (SE 8.6; $p = 0.005$). This pattern is consistent with shared ethnic attachment giving ethnic obligations precedence over religious appeals in homogeneous groups. In mixed groups, stronger attachment to different territories instead leaves more room for a religious framing to organize cooperation.

Could these patterns reflect individual receptiveness to religion rather than the group context? Embeddedness in the chief's ethnic network of reciprocity might weaken residents' responsiveness to religious appeals. In homogeneous groups, own ethnic attachment has small positive correlations with intrinsic religiosity ($r = 0.069$, $p = 0.150$) and extrinsic religiosity ($r = 0.067$, $p = 0.123$; Appendix Table C-17). Ethnic attachment and Christian identification are positively correlated after removing each identity from the other's ranking ($r = 0.173$, $p < 0.001$). The group-score comparisons also provide little support for lower religiosity as an explanation (Appendix Table C-18).

We return to the majority-ethnicity explanation discussed in Section 4.2. By design, homogeneous teams consist entirely of residents from the neighborhoods' ethnic majority. Mixed teams can contain residents from this majority together with residents from other origins. We can therefore compare pastors' influence on the chief's coethnics assigned to homogeneous and mixed teams to ask whether their ethnic tie to the chief or the composition of the group limits the pastor's influence. Appendix Table C-7 shows that pastors raise participation by 12.5 points among the chief's coethnics placed in mixed groups (SE 6.7; $p = 0.063$), while their effect in homogeneous groups is 1.7 points (SE 2.4). The 10.8-point difference is positive but imprecise (SE 7.1; $p = 0.129$). The estimates suggest that pastors can mobilize majority residents in mixed groups. This is consistent with shared origins within the group limiting the effectiveness of religious appeals.

Pastors reach intrinsically religious residents. Within a given group context, residents' religious commitment may still shape their response to pastors. Following Allport and Ross (1967), we distinguish between intrinsic religiosity, where faith is internalized as a guide to life, and extrinsic religiosity, where religion serves other personal or social needs. Extrinsic religiosity includes personal reliance, such as seeking comfort or support from religion, and social reliance, such as using the church for social contact. If pastors work through a religious frame of obligation that resonates with the audience listening to the appeal, we should expect greater influence among residents for whom religion guides personal conduct. We average the intrinsic and extrinsic items separately and divide residents into lower and higher categories.

Figure 6B and Appendix Table C-19 present the results. In mixed groups, pastors raise participation by 9.9 points among residents with higher intrinsic religiosity (SE 4.3), compared with 0.7 points among those with lower intrinsic religiosity (SE 5.0). The 9.2-point difference is positive but imprecise (SE 7.4; $p = 0.213$). We find no clear evidence that chief effects vary with intrinsic religiosity in either homogeneous or mixed groups (Appendix Table C-19). Extrinsic religiosity provides a useful comparison. Appendix Table C-20 shows pastor effects in mixed groups of 7.7 points among residents with lower extrinsic religiosity (SE 4.3) and 3.6 points among those with higher extrinsic religiosity (SE 3.9). The -4.1-point difference is imprecise (SE 6.0; $p = 0.498$). The pooled estimates therefore provide no clear evidence that responsiveness to pastors increases with extrinsic religiosity. Residents whose faith guides their conduct appear more responsive to pastors, as we would expect if they feel an internal religious obligation to contribute, but the differences by religiosity are imprecise.

Which traits make leaders effective? The comparisons so far point to ethnic ties that give chiefs standing with individual residents, and to group context and receptiveness to religious obligation for pastors. We next ask whether leaders' own characteristics and denominational ties help account for these differences (Appendix Section C.6). Chiefs' effects are suggestively stronger among those with deeper local roots in homogeneous groups, while pastors' effects are stronger among those with greater conviction in divine moral enforcement in mixed groups. Pastors also reach residents outside their denomination in mixed groups, with imprecise evidence that denominational ties extend their reach into homogeneous groups. These comparisons suggest that local roots matter for chiefs, while conviction in religious obligation matters for pastors.

Taken together, these results support the interpretation that chiefs' influence operates through established relationships in the neighborhood, while pastors depend on religious claims regarding divine judgment.

5 Source of Enforcement Results

After investigating *whose* participation a leader can enforce, a question that naturally follows is *how* participation is enforced. The preceding results suggest that leaders make participation consequential through different channels. Yet the evidence is not enough for a conclusive interpretation. Do residents contribute

because they expect others to judge their behavior? Or do they contribute even when others are not watching because of an internalized obligation?

For chiefs, their greater influence over residents who recognize their capacity to sanction and reward is consistent with their role as *external*, social enforcers of behavior (Kosfeld and Rustagi, 2015). If this account is true, increased public accountability should strengthen their capacity to enforce behavior. For pastors, their larger influence over residents who score higher in intrinsic religiosity points toward *internal* religious motivations (Shariff and Norenzayan, 2007). These internal motivations do not require others to observe participation, so public accountability should not strengthen pastors' influence. Pastors' effect on participants could even decline if disclosure shifts attention toward social image concerns that provide weaker incentives in ethnically mixed groups.

We next use our experimental variation in public accountability to disentangle these sources of enforcement.

5.1 Empirical Strategy

To distinguish between internal and external sources of enforcement, we randomized whether meetings disclosed individual information about peers' contributions. In teams with disclosure, each member's contribution record was public at every meeting, while teams without disclosure could only see the total group contribution. This experimental arm changed how much participants knew about their peers' participation. Public accountability can strengthen external enforcement by making behavior visible to others, but it should not strengthen internal enforcement channels because individuals already observe their own actions.

We now set $e(i) = (q(i), c(i))$ and $\mathcal{E} = \{(0, H), (1, H), (0, M), (1, M)\}$, producing leader effects in each composition-by-disclosure environment while retaining the sample, baseline control, and fixed effects from Equation (1). This specification asks whether the two leaders' effectiveness also depends on the visibility of individual participation. We expected chiefs to be most effective in homogeneous groups with disclosure, where the group falls within their authority and they can evaluate individual behavior. Pastors should rely less on disclosure because their authority can also make participation binding through internal accountability.

5.2 Main Results

Figure 7 and Appendix Table D-1 show how leader effects across homogeneous and mixed teams vary under the two disclosure conditions. In homogeneous groups with disclosure, assigning a traditional leader increases participation by 10.6 percentage points (SE 3.5; $p = 0.002$), while the religious-leader effect is 1.7 points (SE 3.1; $p = 0.579$); the direct traditional-over-religious gap is 8.9 points (SE 3.7; $p = 0.016$). In homogeneous groups without disclosure, neither leader has a statistically significant effect relative to no leader. Pastors show the opposite pattern in mixed groups. In mixed groups with disclosure, neither leader has a statistically significant effect. The pastor's point estimate in mixed groups is larger without disclosure, although the difference is imprecise. In mixed groups without public disclosure of individual participation,

religious leadership increases participation by 9.0 percentage points (SE 3.8; $p = 0.019$) relative to having no leader. The traditional-leader effect in this environment is -1.1 points (SE 3.5; $p = 0.751$). The traditional-minus-religious gap in mixed groups without public accountability is -10.1 points (SE 4.2; $p = 0.016$). The cross-environment contrast summarizes how the relative effectiveness of chiefs and pastors changes between these two environments. It equals $\widehat{\Delta}_E = 8.9 - (-10.1) = 19.0$ percentage points (SE 5.6; $p < 0.001$).

These results reinforce the interpretation that religious leaders are weak mobilizers of homogeneous groups and that chiefs are weak mobilizers of mixed groups due to their mismatched scope. They also provide evidence that chiefs and pastors mobilize through distinct enforcement channels. This second reversal, in which pastors hold an advantage in private disclosure and chiefs in public disclosure, is difficult to reconcile with a common social mechanism that operates more strongly for one leader than the other. Chiefs in same-origin groups appear to rely on external enforcement, with residents responding to the chief's judgment of their behavior when participation is public. In contrast, pastors appear to work through internal religious accountability, using a shared Christian framework that leads residents to evaluate their own behavior before God, which becomes most relevant when their actions are private. When actions are public, pastors appeals may become less effective because behavior is more strongly influenced by the existing social norm and religious motivations become secondary. This is our *source of enforcement* result.

The response remains concentrated in labor. The environment gradient also appears across the alternative labor measures introduced in Section 4.2. It corresponds to 1.10 additional days of participation, with adjusted p-values below 0.004 across the four labor outcomes (Appendix Table D-2). Other definitions of labor contributions, including effort ratings, move in the same direction. Yet cash and material contributions are less frequent and do not show the same pattern (Appendix Table D-3).

Chiefs may become counterproductive when individual records are private. In homogeneous groups without disclosure, the chief effect falls from near zero in the main window to -7.4 percentage points in the final week (SE 4.1; Appendix Figure D-2 and Table D-5). Although imprecisely estimated, this negative effect suggests that assigning a chief may eventually reduce participation. One possible interpretation is free riding: if residents expect the chief to contribute more to the group's work, they may reduce their own effort when individual participation is not publicly disclosed.

Alternative explanations: endogeneous leader responses. Most alternative explanations that were discussed as potential accounts for earlier results could also explain these patterns. Rather than an account in which teammates respond to the reputational exposure to chiefs and intrinsic incentives under pastors, one could argue, for example, that the chiefs focus their efforts on sensitizing participants under public disclosure, while pastors would employ more effort to sensitize participants when records are not public. Hence, we revisit the earlier outcomes on the behavior of leaders to further investigate those possibilities. Panel

D of Appendix Tables B-9, C-5, and C-6 presents the results.

We find some evidence for different responses of leaders depending on the disclosure environment, but they cannot fully explain the patterns or compete with our favored interpretation. When it comes to behavior during debriefing meetings, we find that the only significant difference arises among pastors. Although they do not change the topics they raise in their speeches on average, we find that in mixed groups, pastors speak for half a minute less in a public disclosure environment compared to the average of 3.75 minutes spoken in a private disclosure environment (-0.56 minutes; $p = 0.047$). This result helps us understand the more muted impact of pastors under public disclosure, but it cannot fully account for the entire pattern of results.

Our analysis of salongo co-arrival outcomes provides little support for a leader-response account. Estimates are small and imprecise. In homogeneous groups with public disclosure, the roughly 10 percentage point increase in participation induced by chief presence includes an estimated 2.33 percentage point increase in participation accompanied by co-arrival with the chief when we compare it to how chiefs behave in the same homogeneous composition without disclosure. This provides weak evidence consistent with chiefs using public disclosure to identify and bring teammates to salongo, although co-arrival alone does not establish active mobilization. Nevertheless, the estimated increase in co-arrival is too small to account for the full participation effect.

Lastly, our analysis of the survey reports about leader contact also yields little support for a leader-driven account of participation outcomes. The only statistically significant estimate is a 14.06 percentage point ($p = 0.023$) increase in the probability of being asked by the chief to participate in mixed groups with public disclosure, relative to the 7 percent request rate in mixed groups with a chief but no disclosure. Therefore, if anything, the evidence we find is consistent with chiefs trying to compensate for their lack of influence in mixed groups when they are able to see how much each member is contributing. The pattern of contact rates does not suggest that the additional effect is due to a leader-effort account, but rather to how participants respond to leaders' presence under distinct disclosure environments.

Alternative explanations: other accounts. The extended heterogeneous treatment effects specification that compares the influence of pastors on majority-ethnicity residents still supports the conclusion that pastors' null in homogeneous groups is about group composition rather than an individual trait explanation (Appendix Table D-7). The restriction to pastors belonging to the majority ethnicity again yields no detectable pastor effect in homogeneous groups. If anything, it reveals that the null pastor effect in mixed groups with public disclosure is not robust, as estimates in the restricted sample show similar pastor effects in mixed groups under public and private disclosure (Appendix Table D-8). Leadership by example again receives little support (Appendix Table D-9). The pretreatment data show no systematic increase before the first meeting, and allowing cell-specific baseline adjustment leaves the estimates essentially unchanged (Appendix Tables D-10 and C-11). Among the 809 initially assigned residents, chiefs still raise participa-

tion in homogeneous groups with disclosure and pastors in mixed groups without disclosure, although the pooled composition contrast is less precise (Appendix Table D-11). Overall, we find no robust evidence that alternative explanations are driving our results.

5.3 Heterogeneous Treatment Effects: Who Responds To Revealed Records?

Chiefs mobilize homogeneous groups when individual contributions are public, while pastors mobilize mixed groups when records remain private. These patterns are consistent with chiefs enforcing contributions through social judgment and pastors invoking an internal obligation before God. The next comparison uses residents' participation before the meetings to see whether leaders affect those with weak and strong records differently, and whether disclosing those records changes the response.

Past participation and the response to disclosure. First meetings under the public information condition affect participants in two ways: (i) it establishes what the reporting conditions for future meetings will be, and (ii) it also provides information on who shirks and who contributes to everyone who attended the meeting. The treatment provides a signal that influences the social image of every group participant based on how much they have contributed up to that point. Shirkers will be the ones who suffer the largest damage to their standing since the group will find out they did not contribute. If chiefs operate through social judgment, the chiefs' effect should be strongest on exactly these shirkers, where the social standing of the participant is worse.

Under our interpretation, pastors' enforcement operates differently. Intrinsically religious residents will change their behavior because the pastor will remind them of their internal religious motivations. What that motivation is, whether it is a fear of divine punishment or moral convictions associated with religious practice, matters less than where the audience judging residents' behavior is located. In pastors' enforcement, the relevant audience is the self (or God), who always knows whether the resident cooperated or not, regardless of the disclosure condition. Therefore, there is no reason to believe that the pastor effect will also be concentrated among shirkers; it may very well extend across initial participation histories.

We test these predictions by extending Equation (2) from composition c to the full environment $e = (q, c)$, allowing both the category-specific no-leader levels and each leader effect to vary with residents' pretreatment participation in every composition-by-disclosure environment. Figure 8 and Appendix Table D-12 show the results for the two compositions where each leader is effective: chiefs in homogeneous groups and pastors in mixed groups. The top row plots each leader's effect relative to no leader at lower and higher levels of pretreatment participation. The bottom row shows the corresponding no-leader means.

Exposed shirkers drive the chief effect, pastor effects span participation histories. The left column of Figure 8 displays what one would expect from external social enforcement. With disclosure,

chiefs raised participation by 19.0 percentage points (SE 6.1) among residents who participated one day or fewer in the first week, compared with 1.7 points (SE 4.2) among those with higher participation. The difference is −17.3 points (SE 7.9; $p = 0.030$). Without disclosure, the corresponding effects are −5.8 and 7.9 points, and the 13.7-point difference is positive but imprecise ($p = 0.145$). Pastors' effects display no comparable difference based on prior participation. The positive chief effect is concentrated among prior nonparticipants and infrequent participants. The weak visible record gives the chief something to judge and the resident something to repair, while residents with stronger records have less room to improve.

The religious leader effect displays a completely different pattern of results. Pastor effects are positive in both baseline-participation categories when groups are mixed and individual records are not disclosed, at 6.3 points among residents with weaker prior participation (SE 5.9) and 9.4 points among those with stronger records (SE 5.8). Each subgroup estimate is imprecise, while the pooled pastor effect is positive and statistically significant in the main specification. The 3.1-point higher-minus-lower difference is also imprecise (SE 8.9; $p = 0.727$). This result strengthens the interpretation that they operate through internal incentives, but it does not explain why the pastor effect is weakened by public disclosure.

Unpacking the null: public accountability substitutes for the pastor effect. A possible explanation can be seen in the bottom panel of Figure 8. The aggregate no-leader comparison conceals differences between initial shirkers and stronger contributors. In no-leader groups, the implied participation of higher contributors is 74.4 percent without disclosure and 80.6 percent with disclosure, while the corresponding prediction falls for lower contributors. These are standardized no-leader predictions for the residents supporting each leader effect. The implied no-leader difference for higher contributors is 6.3 points (SE 5.0; $p = 0.211$), while the pastor effect falls by 11.7 points across disclosure conditions (SE 7.4; $p = 0.116$; Appendix Table D-12). These differences suggest that public accountability can substitute for internal accountability among stronger contributors, although the pastor-effect contrast is imprecise. The lower predicted participation among weaker contributors is also consistent with a less favorable response under disclosure.

5.4 Policy Implications: Hidden Levers of Fiscal Capacity

As highlighted in Olken and Singhal (2011), standard public finance accounts miss an important source of developing states' fiscal capacity when they omit informal taxation. Our collaboration allows us to partially account for this source of capacity in the studied sample, as described in Appendix D.5.

Our study revealed a surprising demand among citizens for the arrangements facilitated by the city program. Once the city provided basic tools and organizational capacity for neighborhood improvement, residents volunteered their labor. Across the 28 campaign neighborhoods, residents contributed 70,250 hours and completed an estimated 832 worksites. The labor amounts to about 8,800 eight-hour workdays, valued at $26,344 using the historical STEP II public-works rate of $3 per day (Agence Congolaise de Presse, 2021).

Including recorded cash and in-kind contributions brings the total to $27,890—30.5% of Kasaï-Central's reported 2025 property-tax receipts (Direction Générale des Recettes du Kasaï Central, 2026) (Figure 9).

Incorporating our findings into a rough counterfactual that assigns chiefs with public disclosure to homogeneous groups and pastors without disclosure to mixed groups, the program could have mobilized 7,251 additional hours among the 1,040 experimental residents relative to no leaders under the same disclosure conditions. Assuming the six-day effects persist throughout the campaign, this is equivalent to about 75 average worksite workloads. This exercise highlights both the scale of hidden capacity through mobilizing residents and the scope for improvement when leadership and disclosure conditions align with the enforcement channels uncovered by this experiment.

6 Beyond Participation: What Else Changes?

Our results establish when leaders can promote collective action and we develop an interpretation of why they succeed. What are the consequences of this success? The main behavioral outcomes are observed during a short window of time, but sustained cooperation may reshape the social and political views of participants, as well as their relationships. We turn to investigating these possibilities by examining beliefs about cooperation and accountability, the salience of religious and social identities, views of government and other local institutions, and opinions about other members of the group. For the first set of outcomes measured at the individual level, we use the specification in Equation (1) to assess leader treatment effects in different group environments.

More participation does not change beliefs or political attitudes. The first set of outcomes captures beliefs about social and divine accountability and the identities residents consider most important. The second covers government services and accountability, institutional performance, trust in state and non-state actors, and evaluations of the locality chief. Across these outcomes, we find no robust patterns. Appendix Tables E-1, E-2, and E-3 report the full results. Additional participation therefore changed behavior without broadly changing residents' beliefs, identities, or political attitudes.

Sustained cooperation precedes greater willingness to help one another. The clearest changes concern relationships within the group. We compare residents' evaluations of ordinary peers and assigned leaders before the first discussion and at the start of the second meeting. The outcome is the standardized change in these evaluations. To distinguish how leader presence affects evaluations of ordinary peers from evaluations of the leader, we extend equation (1) as follows. Let j index the person being evaluated and let $J_{jg} = 1$ indicate that this person is the assigned leader:

$$Y_{ij} = \lambda_{n(i),q(i),c(i)} + \sum_{e \in \mathcal{E}} \Big[\left(\tau_e^T + \kappa_e^T J_{jg}\right) T_{g(i)} + \left(\tau_e^R + \kappa_e^R J_{jg}\right) R_{g(i)} \Big] \mathbf{1}\{e(i) = e\} + X_i'\beta + \varepsilon_{ij}. \tag{3}$$

Here Y_{ij} is the standardized change in an individual i's evaluation of person j. For leader type $L \in \{T, R\}$, τ_e^L captures the effect of leader assignment on evaluations of ordinary peers in environment e, while $\tau_e^L + \kappa_e^L$ captures the effect on evaluations of the assigned leader. Appendix Tables E-4 and E-5 report all six outcomes across the four experimental environments. Figure 10 focuses on willingness to help, showing separately the environment in which each leader increased participation and pooling the other three environments.

The effects are sizable where leaders sustain participation. In homogeneous groups with disclosure, assigning a chief increases residents' willingness to help an ordinary participant by 0.19 standard deviations. At the average baseline answer of about 4,600 Congolese francs, this effect corresponds to roughly 1,400 francs more willingness to help, or 30 percent of the baseline mean. In mixed groups without disclosure, assigning a pastor increases the willingness to help an ordinary participant by 0.35 standard deviations and the willingness to help the pastor by 0.34 standard deviations. These effects correspond to about 2,760 and 2,680 francs, or 60 and 58 percent of the baseline means. We do not observe systematic changes in the other evaluations of peers. Residents seem more willing to help people with whom they have sustained cooperation, even though they do not come to view those peers more favorably in general.

Pastors gain support when they sustain cooperation. The pattern is different for peers and leaders. Chief assignment mainly increases the willingness to help ordinary peers, while evaluations of the chief change little. Pastor assignment increases the willingness to help both ordinary participants and the pastor himself by about the same amount. This difference fits their initial positions. Chiefs already hold established authority among coethnics, while pastors enter mixed groups without the same standing and can gain it when they successfully sustain cooperation. Religious leaders can therefore leave divided groups more cohesive while also gaining greater support from residents.

7 Conclusion

What can unite ethnically divided societies around collective goals when social enforcement breaks down? In Kananga, we have found that religious authorities succeed in promoting collective action in ethnically heterogeneous groups where traditional authorities have failed. In a large-scale field experiment that involved more than a thousand volunteers and hundreds of small neighborhood improvement projects, they not only succeeded in promoting the participation of fragmented teams; they were especially likely to succeed when information about participation was unavailable and social enforcement was the hardest. Our analysis, therefore, points to two reasons that explain their success: the authority they invoke can work across ethnic relationships, and enforcement occurs internally rather than being dependent on social sanctions.

Perhaps more surprising is the result that religious leaders *fail* to sustain high levels of cooperation precisely in settings where traditional leaders succeed. When the audience is ethnically cohesive *or* public accountability shapes incentives, religious leaders are less effective in inducing cooperation. Evidence suggests that ethnic cohesion limits the influence of religious leaders because ethnic identity organizes cooperation as an ethnically enforced obligation, limiting the scope for framing the activity as a religious duty. Public accountability, in turn, appears to limit the influence of the self-accountability activated by religious moral appeals. Unlike traditional leaders, who display patterns consistent with an authority that depends on social monitoring and is rooted in ties to each resident that may or may not be ethnic, the authority of religious leaders appears to rely on preexisting beliefs and on a setting that allows those beliefs to influence behavior.

These findings expose two dimensions of authority that are rarely studied in tandem and show why future considerations of these dimensions should matter for understanding how collective action operates. Its *scope* determines whom a leader can mobilize; its *source* determines what makes participation incentive compatible—evaluation by others, obligations individuals apply to themselves, or both. Collective action depends on the fit between the leader in charge, the composition of the audience, and whether the enforcement technology is appropriate for the setting. Although extending these considerations to other leaders and institutions requires further research, the implications of this theory of change could apply far beyond religion, customs, and the DRC. More broadly, the same logic becomes relevant where a narrow identity sustained by social accountability coexists with a more encompassing identity whose obligations are internally enforced.

Therefore, beyond the policy implications of the studied framework, which can be applied to any development policy that uses local leaders as intermediaries in program implementation, the paper opens important research avenues by revisiting the relationship between ethnic fragmentation and collective action. Indeed, our paper confirms that fragmentation is a corrosive force for the *social enforcement* of collective action. Yet, by corroding social enforcement, something else also happens: internal motives begin to enforce conduct. Fragmentation does not simply corrode collective action; it changes which institutions can govern behavior. Religion is one widespread institution that enables cooperation in divided settings, potentially among many. What else can govern cooperation when social enforcement breaks down remains the subject of future research.

References Cited

Acemoglu, Daron, Tristan Reed, and James A. Robinson, "Chiefs: Economic Development and Elite Control of Civil Society in Sierra Leone," *Journal of Political Economy*, 2014, *122* (2), 319–368.

Actualite.cd, "Kin Bopeto : les quartiers les plus propres de la capitale seront primés à partir de la semaine prochaine," Actualite.cd October 2019.

Agence Congolaise de Presse, "Kasaï Central : le Projet STEP II construit 27 écoles à Kananga," January 2021. January 22. https://acp.cd/economie/kasai-central-le-projet-step-ii-construit-27-ecoles-a-kananga/.

Akerlof, George A. and Rachel E. Kranton, "Economics and Identity," *The Quarterly Journal of Economics*, 2000, *115* (3), 715–753.

Alesina, Alberto and Eliana La Ferrara, "Participation in Heterogeneous Communities," *The Quarterly Journal of Economics*, 2000, *115* (3), 847–904.

__ and __, "Ethnic Diversity and Economic Performance," *Journal of Economic Literature*, 2005, *43* (3), 762–800.

__, Reza Baqir, and William Easterly, "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 1999, *114* (4), 1243–1284.

Allport, Gordon W. and J. Michael Ross, "Personal Religious Orientation and Prejudice," *Journal of Personality and Social Psychology*, 1967, *5* (4), 432–443.

Andreoni, James and B. Douglas Bernheim, "Social Image and the 50–50 Norm: A Theoretical and Experimental Analysis of Audience Effects," *Econometrica*, 2009, *77* (5), 1607–1636.

Auriol, E., J. Lassébie, A. Panin, E. Raiber, and P. Seabright, "God Insures Those Who Pay? Formal Insurance and Religious Offerings in Ghana," *Quarterly Journal of Economics*, 2020, *135* (4), 1799–1848.

Baldwin, Kate, *The Paradox of Traditional Chiefs in Democratic Africa*, Cambridge: Cambridge University Press, 2016.

__ and John D. Huber, "Economic versus Cultural Differences: Forms of Ethnic Diversity and Public Goods Provision," *American Political Science Review*, 2010, *104* (4), 644–662.

Balán, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel, "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo," *American Economic Review*, March 2022, *112* (3), 762–97.

Banerjee, Abhijit V, Sylvain Chassang, Sergio Montero, and Erik Snowberg, "A theory of experimenters: Robustness, randomization, and balance," *American Economic Review*, 2020, *110* (4), 1206–1230.

Bazzi, Samuel, Arya Gaduh, Alexander D. Rothenberg, and Maisy Wong, "Unity in Diversity? How Intergroup Contact Can Foster Nation Building," *American Economic Review*, 2019, *109* (11), 3978–4025.

__, Gabriel Koehler-Derrick, and Benjamin Marx, "The Institutional Foundations of Religious Politics: Evidence from Indonesia," *The Quarterly Journal of Economics*, 2020, *135* (2), 845–911.

Becker, Sascha O., Jared Rubin, and Ludger Woessmann, "Religion and Growth," *Journal of Economic Literature*, 2024, *62* (3), 1094–1142.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-Dimensional Methods and Inference on Structural and Treatment Effects," *Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

Bénabou, Roland and Jean Tirole, "Incentives and Prosocial Behavior," *American Economic Review*, 2006, *96* (5), 1652–1678.

__ and __, "Mindful Economics: The Production, Consumption, and Value of Beliefs," *Journal of Economic Perspectives*, 2016, *30* (3), 141–164.

Berman, Eli, "Sect, Subsidy, and Sacrifice: An Economist's View of Ultra-Orthodox Jews," *The Quarterly Journal of Economics*, 2000, *115* (3), 905–953.

Blei, David M. and John D. Lafferty, "A Correlated Topic Model of Science," *Annals of Applied Statistics*, 2007, *1* (1), 17–35.

__, Andrew Y. Ng, and Michael I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, 2003, *3*, 993–1022.

Boudreau, Laura, Rocco Macchiavello, Virginia Minni, and Mari Tanaka, "Leaders in Social Movements: Evidence from Unions in Myanmar," *American Economic Review*, 2025, *115* (6), 1975–2000.

Brandily, Paul, Eric Mvukiyehe, Lodewijk Smets, Peter van der Windt, and Marijke Verpoorten, "From Workfare to Economic and Sociopolitical Stability? Evidence from a Randomized Trial in Eastern Congo," *The World Bank Economic Review*, 2025, *39* (3), 711–730.

Bryan, Gharad, James J. Choi, and Dean Karlan, "Randomizing Religion: The Impact of Protestant Evangelism on Economic Outcomes," *The Quarterly Journal of Economics*, 2021, *136* (1), 293–380.

Buccione, Giulia, "Religious Messaging and Adaptation to Water Scarcity: Evidence from Jordan," 2023. Working paper.

Callaghy, Thomas M., *The State–Society Struggle: Zaire in Comparative Perspective*, New York: Columbia University Press, 1984.

Casey, Katherine, Rachel Glennerster, Edward Miguel, and Maarten Voors, "Skill versus Voice in Local Development," *Review of Economics and Statistics*, 2023, *105* (2), 311–326.

Chaney, Eric, "Revolt on the Nile: Economic Shocks, Religion, and Political Power," *Econometrica*, 2013, *81* (5), 2033–2053.

Chen, Daniel L., "Club Goods and Group Identity: Evidence from Islamic Resurgence during the Indonesian Financial Crisis," *Journal of Political Economy*, 2010, *118* (2), 300–354.

de la Sierra, Raúl Sánchez, "Whither Formal Contracts?," *Econometrica*, 2021, *89* (5), 2341–2373.

Dell, Melissa and Ashesh Rambachan, "The Measurement Revolution? Credible Measurement and Inference in the Age of AI," Working Paper 35744, National Bureau of Economic Research 2026.

Depetris-Chauvin, Emilio, Ruben Durante, and Filipe Campante, "Building Nations through Shared Experiences: Evidence from African Football," *American Economic Review*, 2020, *110* (5), 1572–1602.

Devisch, René, "'Pillaging Jesus': Healing Churches and the Villagisation of Kinshasa," *Africa*, 1996, *66* (4), 555–586.

Dhillon, Inderjit S. and Dharmendra S. Modha, "Concept Decompositions for Large Sparse Text Data Using Clustering," *Machine Learning*, 2001, *42*, 143–175.

Direction Générale des Recettes du Kasaï Central, "Rapport des recettes réalisées par service et par acte générateur," Division de Recouvrement, Kananga January 2026. Reported 2025 receipts; property-tax line. Administrative report and accompanying fiscal workbook.

Dunia, Louis, Aimable Amani Lameke, Nathan Nunn, Max Posch, and Raúl Sánchez de la Sierra, "On the Importance of African Traditional Religion for Economic Behavior," NBER Working Paper 31430, National Bureau of Economic Research 2023.

Easterly, William and Ross Levine, "Africa's Growth Tragedy: Policies and Ethnic Divisions," *Quarterly Journal of Economics*, 1997, *112* (4), 1203–1250.

Enke, Benjamin, "Kinship, Cooperation, and the Evolution of Moral Systems," *Quarterly Journal of Economics*, 2019, *134* (2), 953–1019.

Falk, Armin and Michael Kosfeld, "The Hidden Costs of Control," *American Economic Review*, 2006, *96* (5), 1611–1630.

Fehr, Ernst and Bettina Rockenbach, "Detrimental Effects of Sanctions on Human Altruism," *Nature*, 2003, *422* (6928), 137–140.

Fracchia, Mattia, Teresa Molina-Millán, and Pedro C. Vicente, "Motivating Volunteer Health Workers in an African Capital City," *Journal of Development Economics*, 2023, *163*, 103096.

Freeman, Dena, ed., *Pentecostalism and Development*, London: Palgrave Macmillan UK, 2012.

Gentzkow, Matthew, Bryan Kelly, and Matt Taddy, "Text as Data," *Journal of Economic Literature*, 2019, *57* (3), 535–574.

Ghosh, Arkadev, "Religious Divisions and Production Technology: Experimental Evidence from India," *Journal of Political Economy*, 2025, *133* (10), 3249–3304.

_ , Prerna Kundu, Matt Lowe, and Gareth Nellis, "Creating Cohesive Communities: A Youth Camp Experiment in India," *The Review of Economic Studies*, 2026, *93* (1), 438–475.

Glennerster, Rachel, Edward Miguel, and Alexander D. Rothenberg, "Collective Action in Diverse Sierra Leone Communities," *The Economic Journal*, 2013, *123* (568), 285–316.

Gluckman, Max, "Analysis of a Social Situation in Modern Zululand," *Bantu Studies*, 1940, *14* (1), 1–30.

Gneezy, Ayelet, Uri Gneezy, Gerhard Riener, and Leif D. Nelson, "Pay-What-You-Want, Identity, and Self-Signaling in Markets," *Proceedings of the National Academy of Sciences*, 2012, *109* (19), 7236–7240.

Gobbers, Erik, "Ethnic Associations in Katanga Province, the Democratic Republic of Congo: Multi-tier System, Shifting Identities and the Relativity of Autochthony," *The Journal of Modern African Studies*, 2016, *54* (2), 211–236.

Goette, Lorenz, David Huffman, and Stephan Meier, "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence Using Random Assignment to Real Social Groups," *American Economic Review*, 2006, *96* (2), 212–216.

Gorsuch, Richard L. and Susan E. McPherson, "Intrinsic/Extrinsic Measurement: I/E-Revised and Single-Item Scales," *Journal for the Scientific Study of Religion*, 1989, *28* (3), 348–354.

Grootendorst, Maarten, "BERTopic: Neural Topic Modeling with a Class-Based TF-IDF Procedure," arXiv:2203.05794 2022.

Habyarimana, James, Macartan Humphreys, Daniel N. Posner, and Jeremy M. Weinstein, "Why Does Ethnic Diversity Undermine Public Goods Provision?," *American Political Science Review*, 2007, *101* (4), 709–725.

Henn, Soeren J., Gauthier Marchais, Christian Mastaki Mugaruka, and Raúl Sánchez de la Sierra, "From Indirect to Direct Rule: Armed Groups and Customary Chiefs in Eastern Congo," 2026. Working paper. Earlier version circulated as NBER Working Paper 32722.

Henrich, Joseph, Jean Ensminger, Richard McElreath, Abigail Barr, Clark Barrett, Alexander Bolyanatz, Juan Camilo Cardenas, Michael Gurven, Edwins Gwako, Natalie Henrich, Carolyn

Lesorogol, Frank Marlowe, David Tracer, and John Ziker, "Markets, Religion, Community Size, and the Evolution of Fairness and Punishment," *Science*, 2010, *327* (5972), 1480–1484.

Hermalin, Benjamin E., "Toward an Economic Theory of Leadership: Leading by Example," *American Economic Review*, 1998, *88* (5), 1188–1206.

Hjort, Jonas, "Ethnic Divisions and Production in Firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hoffmann, Kasper, Godefroid Kihangu Muzalia, César Muhigirwa Tungali, and Alice Mugoli Nalunva, "The Past in the Present: Ethnicity, Conflict and Politics in Eastern Congo," Insecure Livelihoods Series, Governance in Conflict Network, Ghent University May 2022.

Iannaccone, Laurence R., "Sacrifice and Stigma: Reducing Free-riding in Cults, Communes, and Other Collectives," *Journal of Political Economy*, 1992, *100* (2), 271–291.

Jack, B. Kelsey and Maria P. Recalde, "Leadership and the Voluntary Provision of Public Goods: Field Evidence from Bolivia," *Journal of Public Economics*, 2015, *122*, 80–93.

Kapepula, Gilbert Tshiebue, Max Mbosho Konshi, and Jonathan L. Weigel, "Prosociality and Pentecostalism in the D.R. Congo," *Religion, Brain & Behavior*, 2022, *12* (1–2), 150–170.

Kosfeld, M. and D. Rustagi, "Leader Punishment and Cooperation in Groups: Experimental Field Evidence from Commons Management in Ethiopia," *American Economic Review*, 2015, *105* (2), 747–783.

Laitin, David D., *Hegemony and Culture: Politics and Religious Change among the Yoruba*, Chicago: University of Chicago Press, 1986.

Lang, Martin, Benjamin G. Purzycki, Coren L. Apicella, Quentin D. Atkinson, Alexander Bolyanatz, Emma Cohen, Carla Handley, Eva Kundtová Klocová, Carolyn Lesorogol, Sarah Mathew, Rita A. McNamara, Cristina Moya, Caitlyn D. Placek, Montserrat Soler, Thomas Vardy, Jonathan L. Weigel, Aiyana K. Willard, Dimitris Xygalatas, Ara Norenzayan, and Joseph Henrich, "Moralizing gods, impartiality and religious parochialism across 15 societies," *Proceedings of the Royal Society B: Biological Sciences*, 2019, *286* (1898), 20190202.

Lowe, M., "Has Intergroup Contact Delivered?," *Annual Review of Economics*, 2025, *17*, 321–344.

Lowe, Matt, "Types of Contact: A Field Experiment on Collaborative and Adversarial Caste Integration," *American Economic Review*, 2021, *111* (6), 1807–1844.

Lowes, Sara, Benjamin Marx, and Eduardo Montero, "Religion in Emerging and Developing Regions," *Annual Review of Economics*, 2026, *18*.

_ , Nathan Nunn, James A. Robinson, and Jonathan Weigel, "Understanding Ethnic Identity in Africa: Evidence from the Implicit Association Test (IAT)," *American Economic Review: Papers & Proceedings*, 2015, *105* (5), 340–345.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and Team Performance in a Kenyan Organization," *Journal of Public Economics*, 2021, *197*, 104332.

McCauley, John F., "The Political Mobilization of Ethnic and Religious Identities in Africa," *American Political Science Review*, 2014, *108* (4), 801–816.

McClendon, Gwyneth and Rachel Beatty Riedl, "Religion as a Stimulant of Political Participation: Experimental Evidence from Nairobi, Kenya," *The Journal of Politics*, 2015, *77* (4), 1045–1057.

McClendon, Gwyneth H. and Rachel Beatty Riedl, *From Pews to Politics: Religious Sermons and Political Participation in Africa*, Cambridge: Cambridge University Press, 2019.

Mercy Corps, "Kasaï Conflict Assessment: February–March 2019," Mercy Corps Research Report June 2019.

Michalopoulos, Stelios and Elias Papaioannou, "Historical Legacies and African Development," *Journal of Economic Literature*, 2020, *58* (1), 53–128.

Miguel, Edward, "Tribe or Nation? Nation Building and Public Goods in Kenya versus Tanzania," *World Politics*, 2004, *56* (3), 327–362.

_ and Mary Kay Gugerty, "Ethnic Diversity, Social Sanctions, and Public Goods in Kenya," *Journal of Public Economics*, 2005, *89* (11-12), 2325–2368.

Ministère de l'Habitat et du Développement Urbain, "Concours national "Villes Propres" : pour des villes camerounaises résilientes et durables," Government webpage 2022.

Mousa, Salma, "Building Social Cohesion between Christians and Muslims through Soccer in Post-ISIS Iraq," *Science*, 2020, *369* (6505), 866–870.

Ngalamulume, Kalala, "Mythe, Politique et Histoire: Le Mythe De Mande Katawa Chez Les "Luluwa" Du Kasayi," *History in Africa*, 1995, *22*, 329–347.

Ngoma Mavungu, Marina, Clara Sievert, Xavier Jaravel, Nathan Nunn, and Jonathan L. Weigel, "How Market Access Shapes Wellbeing and Values: Experimental Evidence from the D.R. Congo," 2025. Working paper, November 10, 2025.

Norenzayan, Ara, *Big Gods: How Religion Transformed Cooperation and Conflict*, Princeton: Princeton University Press, 2013.

Olken, Benjamin A. and M. Singhal, "Informal Taxation," *American Economic Journal: Applied Economics*, 2011, *3* (4), 1–28.

Ostrom, Elinor, *Governing the Commons: The Evolution of Institutions for Collective Action*, Cambridge: Cambridge University Press, 1990.

_ , "Collective Action and the Evolution of Social Norms," *Journal of Economic Perspectives*, 2000, *14* (3), 137–158.

Pew Research Center, "Christian Population as Percentages of Total Population by Country," December 2011.

Purzycki, Benjamin Grant, Coren Apicella, Quentin D. Atkinson, Emma Cohen, Rita Anne McNamara, Aiyana K. Willard, Dimitris Xygalatas, Ara Norenzayan, Joseph Henrich et al., "Moralistic Gods, Supernatural Punishment and the Expansion of Human Sociality," *Nature*, 2016, *530* (7590), 327–330.

Pype, Katrien, *The Making of the Pentecostal Melodrama: Religion, Media and Gender in Kinshasa*, Vol. 6 of *Anthropology of Media*, New York: Berghahn Books, 2012.

Quinn, Kevin M., Burt L. Monroe, Michael Colaresi, Michael H. Crespin, and Dragomir R. Radev, "How to Analyze Political Attention with Minimal Assumptions and Costs," *American Journal of Political Science*, 2010, *54* (1), 209–228.

Rao, Gautam, "Familiarity Does Not Breed Contempt: Generosity, Discrimination, and Diversity in Delhi Schools," *American Economic Review*, 2019, *109* (3), 774–809.

République Démocratique du Congo, "Loi organique n° 08/016 du 7 octobre 2008 portant composition, organisation et fonctionnement des entités territoriales décentralisées et leurs rapports avec l'Etat et les provinces," October 2008. Journal officiel de la République Démocratique du Congo, numéro spécial, 10 octobre 2008.

_ , "Loi organique n° 10/011 du 18 mai 2010 portant fixation des subdivisions territoriales à l'intérieur des provinces," May 2010. Journal officiel de la République Démocratique du Congo, 7 juin 2010.

République du Zaïre, "Ordonnance-loi n° 82-006 du 25 février 1982 portant organisation territoriale, politique et administrative de la République," February 1982. Journal officiel de la République du Zaïre.

Robbins, Joel, "The Globalization of Pentecostal and Charismatic Christianity," *Annual Review of Anthropology*, 2004, *33*, 117–143.

Robinson, James A., "Tax Aversion and the Social Contract in Africa," *Journal of African Economies*, 2023, *32* (Supplement_1), i33–i56.

Roy, Moumita and Daniel Houser, "Identity, Leadership, and Cooperation: An Experimental Analysis," *European Economic Review*, 2024, *165*, 104741.

Rwanda Governance Board, "Impact Assessment of Umuganda: 2007–2016," Technical Report, Rwanda Governance Board, Kigali October 2017.

Sambanis, Nicholas and Moses Shayo, "Social Identification and Ethnic Conflict," *American Political Science Review*, 2013, *107* (2), 294–325.

Scacco, Alexandra and Shana S. Warren, "Can Social Contact Reduce Prejudice and Discrimination? Evidence from a Field Experiment in Nigeria," *American Political Science Review*, 2018, *112* (3), 654–677.

Seay, Laura E., "Effective Responses: Protestants, Catholics and the Provision of Health Care in the Post-War Kivus," *Review of African Political Economy*, 2013, *40* (135), 83–97.

Shariff, Azim F. and Ara Norenzayan, "God Is Watching You: Priming God Concepts Increases Prosocial Behavior in an Anonymous Economic Game," *Psychological Science*, 2007, *18* (9), 803–809.

Stearns, Jason, "Setting Fire to Your Own House: Crisis in the Kasai: The Manipulation of Customary Power and the Instrumentalization of Disorder," Technical Report, Congo Research Group, Center on International Cooperation, New York University July 2018.

Tempels, Placide, *Bantu Philosophy*, 2 ed., Paris: Présence Africaine, 1959. Originally published in French in 1945.

Titeca, Kristof, Tom De Herdt, and Inge Wagemakers, "God and Caesar in the Democratic Republic of Congo: Negotiating Church-State Relations through the Management of School Fees in Kinshasa's Catholic Schools," *Review of African Political Economy*, 2013, *40* (135), 116–131.

Tonda, Joseph, *Le Souverain moderne: Le corps du pouvoir en Afrique centrale (Congo, Gabon)* Hommes et sociétés, Paris: Karthala, 2005.

Townsend, Robert M., "Risk and Insurance in Village India," *Econometrica*, 1994, *62* (3), 539–591.

van den Boogaard, Vanessa, Yannick Lokaya Bokasola, Gayatri Sahgal, and Caleb Jérémie Dohou, "Beyond the State, but Reinforcing It? Informal Tax Institutions and State Legitimacy in the Democratic Republic of Congo," WIDER Working Paper 2026/12, UNU-WIDER, Helsinki February 2026.

Vert Togo, "Togo/Concours 'salubrité challenge des quartiers' : près de 500 rues de Lomé assainies," Online news article March 2019.

Wang, Liang, Nan Yang, Xiaolong Huang, Linjun Yang, Rangan Majumder, and Furu Wei, "Multilingual E5 Text Embeddings: A Technical Report," arXiv:2402.05672 2024.

World Bank, "Kananga Emergency Urban Resilience Project (Project ID: P179292)," World Bank Projects & Operations 2022. Retrieved March 19, 2025 from https://projects.worldbank.org/en/projects-operations/project-detail/P179292.

World Christian Database, "World Christianity 2025: Regional Perspectives," *International Bulletin of Mission Research*, 2025, *49* (1), 62–74.

World Health Organization, "Ebola disease caused by Bundibugyo virus, Democratic Republic of the Congo," Disease Outbreak News May 2026. Accessed 19 June 2026.

Figures

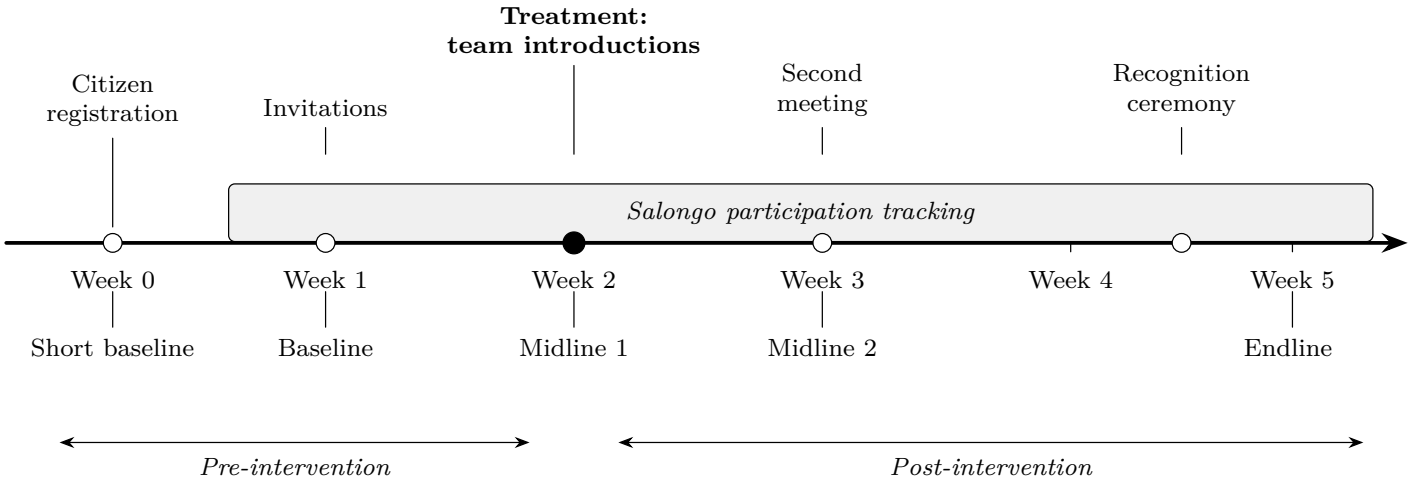
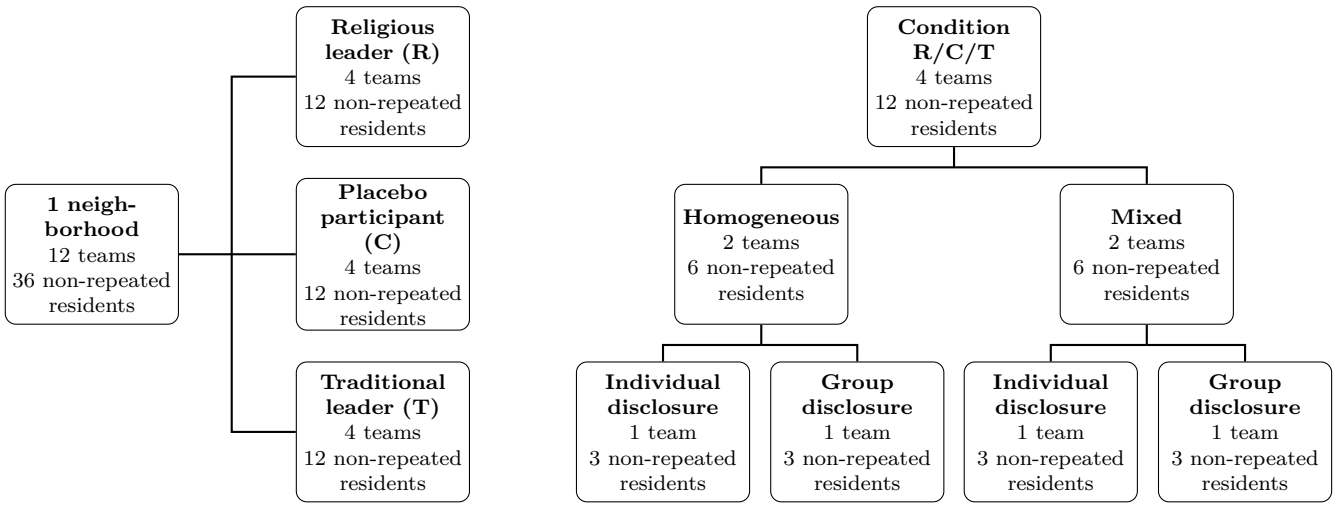


Figure 1: INTERVENTION AND DATA COLLECTION TIMELINE

Notes: The figure shows the sequence of activities during one five-week campaign wave. Residents registered before the city began tracking daily *salongo* participation. Invitations and the baseline survey followed registration. At the first meeting, residents met their assigned teammates and learned the team's leader assignment and whether individual contributions would be disclosed, marking the start of the intervention. The same team met again one week later. The city held the recognition ceremony near the end of the tracking period, and the endline followed. Survey timing appears below the timeline. The intervention is described in Section 3; Appendix Figure D-1 shows the outcome windows used in the analysis.



(a) Leader experiment **(b) Within-leader variation**

Figure 2: EXPERIMENTAL DESIGN

Notes: Each neighborhood contains twelve teams. Panel (a) assigns four teams to a traditional leader, four to a religious leader, and four to no leader. Panel (b) divides the four teams in each condition evenly between homogeneous and mixed groups and between individual and group disclosure. The diagram shows the standard four-person team used from Wave 3 onward. The slightly different no-leader roster used in Waves 1–2 is explained in Appendix Tables B-5–B-7. The figure is discussed in Section 3; balance and implementation checks appear in Appendix Tables B-4–B-7.

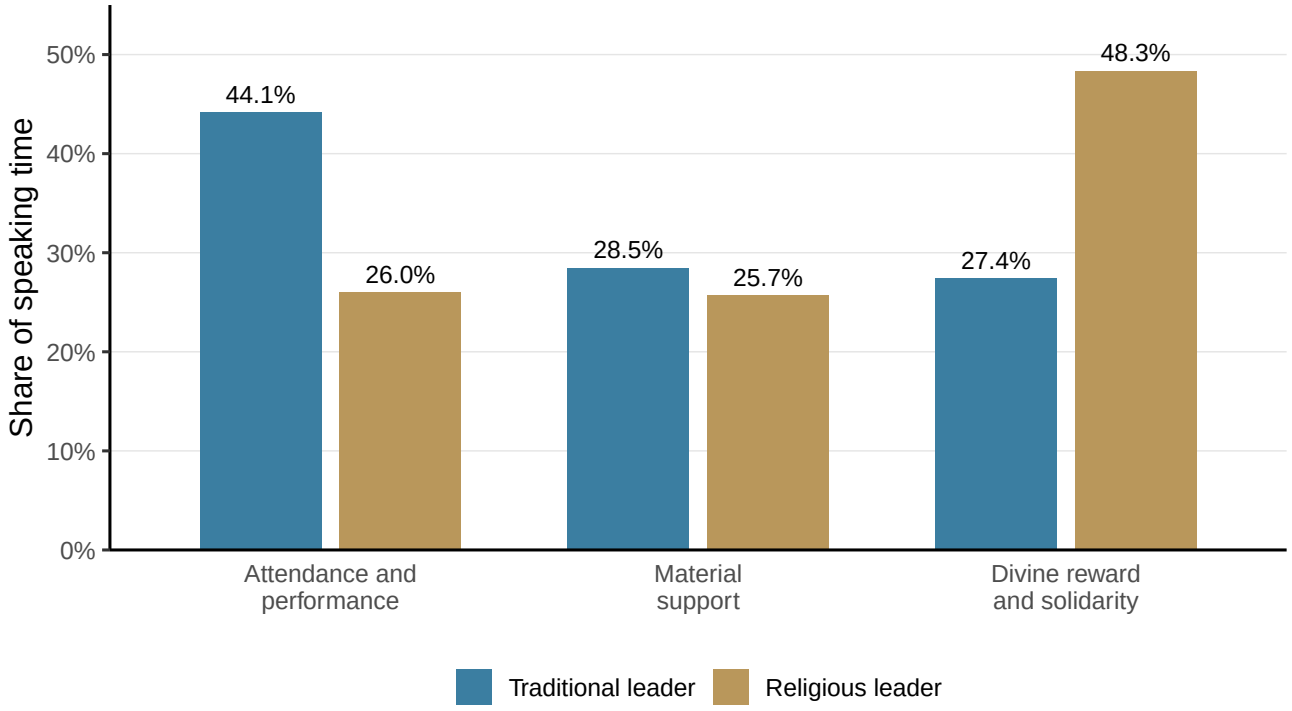


Figure 3: TRADITIONAL LEADERS STRESS PERFORMANCE; RELIGIOUS LEADERS STRESS DIVINE REWARD

Notes: Bars show the mean share of speaking time devoted to each topic during the first recorded meeting, by leader type and pooling the four experimental environments. Each classified passage is assigned to one topic and weighted by its duration. Shares are calculated within each leader-meeting observation and then averaged, giving each observation equal weight. The three shares sum to 100 percent within leader type. Blue bars report traditional leaders and gold bars religious leaders. Section 3.5 discusses the figure. Appendix Table B-9 reports speaking time and topic shares separately by environment; Appendix Table B-8 gives characteristic words and translated examples.

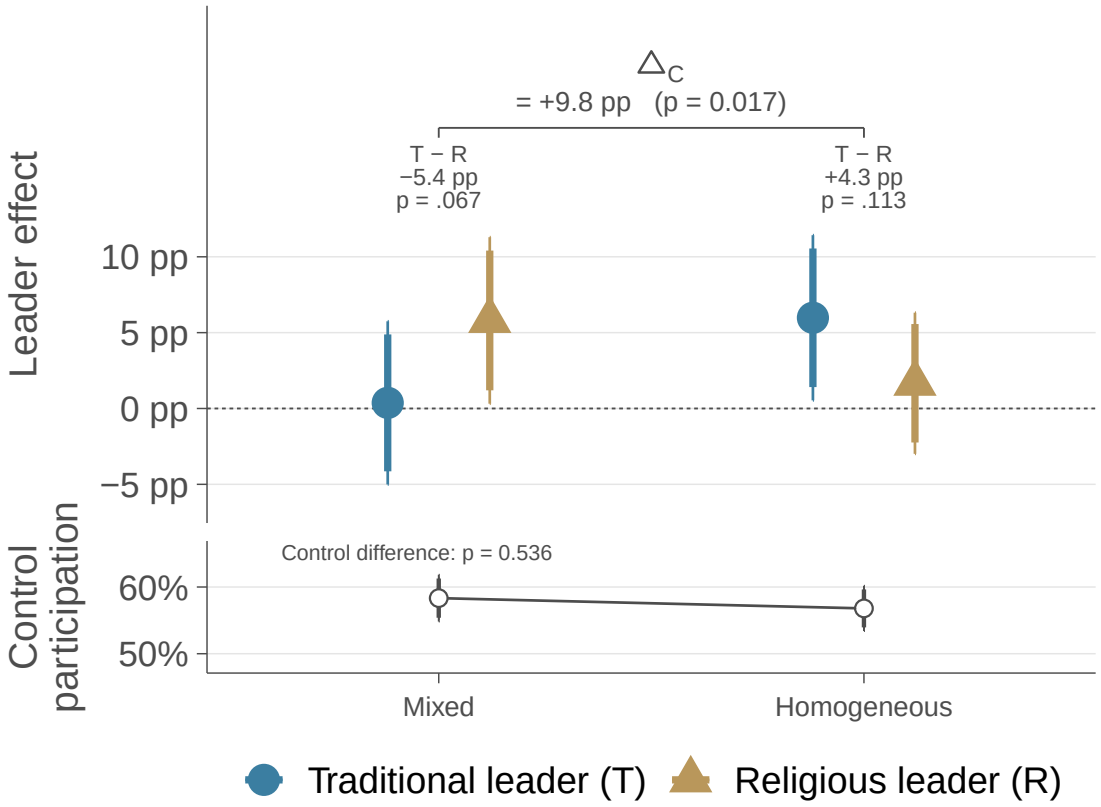


Figure 4: LEADER EFFECTS BY GROUP COMPOSITION

Notes: Blue circles and gold triangles show how assigning a traditional leader or religious leader changes participation relative to no leader, pooling across disclosure conditions. The upper panel compares mixed and homogeneous groups; T–R labels compare the leaders directly, and Δ_C measures the change in this gap from mixed to homogeneous groups. Open circles below show no-leader participation. Thick and thin bars are 90 and 95 percent confidence intervals. Estimates adjust for pretreatment participation, 25-meter spillover exposure, and 11 LASSO-selected baseline controls and compare groups within the same neighborhood, composition, and disclosure condition. Section 4.2 discusses the figure; Appendix Table C-1 reports the estimates and specification checks.

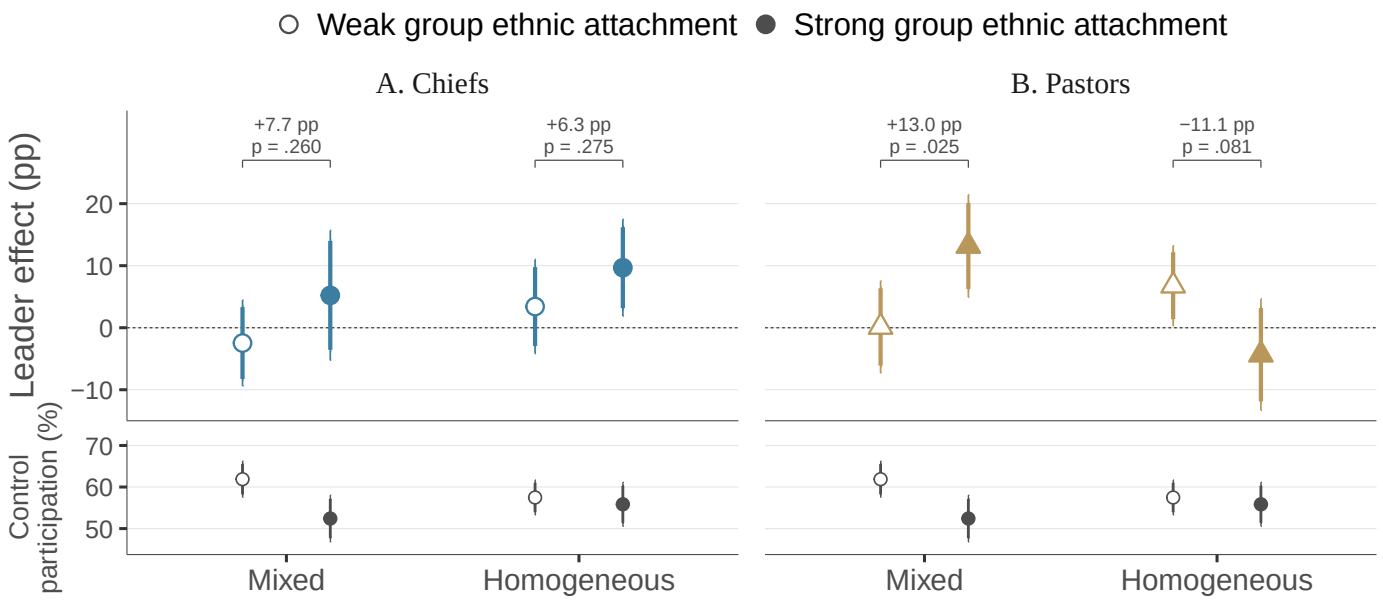


Figure 5: GROUP ETHNIC ATTACHMENT AND LEADER EFFECTIVENESS

Notes: Both panels compare groups above and at or below the group median of ethnic attachment, 10/3. Before treatment, residents ranked their identities; own-territory scores run from 1–6, with higher scores indicating greater importance. We average the available scores of ordinary members, excluding leaders: three in 330 groups and two in six (1,040 residents). Open and filled markers show the lower and higher categories. Lower plots average fitted predictions over actual no-leader residents within each observed composition and attachment category. Blue circles in Panel A report traditional leader effects and gold triangles in Panel B report religious leader effects relative to no leader, pooling disclosure. Thick and thin bars are 90 and 95 percent group-clustered confidence intervals. Estimates include baseline participation, 25-meter spillover exposure, 11 LASSO-selected baseline controls, and neighborhood-by-composition-by-disclosure fixed effects. Annotations above pairs report higher-minus-lower differences. Section 4.3 discusses the figure; Appendix Table C-12 reports both leader types and both specifications.

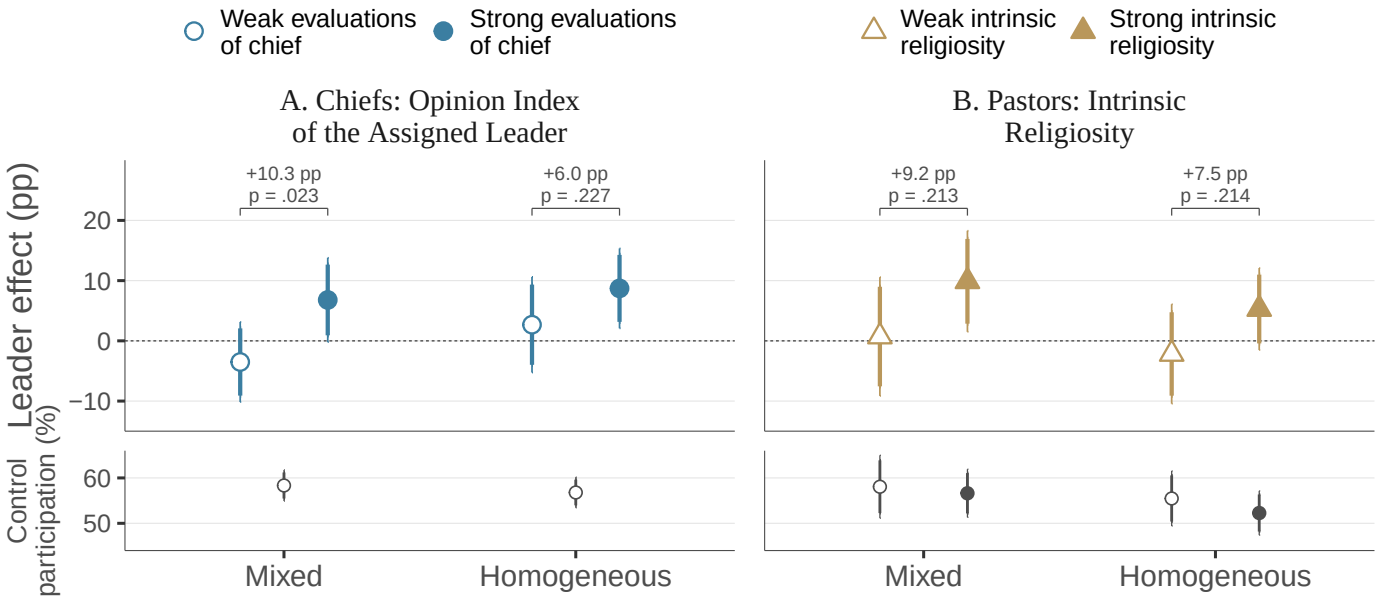


Figure 6: CHANNELS OF AUTHORITY AND LEADER EFFECTIVENESS

Notes: Panel A compares residents at or below the median opinion index with those above it. The index combines pre-discussion ratings of the assigned leader's trustworthiness, hard work, morality, and inspiration. Missing values for 33 residents are filled with their group's observed mean; three residents without an observed group score are excluded. Panel B compares residents above and at or below 3.2 on intrinsic religiosity, the cutpoint that best balances category sizes (498 lower, 536 higher). We measure intrinsic religiosity from five questions, asked at baseline, about faith as a guide to life. After reversing two items, we take the average for residents with at least three valid answers, leaving 1,034 residents. Open and filled markers show the lower and higher categories. Lower plots pool no-leader predictions within composition in Panel A, where assessments are undefined without a leader; Panel B uses common no-leader predictions. Blue circles in Panel A report traditional leader effects and gold triangles in Panel B report religious leader effects relative to no leader, pooling disclosure. Thick and thin bars are 90 and 95 percent group-clustered confidence intervals. Estimates include baseline participation, 25-meter spillover exposure, 11 LASSO-selected baseline controls, and neighborhood-by-composition-by-disclosure fixed effects. Annotations above pairs report higher-minus-lower differences. Section 4.3 discusses the figure; Appendix Tables C-14 and C-19 report both leader types and both specifications for Panels A and B, respectively.

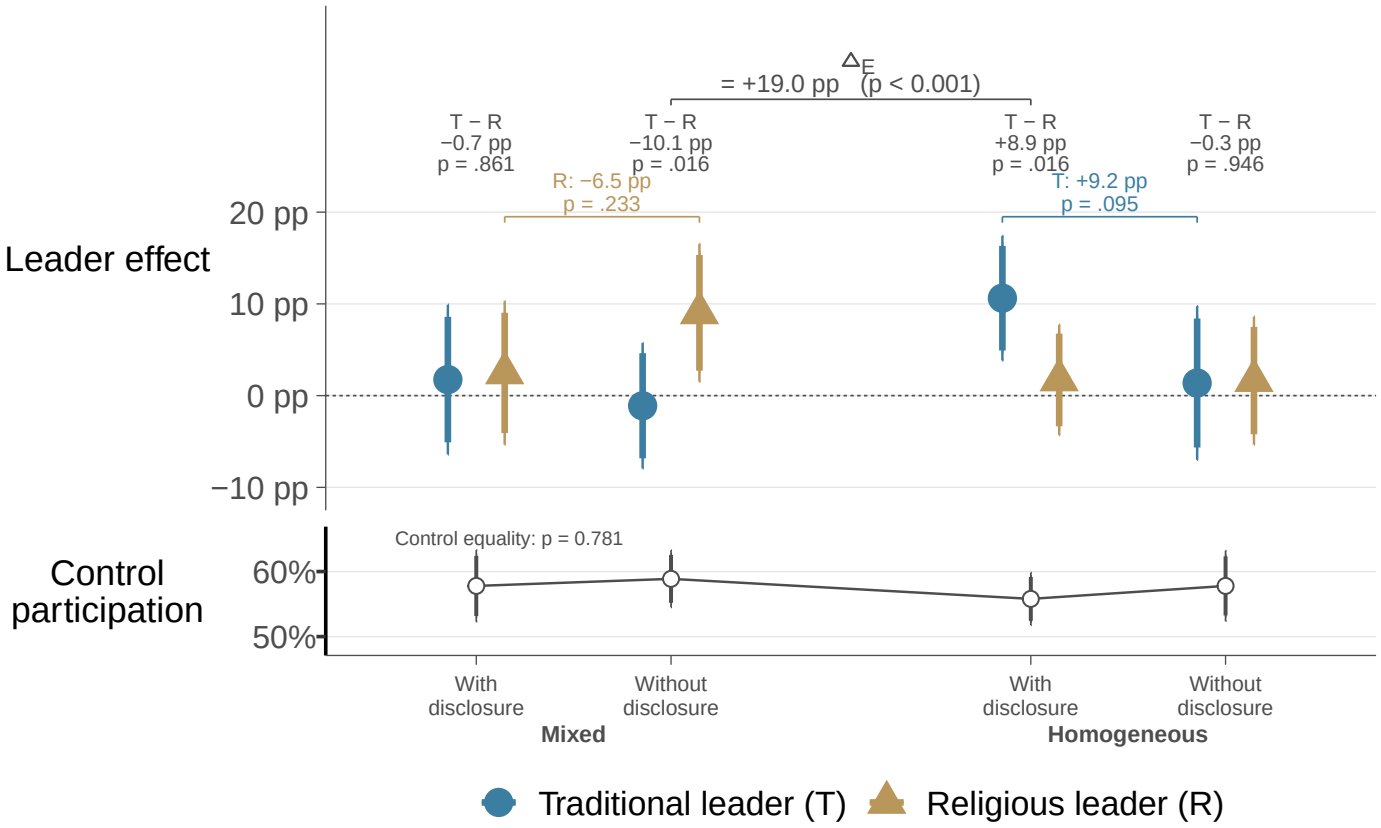


Figure 7: LEADER EFFECTS BY COMPOSITION AND DISCLOSURE

Notes: Blue circles and gold triangles show traditional leader and religious leader effects relative to no leader in each composition-by-disclosure environment. Traditional leaders raise participation most in homogeneous groups with disclosure, and religious leaders in mixed groups without disclosure. T–R labels compare the leaders directly; Δ_E is the difference between their gaps in these two environments. Colored brackets report disclosure-minus-no-disclosure differences. Open circles below show no-leader participation. Thick and thin bars are 90 and 95 percent confidence intervals. Estimates use the same sample, pretreatment participation control, spillover exposure, LASSO-selected baseline controls, and neighborhood-by-composition-by-disclosure fixed effects as Figure 4. Section 5 discusses the figure; Appendix Table D-1 reports the estimates and specification checks.

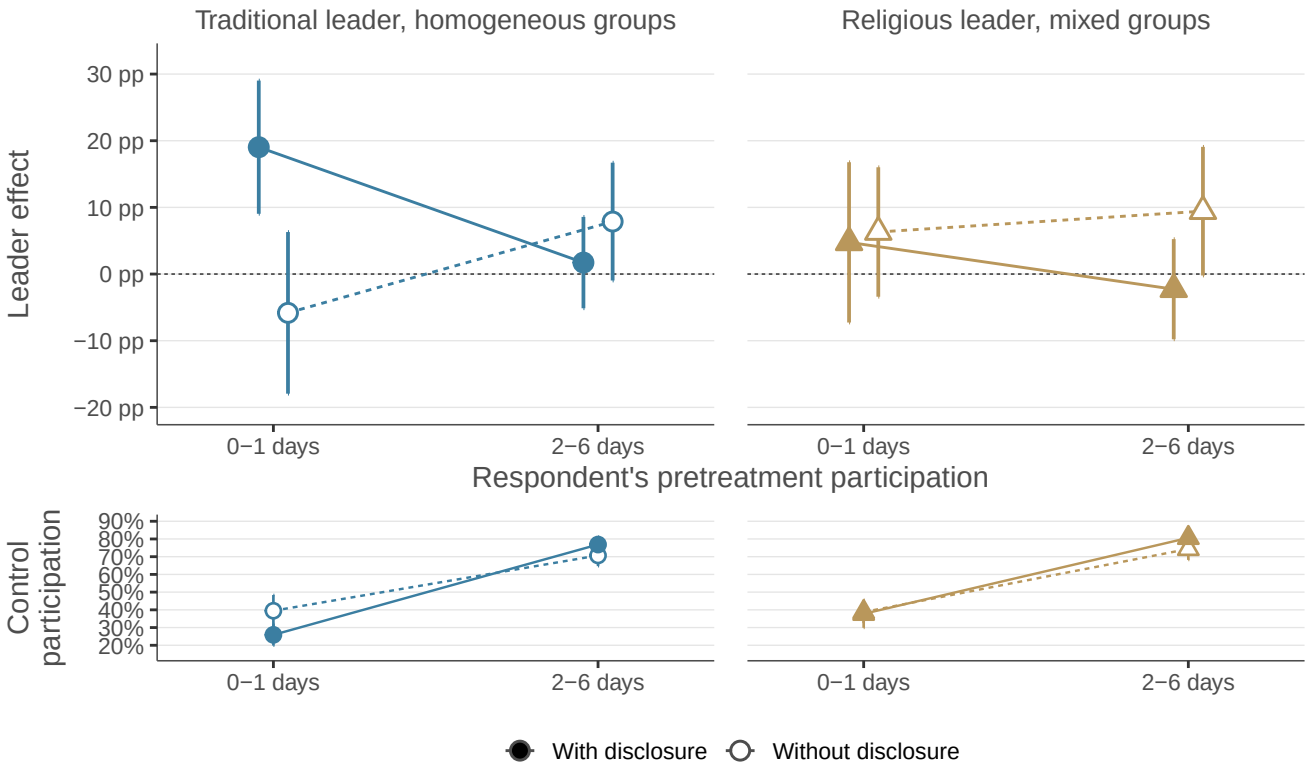


Figure 8: TRADITIONAL LEADERS DRAW IN PRIOR NONPARTICIPANTS UNDER DISCLOSURE; RELIGIOUS LEADER EFFECTS SPAN PARTICIPATION HISTORIES

Notes: Read each column from top to bottom. The upper panels compare leader effects among residents with lower and higher pretreatment participation. Below them, we show predicted participation without a leader for the same types of residents. The left column reports traditional leaders in homogeneous groups and the right column religious leaders in mixed groups; blue circles and gold triangles distinguish leader types. Filled markers and solid lines indicate disclosure; open markers and dashed lines indicate no disclosure. Lower participation means zero or one of the six eligible pre-meeting days, and higher participation means two to six days. With disclosure, the traditional leader effect is concentrated among prior nonparticipants and infrequent participants. Without disclosure, the religious leader effect extends across both participation groups. The no-leader means help distinguish differences in leader influence from differences in participation without a leader. Estimates include baseline participation, 25-meter spillover exposure, 11 LASSO-selected baseline controls, and neighborhood-by-composition-by-disclosure fixed effects. Lines are 90 percent confidence intervals. The figure is discussed in Section 5.3; Appendix Table D-12 reports the exact plotted specification and direct disclosure contrasts.

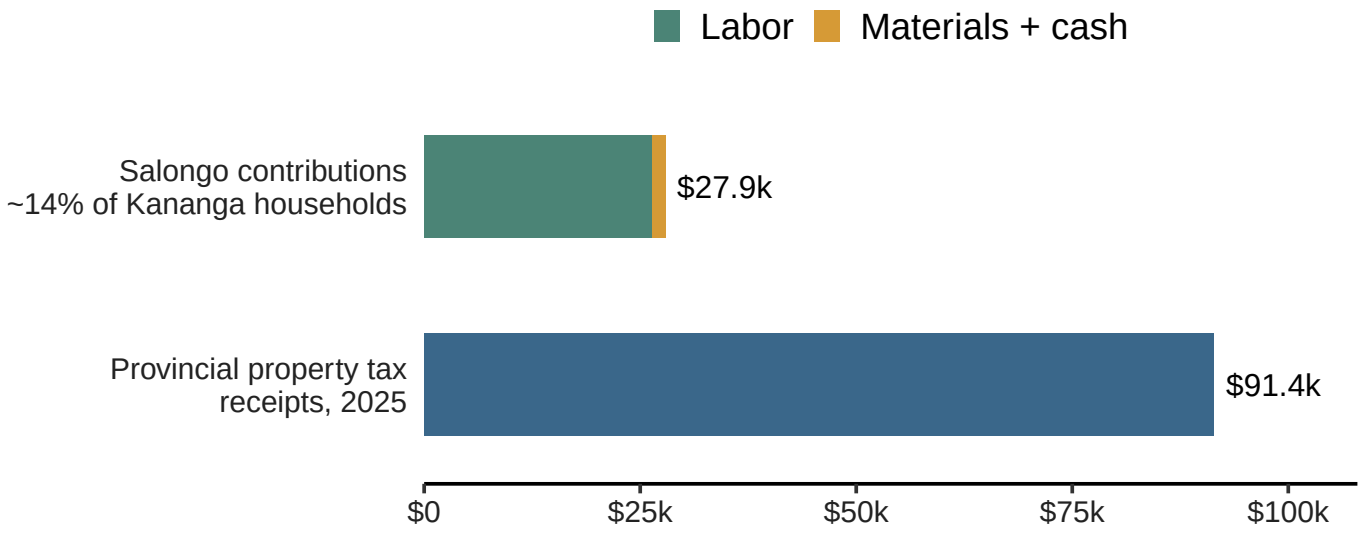


Figure 9 – SALONGO CONTRIBUTIONS AND PROPERTY-TAX RECEIPTS

Notes: Recorded salongo contributions combine labor valued at $0.375/hour with cash and recorded in-kind contributions. The $27,890 total equals 30.5% of reported 2025 Kasaï-Central provincial property-tax receipts, converted at 2,250 CDF/USD. Salongo contributions cover 28 neighborhoods during August 2025–April 2026. Household coverage uses residential buildings as a proxy: 13,575 of 99,109 in Kananga's matched building-footprint and tax-roll data (13.7%). See Appendix D.5 for the valuation and coverage details.

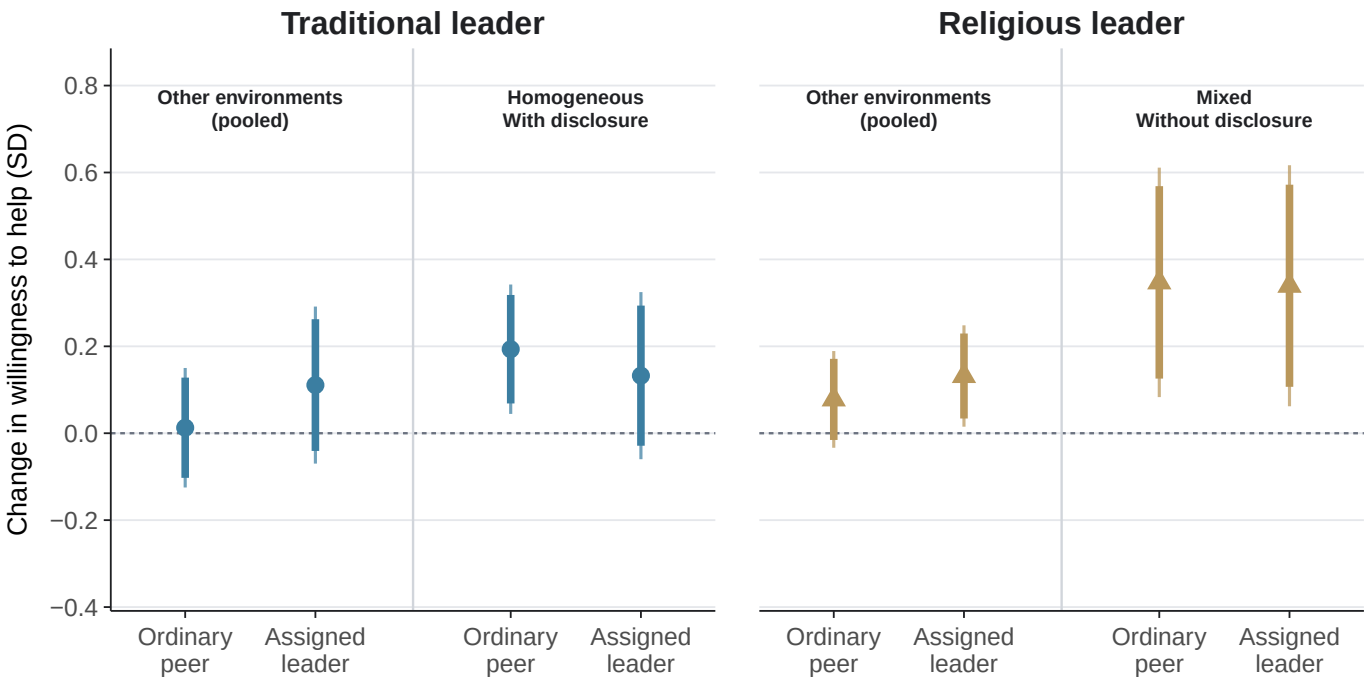


Figure 10: SUCCESSFUL MOBILIZATION STRENGTHENS DIFFERENT RELATIONSHIPS WITHIN THE GROUP

Notes: Before the first discussion and again at the start of Meeting 2, residents reported how much they would help a named group member. The figure shows how leader assignment changes willingness to help over time. The left panel reports traditional leaders and the right panel religious leaders. Within each panel, the right block isolates the environment where that leader raises participation; the left block pools the other three environments. The named person is either an ordinary peer or the assigned leader. Blue circles report traditional leader effects and gold triangles religious leader effects. Effects are measured in baseline standard deviations of the help measure; thick and thin lines are 90 and 95 percent confidence intervals. Willingness-to-help estimates include the 25-meter spillover indicator; outcome-specific LASSO selected no additional baseline controls. Section 6 discusses the figure. Appendix Tables E-4 and E-5 report all outcomes and environments. Panel B of those tables estimates each environment separately, whereas this figure pools the other three environments separately for each leader type; the resulting coefficients therefore differ.

Appendix Contents

Sections

A Context and Leader Characteristics **53**

B Experimental Design and Data **58**
B.1 Fieldwork continuation after Wave 8 . . . 58
B.1.1 Predetermined neighborhood characteristics . . . 58
B.1.2 Power under ten planned waves and seven pre-declaration waves . . . 58
B.1.3 Comparison of Waves 1–7 and Waves 1–8 . . . 61
B.2 Randomization and Team-Construction Protocol . . . 64
B.2.1 Stage 1: Eligible rosters and fixed participants . . . 64
B.2.2 Stage 2: Sampling ordinary residents . . . 64
B.2.3 Stage 3: Forming homogeneous and mixed teams . . . 64
B.2.4 Stage 4: Cumulative rerandomization . . . 65
B.2.5 Stage 5: Information and time slots . . . 65
B.2.6 Stage 6: Unavailable residents and replacements . . . 65
B.2.7 Implementation refinements across waves . . . 66
B.2.8 Pseudocode . . . 66
B.3 Implementation Checks . . . 67
B.4 Meeting Discourse . . . 71

C Scope of Authority **77**
C.1 Spillover Exposure and Radius Robustness . . . 77
C.2 Baseline Control Selection . . . 81
C.3 Main Effects, Alternative Outcomes, and Inference . . . 82
C.4 Alternative Explanations . . . 86
C.4.1 Endogenous Leader Behavior . . . 86
C.4.2 Majority Membership . . . 88
C.4.3 Pastors' Territorial Origins . . . 89
C.4.4 Leadership by Example . . . 90
C.4.5 Pretreatment Participation and Baseline Adjustment . . . 91
C.5 Resident and Group Characteristics . . . 93
C.6 Which Leaders Can Exercise Their Authority? . . . 102

D Source of Enforcement **107**
D.1 Main Effects, Alternative Outcomes, and Inference . . . 107
D.2 Additional Outcome Windows . . . 111
D.3 Alternative Explanations Under Disclosure . . . 114
D.3.1 Pastors' Territorial Origins . . . 117
D.3.2 Leadership by Example . . . 118
D.3.3 Pretreatment Participation . . . 119
D.4 Responses to Revealed Participation Histories . . . 121
D.5 Salongo Contributions and Leadership Alignment . . . 122

E Beyond Participation **125**
E.1 Cooperation Beliefs, Identity, and Governance . . . 125
E.2 Relationships with Peers and Leaders . . . 130

Appendix Tables

A. Context and Leader Characteristics

A-1 Leader Counts and Recruitment Population Shares . . . 53
A-2 Leader Selection and Institutional Embeddedness . . . 55
A-3 Social Proximity and Perceived Authority . . . 56
A-4 Territory-of-origin composition of recovered church congregations . . . 57

B. Experimental Design and Data

B-1 Predetermined Neighborhood Characteristics by Planned Wave . . . 59
B-2 Environment-specific estimates in Waves 1–7 and Waves 1–8 . . . 61
B-3 Mapping from group identifiers to experimental cells . . . 65
B-4 Assignment Balance and Pretreatment Participation . . . 67
B-5 Roster Formation, Replacements, and Composition Fidelity . . . 70
B-6 Treatment Delivery at the Accountability Meetings . . . 70
B-7 Administrative, Survey, and Audio Coverage . . . 70
B-8 Examples of the Three Speech Topics . . . 72
B-9 Leader Speech in the First Meeting . . . 73

C. Scope of Authority

C-1 Leader Effects by Group Composition . . . 82
C-2 Labor Contribution Outcomes by Group Composition . . . 83
C-3 Cash and In-Kind Contributions by Group Composition . . . 84
C-4 Randomization Inference and Alternative Clustering by Group Composition . . . 85
C-5 Arriving Together with the Assigned Leader . . . 86
C-6 Interactions with the Assigned Leader . . . 87
C-7 Shared Territorial Origin with the Neighborhood Chief . . . 88
C-8 Leader Effects Under the Religious-Leader Territory Restriction . . . 89
C-9 Leadership by Example: Leader Pretreatment Participation . . . 90
C-10 Pretreatment Dynamics: Pooled Across Disclosure . . . 91
C-11 Baseline Imbalance and Treatment-Specific Baseline Adjustment . . . 92
C-12 Group Ethnic Identification and Leader Effectiveness . . . 94
C-13 Group Ethnic Identification and Coethnicity in Mixed Groups . . . 95
C-14 Opinion of the Assigned Leader . . . 96
C-15 Perceived Secular Authority of the Assigned Leader . . . 97
C-16 Coethnicity with the Leader . . . 98
C-17 Individual Ethnic Identification and Religiosity . . . 99
C-18 Group Ethnic Identification and Individual Religiosity . . . 99
C-19 Intrinsic Religiosity and Leader Effectiveness . . . 100
C-20 Extrinsic Religiosity and Leader Effectiveness . . . 101
C-21 Leader Effects by Years of Residence on the Avenue . . . 103
C-22 Leader Effects by Belief in Divine Rewards and Punishments . . . 104
C-23 Leader Effects by Joint Chief and Pastor Roles . . . 105
C-24 Denominational Ties and Leader Effectiveness . . . 106

Appendix Tables (continued)

D. Source of Enforcement

D-1 Leader Effects by Composition and Disclosure . . . 107
D-2 Labor Contribution Outcomes by Composition and Disclosure . . . 108
D-3 Cash and In-Kind Contributions by Composition and Disclosure . . . 109
D-4 Randomization Inference and Alternative Clustering by Composition and Disclosure . . . 110
D-5 Robustness to Alternative Post-Outcome Windows . . . 113
D-6 Interactions about Participation and Its Consequences . . . 115
D-7 Pastor Effects by Chief Coethnicity and Disclosure . . . 116
D-8 Leader Effects Under the Religious-Leader Territory Restriction: Disclosure Specification . . . 117
D-9 Leadership by Example: Disclosure Specification . . . 118
D-10 Pretreatment Dynamics . . . 119
D-11 Leader Effects Excluding Replacement Respondents . . . 120
D-12 Leader Effects by Days of Pretreatment Participation . . . 121
D-13 Worksite Workload Benchmarks by Distance Cutoff . . . 123

E. Beyond Participation

E-1 Leader Effects on Cooperation Norms and Divine Monitoring Beliefs . . . 125
E-2 Leader Effects on Identity Salience . . . 127
E-3 Governance and Institutional Outcomes: Tests by Outcome Family . . . 129
E-4 Endpoint Changes in Opinions about Ordinary Peers . . . 130
E-5 Endpoint Changes in Opinions about Assigned Leaders . . . 131

Appendix Figures

A-1 Neighborhood ethnic fractionalization and perceived importance of *salongo* . . . 54
B-1 Power for leader effects and the paper's composition and environment contrasts . . . 60
B-2 Leader Topic Shares Across Text Models . . . 74
B-3 Leader Speech by Environment Across Text Models . . . 75
B-4 Participant Speech Levels Across Text Models . . . 76
C-1 Terminal Spillover Exposure in Three Neighborhoods . . . 77
C-2 Leader Effects Across Spillover Radii . . . 79
C-3 Spillover Exposure Coefficient Across Radii . . . 80
D-1 Outcome windows used in the timing robustness analysis. . . . 111
D-2 Leader Effects Across Post-Treatment Outcome Windows . . . 112
D-3 Activities at observed campaign worksites . . . 124

A Context and Leader Characteristics

Table A-1 – LEADER COUNTS AND RECRUITMENT POPULATION SHARES

	Leader counts		Population
	Religious leaders	Traditional leaders	Share (%)
Panel A. Origin territories			
Dibaya	12 (42.9)	14 (50.0)	36.8
Demba	9 (32.1)	7 (25.0)	25.9
Kazumba	2 (7.1)	5 (17.9)	18.2
Dimbelenge	3 (10.7)	1 (3.6)	5.7
Kananga	1 (3.6)	1 (3.6)	5.3
Other / Not Reported	1 (3.6)	0 (0.0)	8.2
Total	28 (100.0)	28 (100.0)	100.0
Panel B. Main denominations			
Pentecostal, autonomous	5 (17.9)	6 (21.4)	22.8
Pentecostal, Branham	3 (10.7)	3 (10.7)	17.7
Catholic	0 (0.0)	4 (14.3)	11.8
Neo-Apostolic	4 (14.3)	4 (14.3)	10.8
Pentecostal, network	4 (14.3)	0 (0.0)	9.1
Kimbanguist	7 (25.0)	2 (7.1)	6.7
Traditional religion	1 (3.6)	1 (3.6)	5.3
Adventist	0 (0.0)	2 (7.1)	4.1
Protestant/Presbyterian	2 (7.1)	2 (7.1)	4.0
Other / not reported	2 (7.1)	4 (14.3)	7.8
Total	28 (100.0)	28 (100.0)	100.0

Notes: How similar are the leaders to residents recruited in the same neighborhoods? We compare individual records in the full recruitment sample, restricting neighborhoods to final assignment Waves 1–7. The leader columns count each assigned person once and show within-column percentages in parentheses. In the final assignment files, these are the people for whom `leader_religious` or `leader_traditional` equals one. Population percentages use the $N = 2,795$ recruited residents with usable neighborhood information. Territory is their reported origin territory; denomination comes from the coded religion/denomination item, rather than the open-text church name. Other / not reported combines rows with less than 4 percent of the recruitment population.

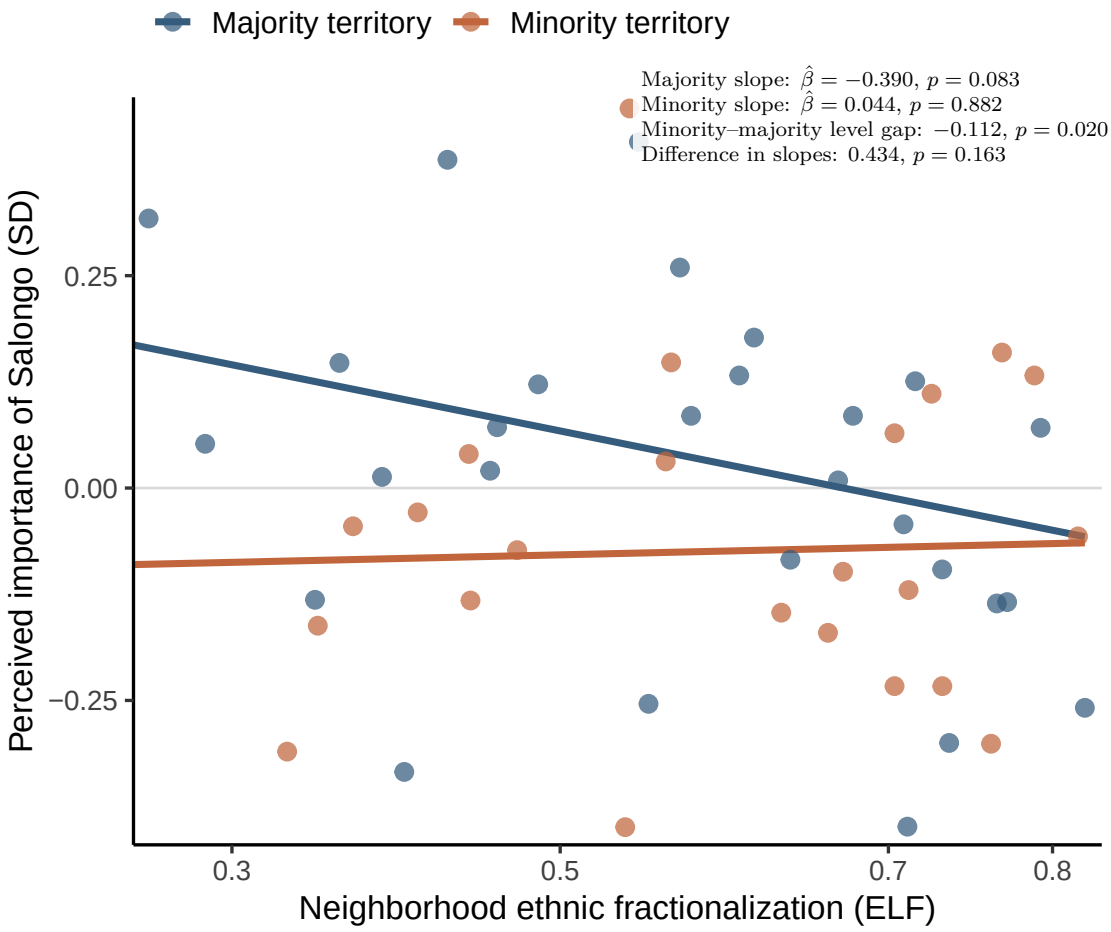


Figure A-1: Neighborhood ethnic fractionalization and perceived importance of *salongo*

Notes: The sample contains 2,791 recruitment-survey respondents in the 28 nonpilot neighborhoods. Neighborhood ethnic fractionalization is the leave-one-out territory-based ELF, $1 - \sum_g s_g^2$, computed from all respondents in the recruitment roster with nonmissing neighborhood and territory. The outcome is the respondent's perceived importance of *salongo*, standardized once over the full estimation sample. The question reads, "How important do people in your community think it is to participate in Salongo?" Responses use a five-point ordered scale anchored at 1, "They think it is not important at all," and 5, "They think it is very important"; values 2–4 are unlabeled intermediate points. Minority status indicates that the respondent's territory is not the modal territory in the neighborhood. The fitted lines come from one individual-level regression of standardized *salongo* importance on mean-centered ELF, minority status, and their interaction, with no additional controls. Standard errors and the displayed tests are clustered by neighborhood. Points are neighborhood-by-status outcome means; the lines are fitted values from the individual-level model. The level gap is evaluated at mean ELF.

Table A-2 – LEADER SELECTION AND INSTITUTIONAL EMBEDDEDNESS

	Observations (1)	Citizen mean (2)	Coeff. on religious leader (3)	Coeff. on traditional leader (4)	p-value of diff., columns 3–4 (5)
Panel A. Demographic and economic selection					
Female	2,681	0.603	**-0.590**	**-0.448**	**0.043**
			[**0.000**]	[**0.000**]	
Age	2,681	34.098	**13.289**	**22.253**	**0.035**
			[**0.000**]	[**0.000**]	
Education (yrs)	2,681	9.531	**1.765**	**1.408**	0.662
			[**0.001**]	[**0.020**]	
Non-clergy salaried job	2,681	0.037	0.033	0.105	0.424
			[0.519]	[0.126]	
Government job	2,681	0.006	0.065	0.030	0.573
			[0.198]	[0.416]	
Log income	1,354	11.558	0.106	**0.417**	0.499
			[0.805]	[**0.007**]	
Avenue tenure	1,327	15.032	**12.880**	**12.487**	0.937
			[**0.002**]	[**0.002**]	
Panel B. Institutional embeddedness					
Traditional authority and local embeddedness					
Any traditional role	1,362	0.047	**0.346**	**0.953**	**0.000**
			[**0.001**]	[**0.000**]	
Traditional head role	1,362	0.026	**0.260**	**0.617**	**0.005**
			[**0.006**]	[**0.000**]	
Secrets network: same territory	1,362	0.295	0.073	**0.163**	0.459
			[0.360]	[**0.039**]	
Material network: same territory	1,362	0.425	0.037	**0.134**	0.311
			[0.632]	[**0.043**]	
Planned salongo hours	1,350	16.807	**6.837**	**10.051**	0.296
			[**0.036**]	[**0.000**]	
Religious authority and church embeddedness					
Any religious role	1,362	0.139	**0.862**	0.147	**0.000**
			[**0.000**]	[0.090]	
Religious head role	1,362	0.053	**0.662**	**0.197**	**0.001**
			[**0.000**]	[**0.024**]	
Secrets network: same church	1,362	0.177	0.049	-0.004	0.601
			[0.471]	[0.956]	
Material network: same church	1,362	0.223	0.091	**-0.128**	**0.008**
			[0.126]	[**0.006**]	
Sunday attendance	1,362	3.466	**0.319**	0.176	0.381
			[**0.034**]	[0.113]	
Weekday attendance	1,362	1.590	**0.877**	**1.055**	0.742
			[**0.030**]	[**0.002**]	
Log religious contribution	1,317	6.244	**1.482**	0.882	0.296
			[**0.000**]	[0.089]	
Knowledge of local authorities					
Knows authorities	1,362	0.292	**0.315**	**0.434**	0.087
			[**0.000**]	[**0.000**]	
Knows quartier chief	1,362	0.309	**0.298**	**0.477**	0.057
			[**0.001**]	[**0.000**]	
Knows locality chief	1,362	0.352	**0.256**	**0.470**	0.083
			[**0.008**]	[**0.000**]	
Knows chief phone	1,362	0.071	**0.322**	**0.358**	0.802
			[**0.001**]	[**0.001**]	
Chief help index	1,362	1.499	0.170	**0.705**	0.067
			[0.491]	[**0.003**]	
Panel C. Psychological traits					
Standard psychological measures					
Big Five Inventory (main index)	1,362	-0.005	0.019	**0.226**	0.086
			[0.838]	[**0.025**]	
Extraversion	1,358	3.210	-0.013	0.112	0.346
			[0.884]	[0.298]	
Agreeableness	1,361	3.844	0.030	0.120	0.593
			[0.790]	[0.255]	
Conscientiousness	1,360	3.905	0.023	**0.273**	0.085
			[0.846]	[**0.004**]	
Neuroticism	1,362	2.872	-0.104	-0.175	0.659
			[0.464]	[0.116]	
Openness	1,362	3.192	-0.085	0.058	0.476
			[0.497]	[0.733]	
Emotional intelligence	1,362	4.296	-0.014	0.007	0.803
			[0.870]	[0.936]	
Locus of control	1,362	4.069	-0.140	0.324	0.091
			[0.612]	[0.091]	
Social desirability	1,362	0.560	**0.055**	0.025	0.322
			[**0.022**]	[0.278]	
Religious psychological measures					
Positive RCOPE	1,362	3.794	**0.106**	-0.073	0.073
			[**0.012**]	[0.448]	
Negative RCOPE	1,362	3.540	**-0.605**	-0.162	0.115
			[**0.005**]	[0.249]	
Intrinsic religiosity	1,362	3.326	0.109	0.009	0.297
			[0.077]	[0.888]	
Extrinsic religiosity: personal	1,358	4.416	-0.060	-0.024	0.839
			[0.586]	[0.863]	
Extrinsic religiosity: social	1,362	2.678	-0.320	0.133	0.079
			[0.094]	[0.491]	

Notes: We compare each leader type with ordinary residents of the same locality. Each row reports a separate respondent-level regression of the listed characteristic on religious- and traditional-leader indicators and locality fixed effects. Locality-clustered p-values are reported in square brackets. Column 5 reports the p-value for equality of the two leader coefficients. We identify the leaders selected to join the experimental teams from the selected meeting/panel leader IDs. The first five demographic and employment rows use the recruitment survey (N = 2,681). The remaining rows combine the study panel with the capped sample of additional baseline-control respondents, giving up to 1,362 observations depending on data availability. Network shares code missing network counts as zero and code shares as zero for respondents with a zero-size network.

Table A-3 – SOCIAL PROXIMITY AND PERCEIVED AUTHORITY

Outcome	N	Non-leader		Traditional leader		Religious leader	
		Different ethnicity mean	Same-ethnicity premium	Different ethnicity	Same ethnicity	Different ethnicity	Same ethnicity
Panel A. Initial Social Cohesion and Pre-Existing Ties							
Know							
Know other person	2,894	2.938	**0.131**	**0.383**	**0.466**	**0.294**	**0.491**
			[0.042]	[0.000]	[0.000]	[0.000]	[0.000]
Pre-existing ties							
Same church	2,894	0.024	0.019	0.003	0.039	0.036	0.041
			[0.061]	[0.882]	[0.071]	[0.056]	[0.064]
Same groupement	2,894	0.000	**0.057**	-0.003	**0.087**	-0.001	**0.048**
			[0.000]	[0.402]	[0.000]	[0.831]	[0.005]
Same village	2,894	0.000	**0.046**	-0.003	**0.060**	0.001	**0.048**
			[0.000]	[0.321]	[0.002]	[0.760]	[0.005]
Close family	2,894	0.010	0.013	-0.002	0.024	0.002	0.013
			[0.076]	[0.806]	[0.176]	[0.799]	[0.323]
Extended family	2,894	0.017	0.020	0.008	**0.050**	0.001	0.033
			[0.060]	[0.608]	[0.024]	[0.924]	[0.093]
Reciprocity							
Mutual sickness support	2,894	1.157	**0.168**	**0.309**	**0.310**	0.163	**0.242**
			[0.042]	[0.005]	[0.001]	[0.111]	[0.038]
Willing to help them	2,894	1.268	**0.168**	0.179	**0.241**	0.103	0.141
			[0.033]	[0.077]	[0.017]	[0.304]	[0.209]
Expected help from them	2,894	1.056	0.168	**0.438**	**0.379**	0.224	**0.343**
			[0.094]	[0.002]	[0.003]	[0.058]	[0.011]
Panel B. Opinions and Authority							
Motivates respondent	2,894	4.041	0.112	0.173	**0.393**	**0.239**	**0.329**
			[0.057]	[0.051]	[0.000]	[0.001]	[0.000]
Moral example	2,894	4.119	**0.160**	0.167	**0.352**	**0.243**	**0.293**
			[0.005]	[0.056]	[0.000]	[0.000]	[0.000]
Trustworthy	2,894	3.845	0.091	0.118	**0.344**	**0.254**	**0.321**
			[0.138]	[0.158]	[0.000]	[0.000]	[0.000]
Hardworking	2,894	4.043	0.046	0.119	**0.340**	**0.301**	**0.268**
			[0.415]	[0.215]	[0.000]	[0.000]	[0.001]
Seen as authority	2,894	3.586	0.084	**0.212**	**0.253**	0.117	**0.210**
			[0.134]	[0.002]	[0.000]	[0.092]	[0.016]
Panel C. Participation and Material Expectations							
Trust they will join	2,894	3.160	0.098	**0.224**	**0.330**	**0.240**	**0.246**
			[0.067]	[0.004]	[0.000]	[0.000]	[0.004]
Expected money contribution	2,485	0.109	-0.054	**1.056**	**0.698**	**0.939**	**1.209**
			[0.807]	[0.002]	[0.014]	[0.001]	[0.001]
Expected work days	2,691	3.253	0.071	**0.485**	**0.580**	0.271	0.349
			[0.537]	[0.002]	[0.000]	[0.055]	[0.059]

Notes: Each respondent in this table is an ordinary resident evaluating a particular teammate. We estimate a separate regression for each outcome to compare these evaluations by the teammate's leader status and ethnic ties to the respondent. An evaluation of an ordinary teammate from a different ethnicity is the omitted category. The first data column reports its mean, and coefficients in the remaining columns measure differences from this group. Standard errors allow correlation both within respondent and within accountability group in the same neighborhood. Brackets give two-sided p-values, and coefficients are bold when $p < 0.05$.

For familiarity ("Know other person") and trust in participation, we show comparison means on the original 1–4 scales, while Panel B uses the original 1–5 scales. The coefficients for these ratings are in standard deviations of the relevant outcome. For pre-existing ties, which are binary, they are in probability units. We transform monetary answers in CDF using the inverse hyperbolic sine and estimate coefficients on that scale. To express comparison means in USD, we take the hyperbolic sine of mean transformed CDF and divide by 2250. The conversion is $\sinh(\overline{\text{asinh}(\text{CDF})})/2250$. To measure mutual sickness support, we take the mean of the two help amounts after transforming them, then use this mean for estimation and back-transformation. Expected work is in days; we convert earlier answers given in hours or minutes using six hours for each day.

All regressions include neighborhood fixed effects. We select controls separately for each outcome, after removing neighborhood means, using post-double-selection LASSO. A control is kept if it is selected for the outcome or for the coethnicity and leader regressors. We let LASSO choose from the respondent's baseline social-desirability index, sex, age, education, wealth stratum, territory, and religion; indicators for whether respondent and teammate match in location, sex, age, education, and wealth; and agreement between the respondent's report of the teammate's religion and the teammate's recorded religion.

Table A-4 – TERRITORY-OF-ORIGIN COMPOSITION OF RECOVERED CHURCH CONGREGATIONS

	Full recruitment roster		Randomized study sample	
	Exact normalized	Strict fuzzy	Exact normalized	Strict fuzzy
Recovered church clusters	66	71	10	12
Observed congregants in recovered clusters	513	555	67	79
Church clusters spanning multiple origin territories	60	64	10	12
Share of church clusters spanning multiple territories	90.9%	90.1%	100.0%	100.0%
Share of congregants in those church clusters	91.6%	91.4%	100.0%	100.0%
Same-territory pairs within church clusters	48.2%	48.4%	53.2%	55.4%
Expected same-territory pairs under no sorting	48.8%	48.5%	53.6%	55.2%
Observed minus expected same-territory pairs	-0.6%	-0.1%	-0.4%	0.2%
Permutation p-value: positive territorial sorting	0.696	0.570	0.566	0.485

Notes: To compare ethnic sorting within congregations, we need to establish which reported church names refer to the same congregation. We use Christian respondents in the cumulative Waves 1–7 recruitment roster who report a valid territory, neighborhood, and free-text church name. As in the experiment, two residents are coethnic when their normalized reported territories of origin match exactly.

We link church names within neighborhood-by-denomination blocks. Exact matches have identical sorted sets of tokens after lowercasing, removing accents and punctuation, standardizing saint abbreviations, and removing generic church and denomination words. Names that give no information about the congregation are excluded. Strict fuzzy matching also links names to reference names ranked by frequency. It requires Jaro–Winkler similarity of at least 0.97, taking the maximum over the original and sorted token orders. The additional restrictions depend on the names being compared. A single-token match requires at least five characters in each name and similarity of 0.99 or higher. We impose token Jaccard similarity of at least 0.80 when two multi-token names have different token counts. Could the names nevertheless refer to different churches? Conflicting numeric identifiers rule out the match, and we require a direct match with the reference name rather than a chain through intermediate names.

We compare congregations with at least five observed members against a no-sorting allocation of territories. For this comparison, we keep the recovered church sizes and draw new territorial assignments within neighborhood-by-denomination blocks; 999 draws give the one-sided permutation test. A recovered church's size is its membership observed during recruitment. It need not equal total congregation membership, and a congregation drawing residents from several neighborhoods can be split by the restriction to matching names within neighborhood.

B Experimental Design and Data

B.1 Fieldwork continuation after Wave 8

On Friday, May 15, 2026, the Democratic Republic of the Congo declared a Bundibugyo Ebola outbreak after laboratory confirmation in Ituri Province; on May 17, the World Health Organization (WHO) determined that the event constituted a Public Health Emergency of International Concern. Wave 8 was underway when the outbreak happened. Waves 9 and 10 were scheduled for the upcoming months of June, July, and August and would have required the repeated convening of new groups for collective work and program meetings. The institutional response to the outbreak underscored the risks of initiating such activities. WHO recommended considering the postponement of mass gatherings; UN staff in Bunia were instructed to work from home and avoid physical contact and crowded areas; the U.S. Department of State temporarily suspended all visa services at its embassy in Kinshasa; and the African Development Bank introduced remote participation for its Annual Meetings in neighboring Congo while explicitly monitoring risks from the outbreak, despite no confirmed or suspected cases in the host country. Initiating additional waves under these conditions was unwarranted and posed severe ethical risks. We therefore decided to stop the study on May 30, 2026, interrupting the implementation of the upcoming cohorts.[28] The emergency overlapped Wave 8 implementation. At the time of the declaration, Wave 8 was only 30.4 percent complete.

Stopping fieldwork could affect both the neighborhoods in the sample and the precision of their estimates. We investigate the first concern by comparing predetermined characteristics of neighborhoods reached before the emergency with those scheduled later (Section B.1.1), and the second by comparing power for ten planned waves with seven completed waves (Section B.1.2). The main analysis uses Waves 1–7. We then add the disrupted Wave 8, comparing estimates for Waves 1–8 with those for Waves 1–7.

B.1.1 Predetermined neighborhood characteristics

The ten-wave rollout order was randomly assigned before implementation. Table B-1 compares property values, building size and counts, and commune composition for neighborhoods assigned to Waves 1–7 and Waves 8–10.

None of the 12 characteristic-specific tests rejects at the 5 percent level. The exact max-statistic test across all 12 characteristics yields $p = 0.787$. We find no systematic differences in the observed characteristics of early- and late-wave neighborhoods. These tests have limited power, however, and cannot rule out differences in characteristics we did not measure. Within each realized neighborhood, treatment remains randomized, and the specifications absorb neighborhood $\times$ disclosure $\times$ composition fixed effects.

B.1.2 Power under ten planned waves and seven pre-declaration waves

How much precision would the remaining waves have added? Figure B-1 compares the planned ten-wave design with Waves 1–7, completed before the national declaration. We calibrate this exercise with the municipal field pilot conducted before July 1, 2025, well before the present experiment. Its 251 participants in four neighborhoods have a six-day baseline week and a separate six-day outcome week immediately after it. We exclude the earlier pilot neighborhood of Ngdilayi because its calendar does not contain these consecutive weeks.

The preregistration planned 40 neighborhoods and 480 groups, with a minimum detectable effect (MDE) of 0.20 outcome standard deviations. We keep the same MDE for both designs. In every wave, three ordinary residents enter the analysis from each chief or pastor group. No-leader groups contribute four ordinary residents in Waves 1–2 and three in subsequent waves. Thus, seven waves give 1,040 observations in 336 groups, compared with 1,472 in 480 groups under ten waves. In each simulation, we estimate the paper's model with baseline participation and neighborhood $\times$ disclosure $\times$ composition fixed effects, clustering standard errors by group within neighborhood.

The curves for seven waves describe the Waves 1–7 main analysis. We show the magnitudes of those estimates as gold bars for scale, then report below the same effects with all Wave 8 observations added.

The simulated standard errors for these four leader effects and two delta contrasts are 17–19 percent larger with seven waves. Ten waves would have helped, but power for T_{H1}, R_{M0}, Δ_C, and Δ_E at the 0.20-SD effect size would still fall short of 80 percent.

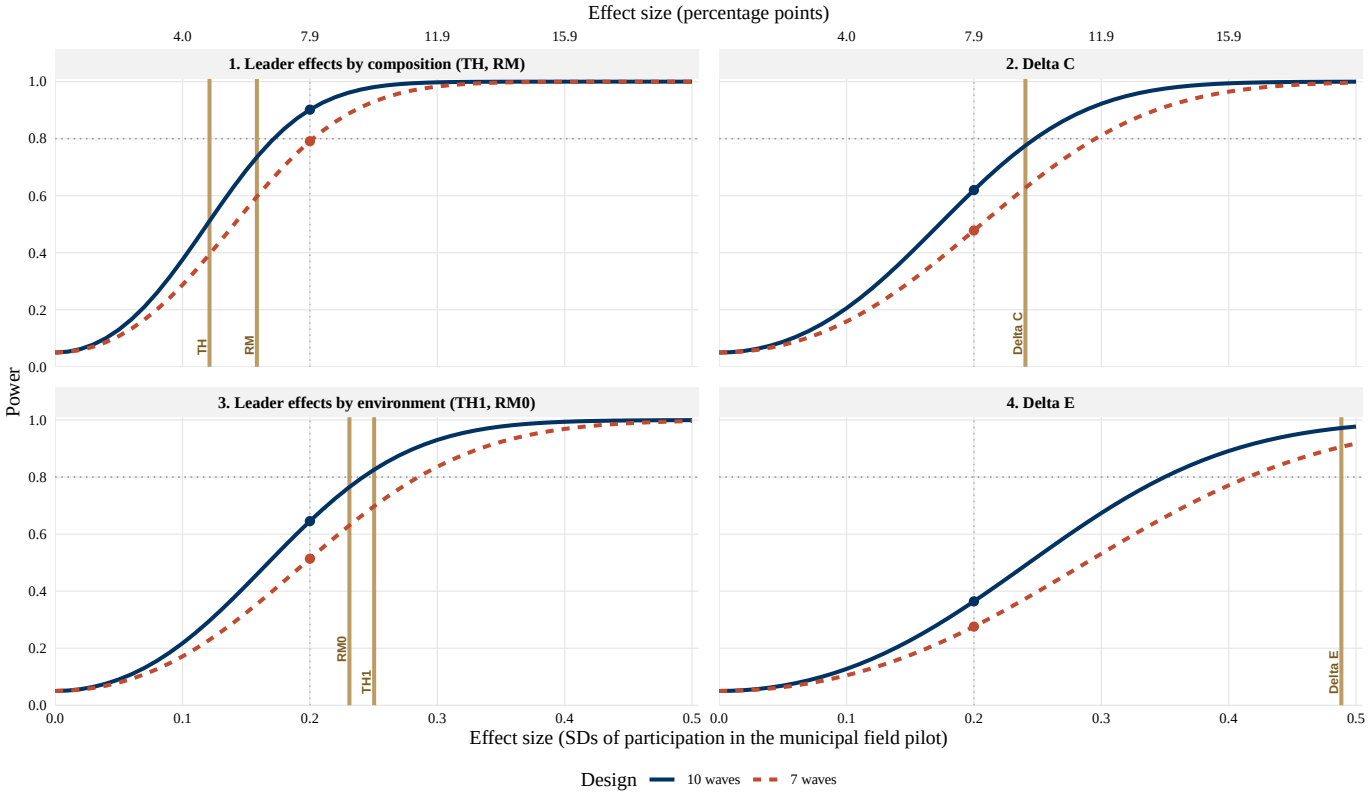
[28] WHO, Disease Outbreak News, May 16, 2026; WHO, PHEIC statement, May 17, 2026; Associated Press, May 18, 2026; U.S. Department of State, May 18, 2026; African Development Bank, May 21, 2026; and WHO, Disease Outbreak News, July 3, 2026.

Table B-1 – PREDETERMINED NEIGHBORHOOD CHARACTERISTICS BY PLANNED WAVE

	Waves 1–7	Waves 8–10	Difference	Std. diff.	RI p
Mean predicted property value (USD)	2612.09	2183.94	-428.15	-0.47	0.358
Median predicted property value (USD)	1359.32	1165.25	-194.07	-0.66	0.187
Mean building area	39.70	38.13	-1.57	-0.34	0.518
Median building area	34.59	33.50	-1.09	-0.35	0.496
Total built area	19261.66	18624.36	-637.30	-0.12	0.810
Number of main properties	313.86	322.00	8.14	0.15	0.769
Number of buildings	484.82	473.17	-11.65	-0.12	0.803
Commune: Kananga (share)	0.11	0.25	0.14	0.50	0.429
Commune: Katoka (share)	0.21	0.08	-0.13	-0.39	0.500
Commune: Lukonga (share)	0.36	0.42	0.06	0.13	1.000
Commune: Ndesha (share)	0.07	0.08	0.01	0.06	1.000
Commune: Nganza (share)	0.25	0.17	-0.08	-0.20	0.868
Joint max-statistic test			Maximum |standardized difference| = 0.66		0.787

Notes: The sample contains the 40 treated neighborhoods assigned ex ante to ten rollout waves: 28 neighborhoods in Waves 1–7 and 12 in Waves 8–10. Each row reports the early-wave mean, late-wave mean, their difference (late minus early), the difference standardized by the across-pair standard deviation, and a two-sided exact randomization-inference p-value. Wave assignment occurred at the neighborhood-pair level. Commune rows are neighborhood shares. The final row reports the randomization p-value for the maximum absolute standardized difference across all 12 predetermined characteristics.

What did the shorter rollout cost in precision? At the preregistered MDE of 0.20 outcome standard deviations, or 7.95 percentage points, power for T_H and R_M would be 0.902 with ten waves, compared with 0.791–0.793 with seven. For T_{H1} and R_{M0}, ten waves give 0.646–0.647 and seven give 0.514–0.518. For the composition contrast Δ_C, power is 0.620 with ten waves and 0.478 with seven. For the environment contrast Δ_E, it is 0.364 with ten and 0.276 with seven.



Dotted line: preregistered MDE benchmark (0.20 SD); gold bars: magnitudes of the labeled Waves 1−7 estimates

Figure B-1 – POWER FOR LEADER EFFECTS AND THE PAPER'S COMPOSITION AND ENVIRONMENT CONTRASTS

Notes: T_H is the traditional-leader effect in homogeneous groups, and R_M is the religious-leader effect in mixed groups. T_{H1} is the traditional-leader effect in homogeneous groups under disclosure, and R_{M0} is the religious-leader effect in mixed groups without disclosure. The four panels report power for (1) T_H and R_M, (2) $\Delta_C = (\tau^T_H - \tau^T_M) - (\tau^R_H - \tau^R_M)$, (3) T_{H1} and R_{M0}, and (4) $\Delta_E = (\tau^T_{H1} - \tau^T_{M0}) - (\tau^R_{H1} - \tau^R_{M0})$. Panels 1 and 3 plot the lower-power curve across their two stated cells; Panels 2 and 4 use the full clustered covariance matrix for the stated linear contrast. The bottom axis reports standardized effects; the top axis translates them into participation percentage points using the outcome standard deviation from the municipal field pilot. The horizontal dotted line marks 80% power, and the vertical dotted line marks the preregistered 0.20-standard-deviation MDE benchmark (7.95 percentage points). Gold bars mark the magnitudes of the Waves 1–7 T_H and R_M estimates in Panel 1, Δ_C in Panel 2, T_{H1} and R_{M0} in Panel 3, and Δ_E in Panel 4. The ten- and seven-wave SEs are, respectively, 0.061 SD (2.42–2.43 percentage points) and 0.072 SD (2.85–2.86 percentage points) for T_H and R_M; 0.088 SD (3.51 percentage points) and 0.105 SD (4.17 percentage points) for Δ_C; 0.086 SD (3.41 percentage points) and 0.101 SD (4.00–4.02 percentage points) for T_{H1} and R_{M0}; and 0.124 SD (4.94 percentage points) and 0.148 SD (5.86 percentage points) for Δ_E. The outcome and pretreatment covariate are separate consecutive six-day participation weeks drawn from the municipal field pilot conducted before July 1, 2025. The simulation applies the group sizes implemented in the randomization scripts and uses 1,000 bootstrap designs; the largest Monte Carlo standard error among the reported power calculations is 0.003. Every simulated ANCOVA absorbs neighborhood × disclosure × composition fixed effects, identifies leader effects within the corresponding fixed-effect cell, and controls for baseline-week participation. Standard errors cluster by group within neighborhood using the paper's small-sample correction.

B.1.3 Comparison of Waves 1–7 and Waves 1–8

Table B-2 – ENVIRONMENT-SPECIFIC ESTIMATES IN WAVES 1–7 AND WAVES 1–8

	Waves 1–7			Waves 1–8		
	Baseline adjustment	+ Spillover	+ LASSO	Baseline adjustment	+ Spillover	+ LASSO
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed, without disclosure	-0.29	-0.31	-1.11	0.45	0.45	0.07
	(3.48)	(3.48)	(3.48)	(3.39)	(3.39)	(3.38)
Mixed, with disclosure	1.75	1.67	1.75	5.53	5.35	5.68
	(4.53)	(4.52)	(4.16)	(4.32)	(4.32)	(4.08)
Homogeneous, without disclosure	-0.31	-0.34	1.37	-3.47	-3.65	-2.30
	(4.46)	(4.45)	(4.27)	(4.43)	(4.44)	(4.37)
Homogeneous, with disclosure	9.95***	10.13***	10.62***	7.63**	7.89**	8.41**
	(3.63)	(3.63)	(3.47)	(3.73)	(3.72)	(3.54)
Religious leader						
Mixed, without disclosure	9.18**	9.16**	9.02**	6.80*	6.88*	7.03*
	(3.97)	(3.98)	(3.83)	(3.78)	(3.79)	(3.65)
Mixed, with disclosure	3.41	3.40	2.48	5.57	5.51	4.77
	(4.30)	(4.32)	(3.99)	(4.01)	(4.04)	(3.80)
Homogeneous, without disclosure	1.66	1.67	1.64	3.29	3.21	3.22
	(3.81)	(3.80)	(3.56)	(3.75)	(3.74)	(3.58)
Homogeneous, with disclosure	0.01	-0.01	1.71	-0.72	-0.72	0.70
	(3.27)	(3.26)	(3.07)	(3.13)	(3.13)	(2.95)
Panel B. Treatment-effect contrasts						
Mixed, with: T − R	-1.66	-1.73	-0.73	-0.04	-0.16	0.90
	(4.40)	(4.41)	(4.15)	(4.15)	(4.17)	(4.02)
Mixed, without: T − R	-9.47**	-9.47**	-10.13**	-6.34*	-6.43*	-6.96*
	(3.98)	(4.01)	(4.17)	(3.84)	(3.87)	(4.00)
Homogeneous, with: T − R	9.94**	10.13**	8.91**	8.35**	8.60**	7.71**
	(4.03)	(4.00)	(3.67)	(3.98)	(3.96)	(3.63)
Homogeneous, without: T − R	-1.97	-2.01	-0.27	-6.77	-6.86	-5.52
	(4.13)	(4.12)	(3.90)	(4.30)	(4.28)	(4.21)
Δ_E: H, with − M, without	19.41***	19.60***	19.04***	14.69***	15.04***	14.67***
	(5.67)	(5.69)	(5.57)	(5.55)	(5.56)	(5.41)
Observations	1,040	1,040	1,040	1,184	1,184	1,184
Group-neighborhood clusters	336	336	336	384	384	384
Neighborhoods	28	28	28	32	32	32
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes	Yes
Pretreatment participation control	Yes	Yes	Yes	Yes	Yes	Yes
25m spillover control	No	Yes	Yes	No	Yes	Yes
LASSO-selected baseline controls	No	No	Yes	No	No	Yes

Notes: Panels A and B reproduce the environment-specific half of the main results table for Waves 1–7 and the pooled Waves 1–8 sample. Panel C, on the following page, gives no-leader means for all six specifications. Each observation is a respondent's average salongo participation, in percentage points, over the six eligible days after Meeting 1 through the day of Meeting 2 within the assigned treatment cell. Panel A reports leader effects relative to no leader; Panel B reports traditional-minus-religious contrasts and Δ_E. All columns include the ANCOVA pretreatment-participation control and neighborhood × disclosure × composition fixed effects. The spillover columns add the 25m exposure indicator. The columns labeled + LASSO add 11 LASSO-selected baseline controls in both wave windows. Standard errors are clustered by group within neighborhood. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table B-2 – ENVIRONMENT-SPECIFIC ESTIMATES IN WAVES 1–7 AND WAVES 1–8 (continued)

	Waves 1–7			Waves 1–8		
	Baseline adjustment	+ Spillover	+ LASSO	Baseline adjustment	+ Spillover	+ LASSO
Panel C. No-leader counterfactual means						
Mixed, without disclosure	58.88	58.88	58.88	57.53	57.53	57.53
	(2.46)	(2.44)	(2.24)	(2.36)	(2.34)	(2.15)
Mixed, with disclosure	57.79	57.79	57.79	54.65	54.65	54.65
	(3.14)	(3.14)	(2.81)	(2.97)	(2.97)	(2.73)
Homogeneous, without disclosure	57.79	57.79	57.79	57.69	57.69	57.69
	(2.91)	(2.91)	(2.76)	(2.74)	(2.75)	(2.64)
Homogeneous, with disclosure	55.80	55.80	55.80	56.09	56.09	56.09
	(1.96)	(1.98)	(2.05)	(1.97)	(1.99)	(2.03)

Notes: These are the model-implied no-leader participation means, in percentage points, for the six specifications on the preceding page. Predictions are averaged over the no-leader observations in each environment in the corresponding column's estimation sample. Parentheses report group-clustered standard errors.

Table B-2 compares Waves 1–7 with all eight collected waves while holding the estimator fixed. The columns labeled + LASSO use the same 11 LASSO-selected baseline controls in both samples, selected in Waves 1–7 and held fixed when adding Wave 8. The Waves 1–8 columns retain the complete scheduled Wave 8 six-day outcome window, including Friday, May 15, and Saturday, May 16.

Across the three specifications, Δ_E is 19.04–19.60 percentage points in Waves 1–7 and 14.67–15.04 percentage points in Waves 1–8 (SE 5.41–5.56; p-values 0.007–0.008). Adding Wave 8 therefore reduces the point estimate by roughly 4.4–4.7 percentage points. The Waves 1–8 estimate remains positive, and its confidence interval excludes zero in every specification. The two component leader effects retain their signs. The traditional-leader effect in homogeneous groups under disclosure, T_{H1}, is 7.63–8.41 percentage points (p-values 0.018–0.041). The religious-leader effect in mixed groups without disclosure, R_{M0}, is 6.80–7.03 percentage points (p-values 0.055–0.073).

B.2 Randomization and Team-Construction Protocol

Assignment proceeded in waves of four neighborhoods, with twelve four-person accountability teams in each neighborhood. The twelve teams covered all cells of the $3 \times 2 \times 2$ design: a traditional leader, religious leader, or no leader; homogeneous or mixed composition; and individual disclosure or no individual disclosure. We first constructed the eligible roster and sampled residents within sex-by-wealth strata. After forming homogeneous and mixed teams, we chose an assignment under the prespecified rerandomization rule, attached the information and scheduling assignments, and replaced unavailable residents from pre-generated queues. These are the six steps described below. All random draws were made by the assignment program and saved with the assignment files for each wave.

B.2.1 Stage 1: Eligible rosters and fixed participants

The assignment was run separately by wave using the recruitment roster that had been completed before that wave's cutoff date. We retained consenting residents with the information needed to invite them and form their teams. We excluded other identified religious and traditional leaders from the ordinary-resident pool so that no-leader teams would not receive a leader through that pool.

The same chief and pastor appeared across teams within a neighborhood. We identified one traditional and one religious leader before assigning teams, and each joined the teams assigned to his leader condition. From Wave 3, a randomly selected ordinary resident from the chief's territorial origin occupied the fourth seat across no-leader teams. We call this resident the fixed placebo participant. Having that person appear again in each no-leader environment, without any leadership role, gave the no-leader condition the same team size and number of repeated participants as the leader conditions.

B.2.2 Stage 2: Sampling ordinary residents

Within each neighborhood, ordinary residents were divided into four sampling strata defined by sex and predicted housing wealth: female–lower wealth, female–higher wealth, male–lower wealth, and male–higher wealth. Higher and lower wealth were defined relative to the sex-specific median within the neighborhood; residents with missing property values entered the lower-wealth stratum.

The program sampled up to 48 ordinary residents per neighborhood, prioritizing the traditional leader's territorial origin for homogeneous teams and drawing across territories for mixed teams. Homogeneous teams were formed first. Mixed teams drew from the remaining sampled residents, including those sharing the traditional leader's territorial origin, while prioritizing territorial diversity. Thirty-six ordinary residents filled the twelve teams; remaining sampled residents entered the baseline-only control group.

The initial wave targeted equal representation of the four sex-by-wealth strata. In subsequent waves, the quotas were adjusted using the cumulative distribution of assignments in completed waves: an overrepresented stratum received fewer positions in the next wave and an underrepresented stratum received more. The adjusted totals were then distributed across the four neighborhoods entering that wave, with each neighborhood receiving the same total number of positions.

When a requested sex-by-wealth cell did not contain enough eligible residents, the program used a fixed priority order: first the opposite sex within the same wealth category, then the same sex in the other wealth category, and finally any remaining eligible resident in the relevant composition pool. These fallbacks preserved team size while keeping the assignment as close as possible to the target margins.

B.2.3 Stage 3: Forming homogeneous and mixed teams

Each of the twelve group identifiers corresponded to one experimental cell, as shown in Table B-3. The order in which group identifiers were filled was randomized. Each group received three ordinary residents. The appropriate traditional leader, religious leader, or fixed placebo participant was then added as the fourth member.

Homogeneous teams. The program filled groups 1–6 from the pool associated with the traditional leader's territorial origin. It assigned residents to teams in randomized group order and one position at a time, which prevented any group identifier from systematically receiving the first available residents. When a territorial-origin or sampling-stratum constraint was binding, the program used the prespecified fallback order described above to preserve team size.

Table B-3 – Mapping from group identifiers to experimental cells

Group identifiers	Composition	Fourth team member	Individual disclosure
1, 2	Homogeneous	Fixed placebo participant	No, Yes
3, 4	Homogeneous	Traditional leader	No, Yes
5, 6	Homogeneous	Religious leader	No, Yes
7, 8	Mixed	Fixed placebo participant	No, Yes
9, 10	Mixed	Traditional leader	No, Yes
11, 12	Mixed	Religious leader	No, Yes

Notes: Within each row, the odd-numbered group received no individual disclosure and the even-numbered group received individual disclosure. In Waves 1–2, no fixed placebo participant was used; the no-leader teams instead contained four independently sampled ordinary residents.

Mixed teams. For a resident to enter the next mixed-team seat, the sex-by-wealth target had to be met. Among eligible residents, the program preferred someone whose territory was not yet in that team; if that was impossible, it drew from the rest of the eligible mixed-team pool. It filled groups 7–12 one seat at a time in randomized group order, obtaining as much territorial diversity in each team as the available roster allowed.

Repeated participants. The traditional leader appeared in four teams in the same neighborhood: two homogeneous and two mixed, split evenly across information conditions. The religious leader and, from Wave 3 onward, the fixed placebo participant were used in the same way. Leaders participated as ordinary team members and received no script or formal power. The estimation sample excludes the traditional leader, religious leader, and fixed placebo participant.

B.2.4 Stage 4: Cumulative rerandomization

The program generated 48 candidate assignments per wave and selected the assignment with the best cumulative covariate balance, including assignments from earlier waves. Balance was measured by summing joint F-statistics for the leader and composition indicators and their interactions, controlling for neighborhood fixed effects, across the 15 balancing covariates.[29]

B.2.5 Stage 5: Information and time slots

The two group identifiers within every leader-by-composition cell were assigned opposite information conditions. Odd-numbered groups received the status-quo information regime, in which individual contributions were not disclosed; even-numbered groups received individual disclosure. Random allocation across group identifiers produced one team in each information condition for every leader-by-composition cell in every neighborhood.

Finally, the twelve group identifiers were mapped without replacement to the twelve available weekday-by-daypart meeting slots. A group identifier had the same slot across the four neighborhoods in a given wave. In the second half of fieldwork, the slot ordering mirrored the first-half ordering, reversing both day and daypart so that leader and information cells were not systematically concentrated in earlier or later slots.

B.2.6 Stage 6: Unavailable residents and replacements

Before field implementation, all eligible residents not selected for the initial roster received a random replacement order within neighborhood and territorial-origin category. If a selected ordinary resident was unavailable, the program assigned the first available person in the corresponding queue, prioritizing the same neighborhood and territorial-origin category before relaxing the match within neighborhood. For mixed teams, the fallback continued to prioritize residents outside the traditional leader's territorial origin. Replacement residents inherited the unavailable resident's group, experimental condition, and meeting slot. Leader replacements were handled separately and preserved leader type.

[29] Waves 1–2 followed an earlier version of the protocol, with minor differences in sampling and the balance criterion. No-leader teams contained four independently sampled residents rather than three residents and a fixed placebo participant.

B.2.7 Implementation refinements across waves

The assignment logic was stable across the main sample, with two implementation refinements:

1. In Waves 1–2, no-leader teams contained four independently sampled ordinary residents. Beginning in Wave 3, one fixed placebo participant occupied the fourth seat in all four no-leader teams in a neighborhood, leaving three non-repeated residents in every experimental team.
2. Beginning in Wave 3, the sampling quotas explicitly offset cumulative sex-by-wealth differences across group identifiers from completed waves. Candidate assignments had already been evaluated cumulatively beginning in Wave 2.

These changes standardized the fourth team position and improved cumulative balance. They did not alter the mapping between group identifiers and the leader, composition, or information treatments.

B.2.8 Pseudocode

Input: eligible recruitment roster for four neighborhoods in wave w; fixed traditional and religious leaders; pretreatment covariates; assignments from waves $1, \ldots, w-1$.

For each neighborhood:
  identify the traditional leader's territorial origin;
  select the fixed placebo participant when $w \geq 3$;
  divide ordinary residents into sex $\times$ wealth strata;
  create candidate pools for homogeneous and mixed teams;
  set stratum quotas, offsetting cumulative imbalances from earlier waves.

For candidate seed $s = 1, \ldots, S_w$**:**
  sample residents within composition pools and strata;
  generate complementary three-person stratum patterns across neighborhoods;
  randomly order homogeneous group identifiers 1-6;
  fill each homogeneous team slot-by-slot, prioritizing the leader's origin;
  randomly order mixed group identifiers 7-12;
  fill each mixed team slot-by-slot, prioritizing a new territory;
  add the traditional leader to groups 3, 4, 9, and 10;
  add the religious leader to groups 5, 6, 11, and 12;
  add the placebo participant to groups 1, 2, 7, and 8 when $w \geq 3$;
  append the candidate to completed-wave assignments;
  score candidate s using the prespecified rule in Stage 4 above.

Select the seed with the lowest prespecified assignment score.
Assign no disclosure to odd group identifiers and individual disclosure to even group identifiers.
Attach the wave's without-replacement time-slot permutation.
Export the assignment roster, seed, assignment diagnostics, and ordered replacement queues.

$S_w = 48$ under the final protocol.

B.3 Implementation Checks

Table B-4 – ASSIGNMENT BALANCE AND PRETREATMENT PARTICIPATION

	No leader				Leader − no leader, within environment								
Characteristic	$M0$ mean	$M1$ $-M0$	$H0$ $-M0$	$H1$ $-M0$	T $M0$	R $M0$	T $M1$	R $M1$	T $H0$	R $H0$	T $H1$	R $H1$	N
Panel A. Characteristics used in rerandomization													
Female	0.565	-0.043	-0.033	-0.022	-0.027	0.021	0.076	-0.007	-0.006	0.017	0.007	-0.005	1,040
	(0.057)	(0.071)	(0.080)	(0.072)	(0.066)	(0.077)	(0.073)	(0.064)	(0.082)	(0.078)	(0.065)	(0.072)	
Predicted property value (USD per cubic meter)	2444.612	-167.307	-75.311	113.918	26.979	-214.419	64.212	624.430	120.720	324.920	-293.164	-290.367	968
	(280.312)	(313.857)	(317.962)	(377.632)	(350.713)	(278.422)	(375.699)	(403.680)	(332.820)	(517.079)	(438.059)	(433.837)	
Community role	0.783	-0.011	0.000	0.054	0.038	0.002	-0.046	-0.058	-0.010	-0.022	-0.040	-0.076	1,040
	(0.049)	(0.058)	(0.055)	(0.053)	(0.056)	(0.059)	(0.062)	(0.066)	(0.052)	(0.051)	(0.046)	(0.054)	
Age	34.554	-1.185	-0.293	-0.815	-0.354	-0.711	1.105	0.081	0.499	2.464	1.581	-0.193	1,040
	(1.770)	(1.877)	(1.956)	(2.006)	(2.334)	(2.238)	(2.100)	(1.794)	(2.069)	(1.760)	(2.163)	(1.771)	
Education (years)	9.511	0.130	-0.130	0.402	0.638	0.007	-0.135	0.532	0.352	0.507	-0.157	-0.538	1,040
	(0.353)	(0.480)	(0.494)	(0.514)	(0.556)	(0.506)	(0.491)	(0.445)	(0.457)	(0.462)	(0.482)	(0.455)	
Government job	0.000	0.011	0.011	0.000	0.012	0.000	-0.011	-0.011	-0.011	0.001	0.012	0.000	1,040
	(0.000)	(0.010)	(0.010)	(0.003)	(0.011)	(0.003)	(0.011)	(0.011)	(0.010)	(0.015)	(0.011)	(0.003)	
Salaried job	0.043	0.000	0.000	0.000	-0.006	-0.018	-0.006	0.030	0.006	0.053	0.006	0.030	1,040
	(0.021)	(0.032)	(0.030)	(0.028)	(0.028)	(0.025)	(0.031)	(0.034)	(0.030)	(0.040)	(0.030)	(0.034)	
Mission church	0.174	-0.011	0.022	0.054	-0.005	-0.005	-0.018	-0.006	0.057	0.009	-0.059	0.024	1,040
	(0.050)	(0.060)	(0.064)	(0.064)	(0.064)	(0.070)	(0.054)	(0.048)	(0.064)	(0.062)	(0.057)	(0.064)	
Revival church	0.489	0.033	0.098	-0.033	-0.024	0.107	0.003	-0.021	-0.075	-0.098	-0.016	0.008	1,040
	(0.056)	(0.077)	(0.076)	(0.076)	(0.079)	(0.085)	(0.064)	(0.074)	(0.069)	(0.073)	(0.063)	(0.062)	
Reported prior salongo participation	0.034	-0.134	-0.098	0.048	0.113	-0.165	0.019	0.217	0.161	-0.031	-0.247**	0.099	1,040
	(0.067)	(0.124)	(0.117)	(0.125)	(0.106)	(0.124)	(0.150)	(0.145)	(0.155)	(0.126)	(0.119)	(0.172)	
Salongo contribution index	0.145	-0.224	-0.319**	-0.032	-0.053	-0.112	0.174	0.088	0.233	0.252*	-0.225*	-0.226	1,040
	(0.123)	(0.147)	(0.136)	(0.164)	(0.164)	(0.171)	(0.131)	(0.130)	(0.142)	(0.143)	(0.135)	(0.141)	
Biometric support index	0.040	-0.075	0.033	0.111	0.028	0.112	0.120	0.121	-0.107	0.047	-0.268*	-0.040	1,040
	(0.120)	(0.146)	(0.139)	(0.136)	(0.152)	(0.152)	(0.142)	(0.132)	(0.122)	(0.152)	(0.146)	(0.144)	
Household size	8.380	-0.815	-1.130*	-0.228	-0.200	-1.212**	0.020	1.305**	0.394	0.835	1.028*	0.349	1,040
	(0.447)	(0.552)	(0.575)	(0.546)	(0.857)	(0.574)	(0.551)	(0.521)	(0.528)	(0.632)	(0.557)	(0.634)	
Adult share	0.426	0.041	0.027	0.002	0.039	0.025	-0.056	-0.033	0.028	0.018	0.003	0.032	1,039
	(0.017)	(0.035)	(0.032)	(0.026)	(0.029)	(0.029)	(0.037)	(0.037)	(0.034)	(0.035)	(0.032)	(0.028)	
Distance to centroid (m)	233.250	-1.463	1.625	-7.836	-4.202	5.058	-5.825	-18.727	-15.452	-1.795	5.550	7.423	1,033
	(31.416)	(15.191)	(14.124)	(13.777)	(15.463)	(15.604)	(14.104)	(14.445)	(12.589)	(12.635)	(12.018)	(13.108)	
Panel B. Additional baseline-survey characteristics													
Demographic and economic characteristics													
Log income	11.474	-0.162	0.243	-0.020	0.008	0.183	0.277	0.209	-0.209	-0.366*	0.411	0.370	1,024
	(0.207)	(0.262)	(0.181)	(0.258)	(0.244)	(0.186)	(0.283)	(0.272)	(0.209)	(0.194)	(0.250)	(0.248)	
Avenue tenure	17.272	-3.351*	1.168	-1.778	-0.400	-2.469	2.079	1.358	-5.445***	-3.081	1.596	-0.589	1,012
	(1.690)	(1.791)	(2.177)	(2.224)	(2.131)	(1.785)	(1.733)	(1.873)	(2.064)	(2.242)	(2.029)	(2.172)	

Continued on next page

Assignment Balance and Pretreatment Participation (continued)

	No leader				Leader − no leader, within environment								
Characteristic	$M0$ mean	$M1$ $-M0$	$H0$ $-M0$	$H1$ $-M0$	T $M0$	R $M0$	T $M1$	R $M1$	T $H0$	R $H0$	T $H1$	R $H1$	N
Traditional authority and local embeddedness													
Any traditional role	0.098	-0.042	-0.043	-0.053	-0.037	-0.050	-0.043	0.052	0.019	0.030	-0.008	-0.019	1,032
	(0.031)	(0.037)	(0.036)	(0.035)	(0.039)	(0.037)	(0.026)	(0.041)	(0.033)	(0.037)	(0.028)	(0.026)	
Traditional head role	0.065	-0.043	-0.043	-0.043	-0.016	-0.064**	-0.009	0.038	-0.009	0.027	-0.009	-0.021	1,032
	(0.029)	(0.030)	(0.030)	(0.031)	(0.034)	(0.027)	(0.019)	(0.028)	(0.019)	(0.031)	(0.019)	(0.016)	
Secrets network: same territory	0.286	-0.065	0.075	0.079	-0.069	-0.053	-0.014	-0.003	0.053	-0.028	0.020	-0.033	1,032
	(0.054)	(0.060)	(0.070)	(0.062)	(0.064)	(0.062)	(0.056)	(0.053)	(0.080)	(0.070)	(0.064)	(0.062)	
Material network: same territory	0.383	-0.101	0.148***	0.176***	-0.074	-0.014	0.077	0.036	0.097*	-0.034	-0.058	0.012	1,032
	(0.051)	(0.064)	(0.056)	(0.064)	(0.057)	(0.065)	(0.065)	(0.061)	(0.055)	(0.060)	(0.060)	(0.070)	
Planned salongo hours	17.293	0.224	-1.174	-0.203	-0.554	-0.812	1.282	-0.365	0.697	0.184	1.279	-1.201	1,022
	(1.704)	(1.706)	(1.391)	(1.588)	(1.873)	(1.688)	(2.358)	(1.987)	(1.623)	(1.403)	(1.632)	(1.651)	
Religious authority and church embeddedness													
Any religious role	0.391	-0.007	0.033	-0.062	-0.042	-0.022	-0.120*	-0.015	0.001	-0.078	0.059	0.081	1,032
	(0.057)	(0.072)	(0.066)	(0.065)	(0.073)	(0.077)	(0.069)	(0.075)	(0.062)	(0.062)	(0.064)	(0.071)	
Religious head role	0.098	-0.043	-0.043	-0.043	-0.025	-0.062*	-0.055**	-0.007	0.019	0.017	-0.018	0.005	1,032
	(0.027)	(0.034)	(0.034)	(0.034)	(0.040)	(0.032)	(0.024)	(0.030)	(0.034)	(0.033)	(0.030)	(0.037)	
Secrets network: same church	0.202	-0.021	-0.042	-0.010	-0.067	-0.031	-0.047	0.036	0.012	0.030	0.030	-0.035	1,032
	(0.042)	(0.049)	(0.049)	(0.051)	(0.047)	(0.057)	(0.048)	(0.054)	(0.056)	(0.044)	(0.055)	(0.049)	
Material network: same church	0.213	0.012	-0.019	0.006	-0.043	0.006	0.009	0.031	0.069	0.038	-0.011	0.033	1,032
	(0.037)	(0.050)	(0.044)	(0.044)	(0.048)	(0.050)	(0.056)	(0.053)	(0.052)	(0.050)	(0.046)	(0.045)	
Sunday attendance	3.467	-0.031	-0.152	-0.097	-0.079	0.178	-0.061	0.209	0.003	0.235	0.231	0.282**	1,032
	(0.087)	(0.144)	(0.148)	(0.136)	(0.127)	(0.124)	(0.162)	(0.145)	(0.162)	(0.147)	(0.151)	(0.135)	
Weekday attendance	1.967	-0.286	-0.413**	-0.198	-0.450*	-0.222	-0.183	0.040	-0.067	0.060	-0.210	-0.079	1,032
	(0.187)	(0.211)	(0.171)	(0.224)	(0.240)	(0.218)	(0.236)	(0.215)	(0.211)	(0.167)	(0.271)	(0.237)	
Log religious contribution	6.365	-0.283	0.023	0.000	-0.046	-0.056	-0.225	0.165	0.132	-0.190	-0.006	0.319	999
	(0.307)	(0.417)	(0.370)	(0.350)	(0.437)	(0.347)	(0.495)	(0.456)	(0.354)	(0.393)	(0.415)	(0.348)	
Knowledge of local authorities													
Knows authorities	0.312	0.019	-0.004	-0.044	-0.015	-0.028	-0.048	0.005	0.016	-0.017	0.035	0.000	1,032
	(0.034)	(0.040)	(0.035)	(0.037)	(0.040)	(0.039)	(0.041)	(0.038)	(0.036)	(0.035)	(0.038)	(0.038)	
Knows quartier chief	0.304	0.015	0.043	0.004	0.001	-0.033	0.005	-0.001	-0.016	-0.017	0.042	-0.022	1,032
	(0.057)	(0.051)	(0.047)	(0.047)	(0.054)	(0.058)	(0.060)	(0.054)	(0.053)	(0.050)	(0.056)	(0.052)	
Knows locality chief	0.359	0.026	0.011	-0.084	-0.041	0.040	-0.033	0.026	-0.023	-0.019	0.053	0.059	1,032
	(0.055)	(0.065)	(0.058)	(0.055)	(0.063)	(0.057)	(0.072)	(0.068)	(0.058)	(0.057)	(0.053)	(0.057)	
Knows chief phone	0.076	-0.021	0.011	-0.032	-0.014	-0.038	0.044	0.091**	-0.012	-0.037	0.019	0.067*	1,032
	(0.025)	(0.035)	(0.037)	(0.033)	(0.034)	(0.031)	(0.036)	(0.042)	(0.040)	(0.034)	(0.028)	(0.039)	
Chief help index	1.446	0.264	-0.029	0.018	0.064	0.103	-0.090	-0.327*	0.317*	0.029	-0.138	0.089	1,032
	(0.152)	(0.174)	(0.172)	(0.174)	(0.182)	(0.161)	(0.194)	(0.172)	(0.170)	(0.160)	(0.174)	(0.190)	

Continued on next page

Assignment Balance and Pretreatment Participation (continued)

	No leader				Leader − no leader, within environment								
Characteristic	$M0$ mean	$M1$ $-M0$	$H0$ $-M0$	$H1$ $-M0$	T $M0$	R $M0$	T $M1$	R $M1$	T $H0$	R $H0$	T $H1$	R $H1$	N
Standard psychological measures													
Big Five Inventory (main index)	0.068	-0.040	-0.109	-0.114	-0.115*	-0.051	-0.047	0.009	0.031	0.114	0.037	0.027	1,032
	(0.048)	(0.067)	(0.071)	(0.074)	(0.065)	(0.072)	(0.070)	(0.076)	(0.075)	(0.072)	(0.079)	(0.073)	
Extraversion	3.261	-0.063	-0.053	-0.047	-0.004	-0.047	0.005	-0.014	0.033	0.006	-0.052	-0.035	1,031
	(0.057)	(0.075)	(0.082)	(0.081)	(0.081)	(0.078)	(0.080)	(0.070)	(0.081)	(0.079)	(0.080)	(0.088)	
Agreeableness	3.957	-0.066	-0.114	-0.176*	-0.263***	-0.012	-0.073	0.000	-0.085	0.054	0.145	0.125	1,032
	(0.064)	(0.089)	(0.092)	(0.099)	(0.091)	(0.091)	(0.092)	(0.101)	(0.097)	(0.096)	(0.116)	(0.094)	
Conscientiousness	3.957	0.039	-0.076	-0.099	-0.079	-0.203*	-0.128	0.014	0.014	0.040	0.165*	0.014	1,030
	(0.069)	(0.100)	(0.091)	(0.092)	(0.088)	(0.104)	(0.105)	(0.097)	(0.084)	(0.099)	(0.095)	(0.096)	
Neuroticism	2.924	0.005	-0.027	-0.039	-0.033	-0.138	-0.176*	-0.220**	-0.068	-0.135	0.139	-0.004	1,032
	(0.062)	(0.091)	(0.107)	(0.100)	(0.087)	(0.096)	(0.093)	(0.098)	(0.115)	(0.114)	(0.102)	(0.110)	
Openness	3.266	-0.027	-0.152	-0.099	-0.084	-0.030	-0.145	-0.199*	0.077	0.170	0.013	-0.007	1,032
	(0.078)	(0.105)	(0.112)	(0.098)	(0.104)	(0.098)	(0.109)	(0.116)	(0.119)	(0.126)	(0.090)	(0.095)	
Emotional intelligence	4.353	-0.064	-0.071	0.028	-0.129*	-0.076	-0.017	0.001	-0.013	0.066	-0.116*	-0.044	1,032
	(0.057)	(0.066)	(0.066)	(0.062)	(0.071)	(0.080)	(0.073)	(0.062)	(0.070)	(0.067)	(0.060)	(0.066)	
Locus of control	4.065	-0.083	-0.163	0.191	0.008	0.149	0.238	-0.125	0.082	0.217	-0.246	0.045	1,032
	(0.144)	(0.171)	(0.171)	(0.175)	(0.174)	(0.173)	(0.186)	(0.182)	(0.179)	(0.204)	(0.182)	(0.185)	
Social desirability	0.574	-0.009	-0.028	-0.031	-0.015	-0.043*	0.000	0.020	0.016	0.031*	0.046**	0.020	1,032
	(0.019)	(0.019)	(0.019)	(0.019)	(0.020)	(0.022)	(0.021)	(0.021)	(0.020)	(0.017)	(0.021)	(0.019)	
Religious psychological measures													
Positive RCOPE	3.850	-0.025	-0.087	-0.080	-0.087	-0.032	0.000	0.012	0.049	-0.071	0.032	0.089	1,032
	(0.043)	(0.053)	(0.064)	(0.067)	(0.063)	(0.062)	(0.051)	(0.049)	(0.068)	(0.078)	(0.071)	(0.065)	
Negative RCOPE	3.391	0.203**	0.220***	0.140	0.121	0.203**	-0.090	-0.012	-0.093	-0.140	-0.083	0.030	1,032
	(0.070)	(0.083)	(0.084)	(0.089)	(0.101)	(0.088)	(0.081)	(0.088)	(0.099)	(0.088)	(0.100)	(0.095)	
Intrinsic religiosity	3.376	-0.046	-0.036	-0.046	-0.034	-0.002	0.005	0.002	-0.078*	-0.023	0.027	-0.053	1,032
	(0.022)	(0.037)	(0.036)	(0.038)	(0.043)	(0.033)	(0.038)	(0.039)	(0.042)	(0.045)	(0.048)	(0.045)	
Extrinsic religiosity: personal	4.446	0.026	-0.043	0.136*	-0.188**	-0.019	-0.040	-0.092	-0.036	-0.035	-0.101	-0.179**	1,030
	(0.060)	(0.081)	(0.079)	(0.074)	(0.091)	(0.092)	(0.090)	(0.080)	(0.083)	(0.094)	(0.084)	(0.086)	
Extrinsic religiosity: social	2.743	-0.139	0.022	-0.157	-0.193	-0.069	0.056	-0.081	-0.111	-0.206	0.166	0.075	1,032
	(0.118)	(0.125)	(0.126)	(0.136)	(0.127)	(0.134)	(0.136)	(0.137)	(0.121)	(0.126)	(0.148)	(0.150)	
Panel C. Post-assignment, pre-intervention participation													
Participation rate over six common pre-treatment days	0.404	-0.069*	-0.029	-0.009	-0.126***	-0.073**	0.002	0.040	-0.022	0.018	-0.085**	-0.068*	1,040
	(0.057)	(0.038)	(0.042)	(0.034)	(0.040)	(0.036)	(0.043)	(0.046)	(0.040)	(0.043)	(0.040)	(0.040)	

Notes: We compare residents' characteristics across the randomized groups before the first meeting. Panel A covers the characteristics used to form the groups, Panel B adds measures from the baseline survey, and Panel C reports administrative participation measured after assignment but before residents learned their treatment. The first column is the average for no-leader residents in mixed groups without disclosure ($M0$). The next three columns compare the other no-leader environments with that group. The final eight columns compare chief and pastor assignments with no leader in the same environment. M and H mean mixed and homogeneous groups; 0 and 1 mean group and individual disclosure. Comparisons are made within neighborhood, and standard errors in parentheses allow for residents assigned to the same group. The table is discussed in Section 3; Figure 2 summarizes the design. Stars use unadjusted two-sided p-values. $^{*}p < 0.10$, $^{**}p < 0.05$, $^{***}p < 0.01$.

Table B-5 – ROSTER FORMATION, REPLACEMENTS, AND COMPOSITION FIDELITY

Leader	Composition	Disclosure	Assigned residents	Replacement share	Mean territories	Final chief-territory share	Composition fidelity
No leader	Mixed	Without disclosure	92	21.7	3.04	22.8	96.4
		With disclosure	92	18.5	3.07	22.8	100.0
	Homogeneous	Without disclosure	92	17.4	1.07	97.8	92.9
		With disclosure	92	22.8	1.07	97.8	92.6
Traditional	Mixed	Without disclosure	84	26.2	2.61	23.8	96.4
		With disclosure	84	27.4	2.82	28.6	100.0
	Homogeneous	Without disclosure	84	21.4	1.07	96.4	92.9
		With disclosure	84	25.0	1.11	95.2	89.3
Religious	Mixed	Without disclosure	84	29.8	2.79	26.2	96.4
		With disclosure	84	21.4	2.82	26.2	100.0
	Homogeneous	Without disclosure	84	15.5	1.07	97.6	92.9
		With disclosure	84	20.2	1.18	94.0	85.7
Total			1,040	22.2	1.98	60.7	94.6

Notes: Did the groups formed for the meetings follow the original assignments? Replacement share counts the percentage of residents whose final roster entry differs from their initial assignment. Mean territories counts the origins represented among ordinary residents, and Final chief-territory share gives the percentage from the chief's territory. We count composition as correct when all observed residents of a homogeneous group come from that territory, or when a mixed group contains at least two territories. Territory is observed in 335 of the 336 groups. All shares are percentages.
There are 92 residents in the no-leader cells rather than 84. In Waves 1–2, a different ordinary resident occupied the fourth seat in each no-leader team, and these residents remain in the analysis. From Wave 3, one person filled the seat in all four no-leader teams and was excluded, as were the repeated chief and pastor. Section 3 discusses the table; Figure 2 places these groups in the experiment.

Table B-6 – TREATMENT DELIVERY AT THE ACCOUNTABILITY MEETINGS

Leader	Composition	Disclosure	Assigned residents	Resident attendance	Realized group size	Full-group share	Leader attendance	M1 announcement verified	M2 record verified
No leader	Mixed	Without disclosure	92	95.7	3.82	82.1	–	96.4	100.0
		With disclosure	92	95.7	3.75	75.0	–	100.0	100.0
	Homogeneous	Without disclosure	92	97.8	3.82	85.7	–	100.0	100.0
		With disclosure	92	91.3	3.64	64.3	–	100.0	100.0
Traditional	Mixed	Without disclosure	84	94.0	3.82	85.7	100.0	100.0	100.0
		With disclosure	84	96.4	3.86	85.7	96.4	96.4	100.0
	Homogeneous	Without disclosure	84	92.9	3.79	78.6	100.0	100.0	100.0
		With disclosure	84	94.0	3.82	82.1	100.0	100.0	100.0
Religious	Mixed	Without disclosure	84	95.2	3.86	89.3	100.0	100.0	100.0
		With disclosure	84	96.4	3.89	89.3	100.0	100.0	100.0
	Homogeneous	Without disclosure	84	96.4	3.89	89.3	100.0	100.0	100.0
		With disclosure	84	94.0	3.82	82.1	100.0	100.0	100.0
Total			1,040	95.0	3.82	82.4	99.6	99.4	100.0

Notes: For resident attendance, we ask whether an ordinary resident from the analysis roster was present. For leader attendance, whether the assigned chief or pastor was present. Realized group size counts attending people from the entire scheduled team, including the leader or no-leader participant, and a full group requires all four. We verify two deliveries at the meetings: announcement of the assigned information at Meeting 1 and presentation of the team's performance record at Meeting 2. These are what the final two columns count; assignment alone is insufficient. Attendance and verification are percentages; size is a count.
Why do no-leader cells contain 92 residents, compared with 84 elsewhere? In Waves 1–2, a different ordinary resident filled the fourth seat of each no-leader team, and we keep these residents in the analysis. Starting in Wave 3, one person instead filled that seat in all four teams; as with the leaders serving repeatedly, we exclude this person. Section 3 discusses the table, and Figure 2 shows the experimental cells.

Table B-7 – ADMINISTRATIVE, SURVEY, AND AUDIO COVERAGE

Leader	Composition	Disclosure	Assigned residents	Baseline survey	Meeting 1 usable	Meeting 2 usable	Endline survey	Audio recorded	Administrative outcome
No leader	Mixed	Without disclosure	92	100.0	95.7	93.5	88.0	92.9	100.0
		With disclosure	92	98.9	95.7	90.2	95.7	96.4	100.0
	Homogeneous	Without disclosure	92	100.0	97.8	94.6	93.5	85.7	100.0
		With disclosure	92	98.9	91.3	92.4	97.8	96.4	100.0
Traditional	Mixed	Without disclosure	84	100.0	94.0	92.9	94.0	96.4	100.0
		With disclosure	84	98.8	96.4	90.5	92.9	96.4	100.0
	Homogeneous	Without disclosure	84	98.8	92.9	85.7	94.0	92.9	100.0
		With disclosure	84	98.8	94.0	92.9	97.6	89.3	100.0
Religious	Mixed	Without disclosure	84	100.0	95.2	89.3	85.7	89.3	100.0
		With disclosure	84	100.0	96.4	94.0	95.2	100.0	100.0
	Homogeneous	Without disclosure	84	100.0	96.4	90.5	96.4	100.0	100.0
		With disclosure	84	98.8	94.0	86.9	86.9	82.1	100.0
Total			1,040	99.4	95.0	91.2	93.2	93.2	100.0

Notes: A form reporting that a resident was unavailable at Meeting 1 or Meeting 2 is not an interview. The survey columns therefore count only usable interviews as a percentage of the relevant residents. For audio, we instead ask what percentage of the randomized groups have a verified recording. For administrative data, we ask what percentage of residents have the participation outcome used in the main results. This outcome uses the same six days for each covered resident, after Meeting 1 through the day of Meeting 2. All entries are coverage percentages for the corresponding residents or groups.
Why do the no-leader cells have 92 residents rather than 84? Waves 1–2 used a different fourth ordinary resident in each no-leader team, so those residents remain in the sample. From Wave 3 onward, the same person joined all four teams. We excluded that repeated participant, as we do the repeated leaders. See Section 3 for discussion and Figure 2 for the experimental cells.

B.4 Meeting Discourse

We classify what was said at the first meeting using the words in the transcripts. After cleaning them, we split speakers' turns into short passages and fit a single-topic mixture model for short texts (Quinn et al., 2010). Which words and phrases occur together determines the three topics. We leave treatment assignments, speaker roles, and participation outcomes out of this exercise, and assign each passage to the topic it most likely belongs to. Characteristic words and representative passages suggest attendance and performance, material support, and divine reward and solidarity as the three interpretations (Table B-8). Additional transcribed meetings use the same fitted model and vocabulary, leaving the earlier assignments fixed. There are 2,691 classified passages from 313 meetings.

We allocate each recorded turn's duration across its passages in proportion to their word counts, then sum minutes by speaker and topic. Figure B-4 uses 969 ordinary participants in recorded meetings, retaining 62 silent participants with zero minutes. Unrecorded meetings are excluded. Regressions compare each leader assignment with no leader within neighborhood-by-composition-by-disclosure cells, with standard errors clustered by assigned group. No pretreatment speech measure is available. These comparisons measure responses in the discussion; they do not identify the causal effect of particular messages on later participation.

Robustness across text models. We compare the single-topic model and its full-corpus refit (Quinn et al., 2010), standard and sparse-prior LDA (Blei et al., 2003), a correlated topic model (Blei and Lafferty, 2007), spherical k-means on multilingual E5 embeddings (Dhillon and Modha, 2001; Wang et al., 2024), and BERTopic with three topics (Grootendorst, 2022). Figures B-2 and B-3 compare leaders' pooled topic shares and their speech within each composition-by-disclosure environment. Figure B-4 compares ordinary participants' speaking levels, including the no-leader arm.

All models classify the same 2,691 passages and use the same analysis samples: 206 speaking leader-meetings and 969 ordinary participants. In models that assign each passage to one topic, we give that topic the passage's full duration. In mixed-membership models, we divide the duration using posterior topic weights. Leader figures show equal-weight means across leader-meetings. Participant levels use the fixed-effect regressions described above, averaging each assignment's predictions over the actual no-leader participants in the corresponding environment. Using the same residents for all three predictions means that differences in predicted levels recover the original treatment effects. Confidence intervals use the full covariance matrix and cluster by neighborhood for leaders and by assigned group for participants. They are pointwise and conditional on the fitted text measures. Chiefs' emphasis on attendance and pastors' emphasis on religious appeals persist across models. Pooling disclosure, pastor assignments increase participants' divine speech in both compositions, alongside more total speech in homogeneous groups and less attendance-and-performance talk in mixed groups. These broad directions persist across models; chiefs do not consistently shift participants toward attendance talk.

Table B-8 – EXAMPLES OF THE THREE SPEECH TOPICS

Topic and typical words	Illustrative speech
Attendance and performance *hours, week, absence, percentage, punctuality*	"Last week we completed only a few hours. We should reach five hours per person each day and come regularly." "I call on those who did not work last week to join us so that we can maintain the places where we live." "Everyone was present this week. Next week, let us work with the same commitment to improve our neighborhood." **Chief:** "Doing this work is like being on a football team: those who want to win compete for the cup. [. . .] We must improve our performance to win the contest. [. . .] We must encourage those who did not come and go and see them to find out why they were absent. We must help one another so that our group does better." **Resident:** "[. . .] I ask that encouragement become a habit in our group. If I am absent, I should be reminded to come and take part in the work [. . .]."
Divine reward and solidarity *God, bless, reward, united, together*	"Salongo is group work, not one person's work. With God's support, let us have patience, peace, and gentleness." "Volunteers should not expect wages, but look to God, who rewards people according to their work." "A person reaps what they sow; God rewards each person according to their deeds." **Pastor:** "This work is voluntary and meant to help everyone. [. . .] It is not a person on earth who can pay us; it is God who will do so. I ask only for each person's willingness to work." **Resident:** "I liked what you said. [. . .] God will pay us, and let us respect the working hours, from nine until three." **Pastor, closing:** "[. . .] May God bless you, and may he also bless our group so that we can move forward."
Requests for material support *tools, sacks, hoes, wheelbarrows, avenues*	"The municipality gave us hoes, machetes, spades, and sacks to improve our environment." "Volunteers brought sacks, and the work is progressing. We still need wheelbarrows, hoes, picks, and more sacks." "We now have the tools, and the people with strength to work are here. We can keep improving the neighborhood." **Resident:** "We have started salongo well, and now we are approaching the second week. We ask for help so that we can maintain our streets: we need tools and empty sacks. We will keep doing this voluntarily to maintain our neighborhood and our streets. We ask you to support us so that we can do this work."

Notes: The translated excerpts let the reader see the different appeals captured by the three topics. Frequent words characterizing a topic appear on the left. On the right, examples of chief and pastor speech assigned to that topic are followed by additional illustrative excerpts with speaker labels. Residents also speak in these exchanges, making requests and responding to leaders. Ellipses indicate omissions within an excerpt, and we shorten the examples to make them readable. Speaker labels let the reader follow the appeals and responses. They do not identify what a particular speech caused. We gave the topic model words from the speeches, without telling it the leader type or treatment assignment.
Section 3.5 discusses these excerpts. For the broader pattern, Figure 3 reports leaders' mean topic shares. Appendix Table B-9 also reports speaking time and distinguishes the experimental environments. Appendix Figure B-4 asks how leader assignment changes ordinary participants' speech.

Table B-9 – LEADER SPEECH IN THE FIRST MEETING

	Speaking time (minutes) (1)	Attendance and performance (%) (2)	Divine reward and solidarity (%) (3)	Material support (%) (4)
I. Group composition (pooled across disclosure)				
Panel A. Adjusted levels				
Traditional leader				
Mixed	2.69	47.18	29.32	23.49
	(0.14)	(3.06)	(2.81)	(2.54)
Homogeneous	2.72	39.59	28.09	32.33
	(0.15)	(3.38)	(3.10)	(2.79)
Religious leader				
Mixed	3.46	22.27	51.01	26.72
	(0.15)	(2.79)	(2.91)	(2.10)
Homogeneous	3.45	27.28	49.27	23.45
	(0.15)	(2.84)	(2.96)	(2.15)
Panel B. Comparisons				
Traditional leader				
Homogeneous minus mixed	0.03	-7.60	-1.24	8.84
	(0.29)	(6.44)	(5.91)	(5.33)
Religious leader				
Homogeneous minus mixed	0.00	5.01	-1.74	-3.27
	(0.29)	(5.63)	(5.87)	(4.25)
II. Group composition by disclosure				
Panel C. Adjusted levels				
Traditional leader				
Mixed, without disclosure	2.65	49.96	29.51	20.53
	(0.17)	(5.53)	(4.15)	(3.63)
Mixed, with disclosure	2.74	44.35	29.14	26.52
	(0.19)	(3.31)	(3.44)	(3.58)
Homogeneous, without disclosure	2.53	38.02	32.82	29.16
	(0.23)	(4.77)	(4.75)	(5.17)
Homogeneous, with disclosure	2.93	41.27	23.16	35.57
	(0.15)	(6.89)	(5.66)	(5.53)
Religious leader				
Mixed, without disclosure	3.75	24.06	47.64	28.29
	(0.22)	(4.62)	(5.69)	(5.03)
Mixed, with disclosure	3.18	20.69	54.05	25.26
	(0.18)	(3.75)	(5.16)	(3.62)
Homogeneous, without disclosure	3.22	28.80	48.48	22.72
	(0.20)	(3.94)	(4.61)	(3.15)
Homogeneous, with disclosure	3.74	25.32	50.32	24.35
	(0.28)	(4.07)	(5.11)	(4.32)
Panel D. Comparisons				
Traditional leader				
Disclosure effect, mixed	0.09	-5.62	-0.37	5.99
	(0.23)	(6.72)	(5.12)	(5.10)
Disclosure effect, homogeneous	0.40	3.25	-9.66	6.41
	(0.24)	(9.66)	(8.37)	(9.12)
Religious leader				
Disclosure effect, mixed	-0.56**	-3.37	6.41	-3.03
	(0.27)	(6.21)	(9.15)	(7.61)
Disclosure effect, homogeneous	0.52	-3.48	1.84	1.64
	(0.37)	(5.59)	(7.69)	(6.09)
Leader–meetings, chiefs	103	103	103	103
Leader–meetings, pastors	103	103	103	103
Neighborhoods per leader type	28	28	28	28

Notes: Each observation is an assigned chief or pastor speaking at a recorded first meeting. We compare homogeneous and mixed groups for each leader type, pooling disclosure conditions, then add disclosure and its interaction with composition to estimate with-minus-without-disclosure differences within each composition. Each regression includes leader fixed effects; to calculate adjusted levels, we use the same observed leader weights within each type. The regressions include neither a pretreatment speech measure nor a participation control.
Speaking time is measured in minutes. To calculate topic shares, we assign each passage to one of the three topics described in Section 3.5 and weight it by its duration. Shares are percentages of modeled speaking time. Additional passages are classified using the same fitted model and vocabulary, leaving earlier assignments unchanged. Differences are in minutes for speaking time and percentage points for topic shares. Standard errors in parentheses are clustered by neighborhood. Stars on differences use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

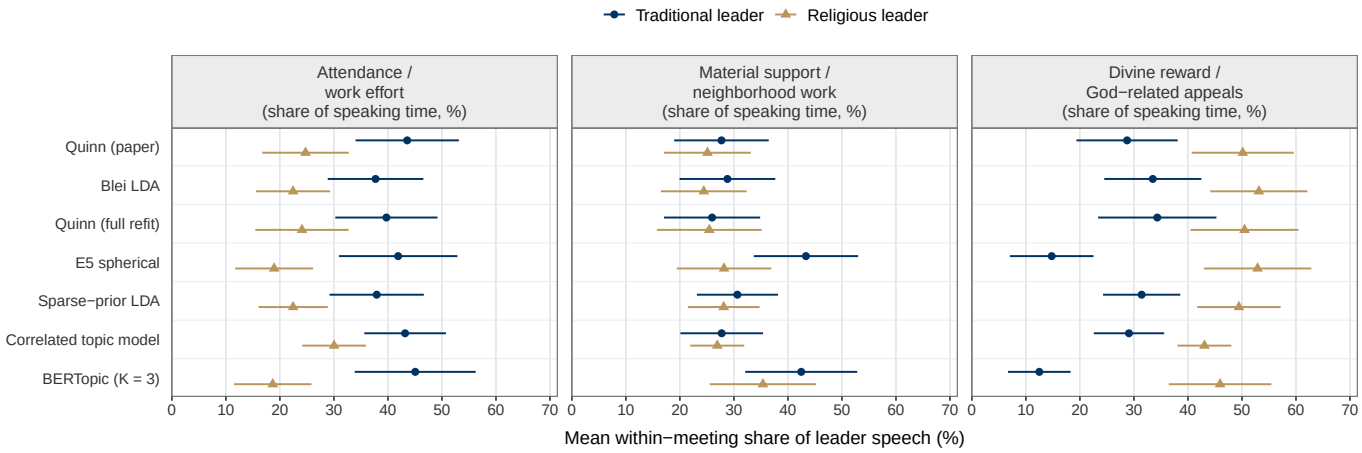


Figure B-2: LEADER TOPIC SHARES ACROSS TEXT MODELS

Notes: Points show mean within-meeting shares of leader speaking time, pooling composition and disclosure. Blue circles denote traditional leaders and gold triangles religious leaders. Each model uses the same 103 speaking leader-meetings of each type. Lines are pointwise 95 percent confidence intervals clustered by neighborhood (28). Topic headings group related themes whose precise meanings can vary across models. Appendix Section B.4 describes the models and aggregation.

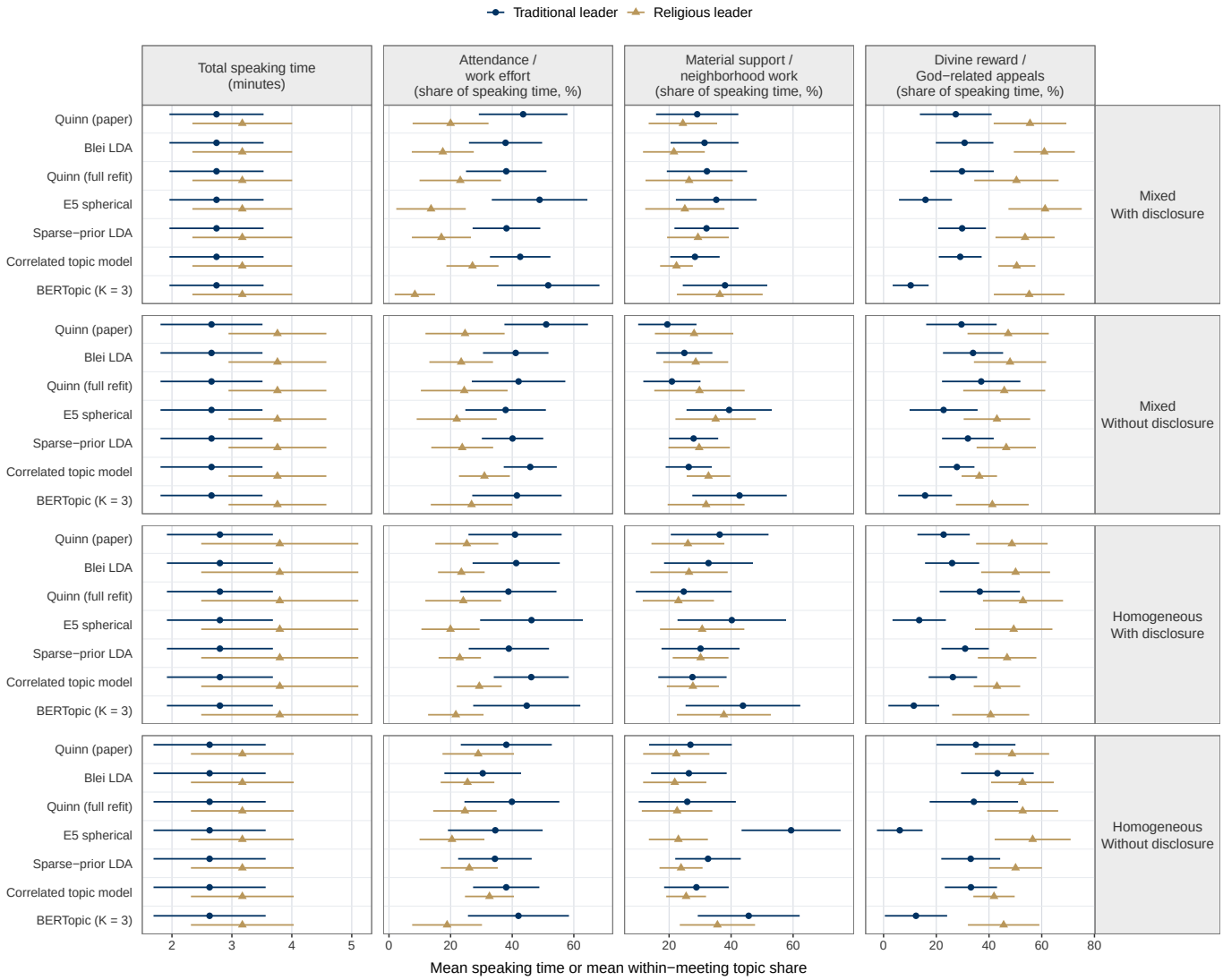


Figure B-3: LEADER SPEECH BY ENVIRONMENT ACROSS TEXT MODELS

Notes: Points show mean total speaking time and mean within-meeting topic shares, separately by group composition and disclosure. Blue circles denote traditional leaders and gold triangles religious leaders. All models use the same 206 speaking leader-meetings; total speaking time is identical across models. Lines are pointwise 95 percent confidence intervals clustered by neighborhood (28), conditional on the fitted text measures. Topic headings group related themes. Appendix Section B.4 describes the methods.

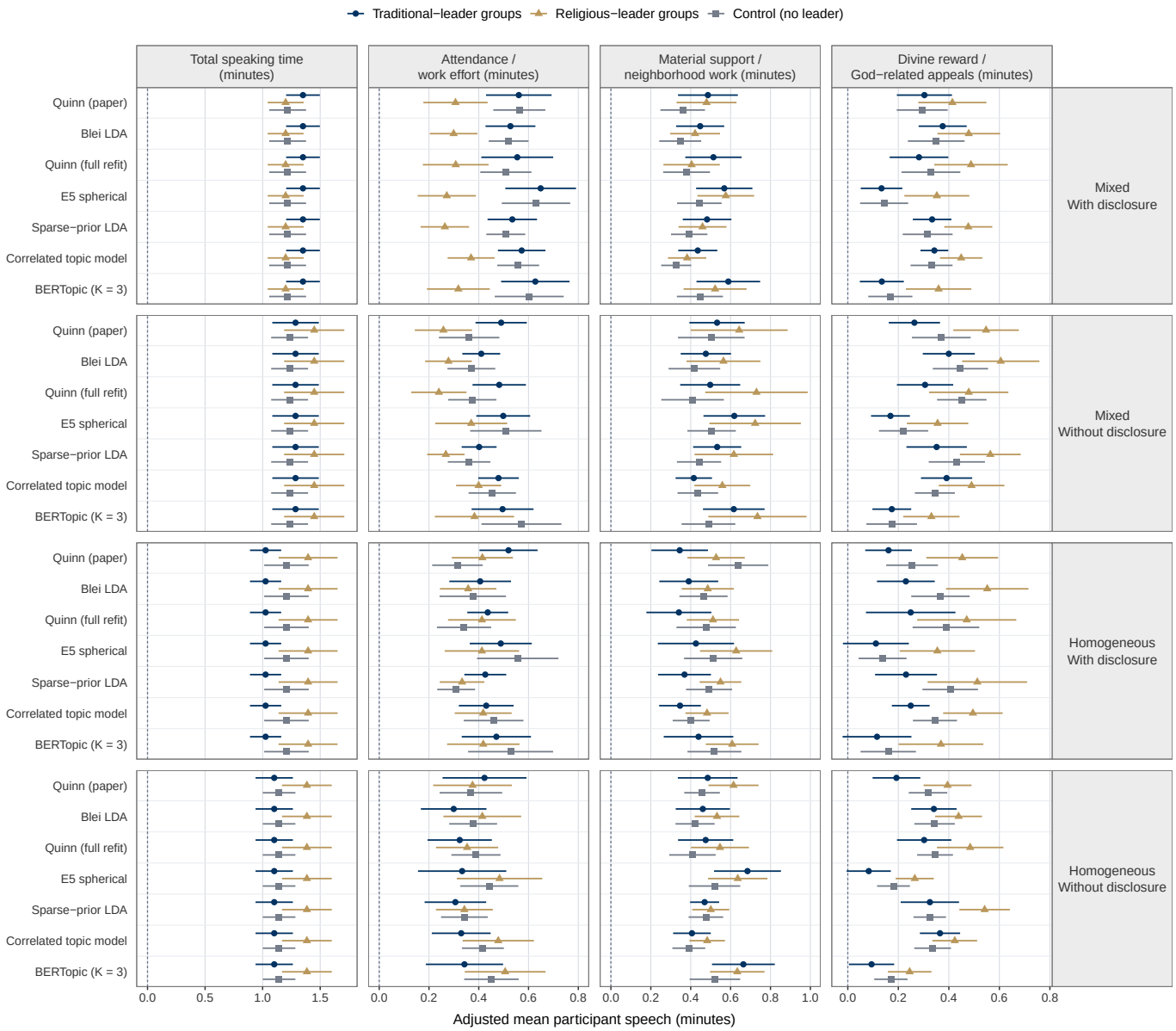


Figure B-4: PARTICIPANT SPEECH LEVELS ACROSS TEXT MODELS

Notes: Points show adjusted ordinary-participant speaking levels in traditional-leader groups (blue circles), religious-leader groups (gold triangles), and no-leader groups (gray squares). The sample contains 969 ordinary participants in 313 groups, including 62 silent participants. We include neighborhood-by-composition-by-disclosure fixed effects. Predictions for all three arms are averaged over the actual no-leader participants in each displayed environment. Lines are pointwise 95 percent confidence intervals using the full group-clustered covariance matrix, conditional on the fitted text measures. Total speaking time is identical across models. Appendix Section B.4 describes the methods.

C Scope of Authority

C.1 Spillover Exposure and Radius Robustness

Could residents respond to the treatment of another accountability group when its members live nearby? If they do, a comparison between our groups also picks up the treatments received by neighboring residents. We examine this possibility by measuring proximity to treated households from other groups and including that exposure in the main specification. We then vary how far we look for those households and compare the resulting estimates.

We can use this exercise to describe the association between exposure and participation, but proximity is not randomized and the coefficient does not identify an indirect treatment effect. Across radii we retain 1,040 participants in 336 assigned groups from 28 neighborhoods. The regression continues to use the same outcome and pretreatment participation control, design fixed effects, 11 LASSO-selected baseline controls, and clustering by group.

Exposure at the end of the wave. A resident is exposed when a study household from another accountability group lies within radius r. We measure this on the final treatment date of the wave, when all groups in that wave have been treated. Write d_{ij} for the distance between respondents i and j and G_i for respondent i's assigned accountability group. Exposure is

$$E_i(r) = \mathbf{1}\left\{\sum_{j \neq i} \mathbf{1}\{d_{ij} \leq r\}\mathbf{1}\{G_j \neq G_i\} > 0\right\}. \tag{4}$$

For j, we use the study households treated in i's wave. Members of i's accountability group do not contribute to exposure, regardless of their distance: they are part of the intervention assigned to i.

Figure C-1 illustrates the 25-meter definition in neighborhoods near the 20th, 50th, and 80th percentiles of exposure at the end of the wave. Some households appear inside a gold buffer but remain classified as unexposed because the nearby treated household belongs to the same assigned group.

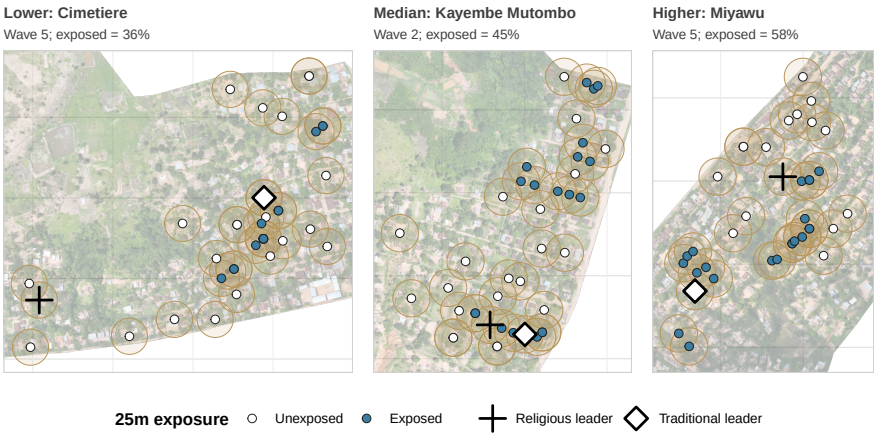


Figure C-1: TERMINAL SPILLOVER EXPOSURE IN THREE NEIGHBORHOODS

Notes: The three neighborhoods are those nearest the 20th, 50th, and 80th percentiles of the share exposed within 25 meters. All panels are measured on the final treatment date. Gold disks are 25-meter buffers around treated study households. Filled points are main-sample respondents with at least one treated out-of-group household within 25 meters; white points are unexposed. Crosses mark religious leaders and diamonds traditional leaders. The property registry TIFF appears as a translucent background. A respondent inside a buffer remains unexposed when the nearby household belongs to the respondent's own assigned group.

Radius choice and additive adjustment. At 25 meters, we have 483 exposed participants out of 1,040, or 46.4 percent, compared with 55.4 percent at 30 meters. We therefore use 25 meters: it comes closest to an equal split of the sample among the radii we tested, and we make the choice without participation

outcomes. Increasing the radius means more exposure and less unexposed comparison, reaching 79.4 percent exposed at 50 meters and 89.4 percent at 70 meters.

Figure C-2 adds $E_i(r)$ to the main specification at each radius while holding the sample and remaining specification fixed. It reports each of the eight leader effects separately by composition and disclosure. The effects change little: across 10–70 meters, each varies by at most 0.51 percentage points. Controlling for nearby households' treatment therefore does not account for the paper's main treatment effects.

Figure C-3 reports the exposure coefficient from the same regressions. At 25 meters, it is 2.08 percentage points (SE 1.99). The exposure measure is therefore not strongly associated with participation at the selected radius.

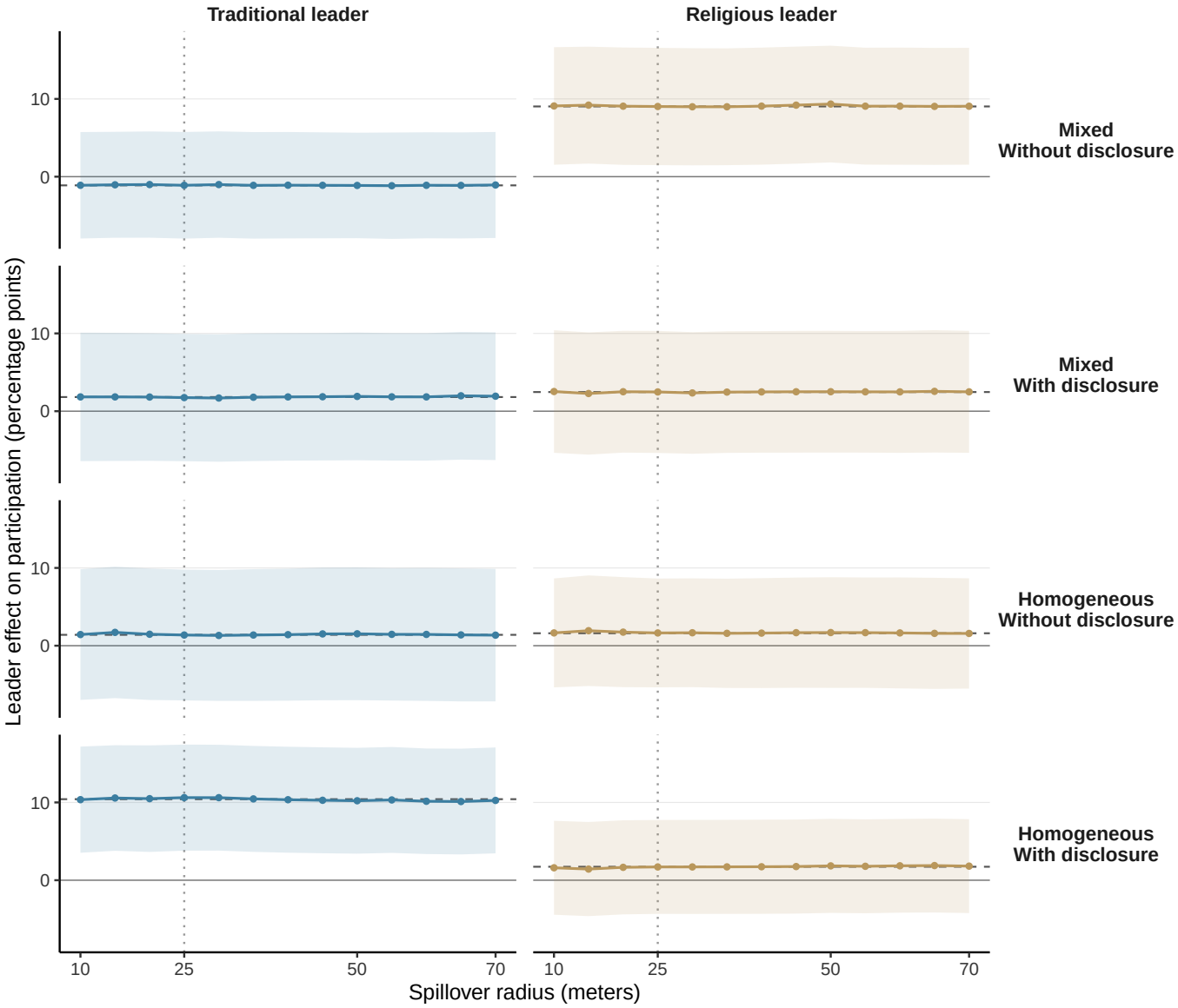


Figure C-2: LEADER EFFECTS ACROSS SPILLOVER RADII

Notes: We compare the leader effects as the radius used to measure spillover exposure increases. Each of the eight panels holds composition and disclosure fixed and shows the effect of a traditional or religious leader relative to no leader. The regression at each radius includes the indicator for exposure at the end of the wave, baseline participation, the same 11 LASSO-selected baseline controls, and neighborhood-by-composition-by-disclosure fixed effects. All regressions use the same 1,040 residents in 336 assigned groups and 28 neighborhoods. Traditional leader effects are blue; religious leader effects are gold. Shaded regions give pointwise 95 percent confidence intervals clustered by assigned group. Horizontal dashed lines provide the comparison without an exposure control, retaining the rest of this specification. The dotted vertical line in each panel marks 25 meters, and the vertical scale is the same throughout.

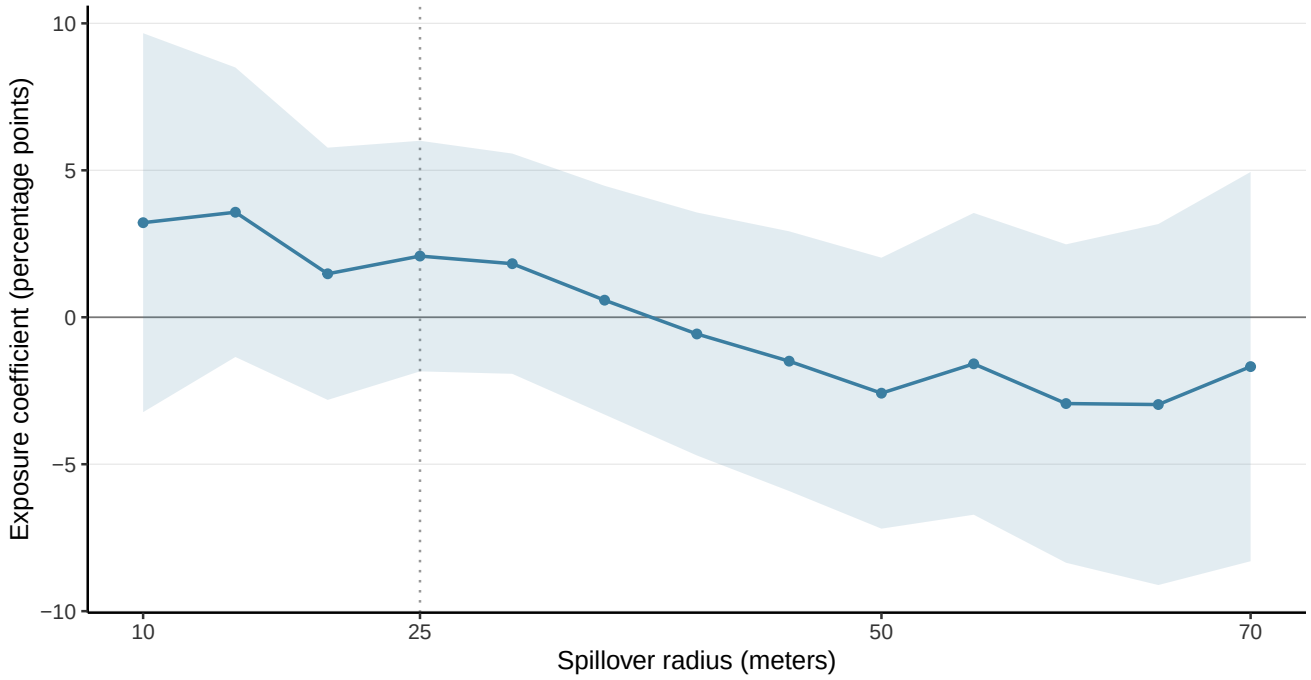


Figure C-3: SPILLOVER EXPOSURE COEFFICIENT ACROSS RADII

Notes: The regressions in Figure C-2 also estimate how participation varies with exposure to other groups. Here we plot that coefficient for the indicator of exposure within each radius at the end of the wave. All eight leader-by-composition-by-disclosure effects remain in the regression, along with baseline participation, the same 11 LASSO-selected baseline controls, and neighborhood-by-composition-by-disclosure fixed effects. We use the same 1,040 residents in 336 assigned groups across 28 neighborhoods. Shading gives pointwise 95 percent confidence intervals, clustering by assigned group; the dotted vertical line marks 25 meters. Residents' places of residence determine exposure. The coefficient therefore describes an association with participation, and the design does not identify an indirect treatment effect.

C.2 Baseline Control Selection

Which baseline characteristics improve prediction of participation? We choose them with single-equation outcome LASSO (Belloni et al., 2014), keeping baseline participation, leader-by-composition effects, design fixed effects, and spillover exposure in the regression throughout. Baseline participation and the design fixed effects remain in all participation analyses, with the preferred specification also retaining the 25-meter exposure measure and 11 selected controls.

Six controls are response-category indicators: chief performance in salongo (category 4), community cooperation (category 2), understanding others' emotions (category 3), government security performance (category 2 or higher), avenue-chief performance (category 2 or higher), and immediate ancestral rewards (category 2 or higher). The five controls outside these response-category indicators are education, gender, age, NGO provision of poor relief, and NGO provision of schools.

The same controls are used for attendance days, hours, effort, asinh hours and alternative observation windows, as well as for the participation robustness and heterogeneity specifications with LASSO controls. The selection is made once for primary participation. When we change the measure, we also use its matched pretreatment outcome; when we change the window, we retain the common pre-meeting participation baseline. Values that are missing are replaced with main-sample means. No missingness indicator was among those selected.

C.3 Main Effects, Alternative Outcomes, and Inference

Table C-1 – LEADER EFFECTS BY GROUP COMPOSITION

	Disclosure pooled		
	Baseline	+ Spillover	+ LASSO
Panel A. Leader effects relative to no leader			
Traditional leader			
Mixed	0.72	0.68	0.37
	(2.86)	(2.86)	(2.74)
Homogeneous	4.82*	4.88*	5.98**
	(2.91)	(2.90)	(2.78)
Religious leader			
Mixed	6.29**	6.28**	5.80**
	(2.94)	(2.95)	(2.80)
Homogeneous	0.83	0.83	1.66
	(2.52)	(2.52)	(2.37)
Panel B. Treatment-effect contrasts			
Mixed: T − R	-5.57*	-5.60*	-5.43*
	(2.98)	(3.00)	(2.96)
Homogeneous: T − R	3.98	4.05	4.32
	(2.94)	(2.92)	(2.72)
Δ_C: composition gradient	9.55**	9.65**	9.75**
	(4.20)	(4.20)	(4.06)
Panel C. No-leader reference means and diagnostics			
Mixed	58.33	58.33	58.33
	(1.99)	(1.99)	(1.79)
Homogeneous	56.79	56.79	56.79
	(1.76)	(1.76)	(1.73)
Wald: all no-leader cells equal	p = 0.562	p = 0.562	p = 0.536
Observations	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336
Neighborhoods	28	28	28
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes
Pretreatment participation control	Yes	Yes	Yes
25m spillover control	No	Yes	Yes
LASSO-selected baseline controls	No	No	Yes

Notes: Panel A reports leader effects relative to no leader; Panel B reports direct leader contrasts; Panel C reports no-leader reference means. Leader effects are pooled across disclosure; Δ_C measures the change in the traditional-minus-religious gap from mixed to homogeneous composition. The outcome is average participation on six eligible days after Meeting 1 through the Meeting 2 day, in percentage points. Every column controls for pretreatment participation and neighborhood-by-composition-by-disclosure fixed effects. The second adds the 25-meter spillover indicator; the third also adds 11 LASSO-selected baseline controls. Group-clustered standard errors appear in parentheses. Section 4.2 and Figure 4 discuss these estimates. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C-2 – LABOR CONTRIBUTION OUTCOMES BY GROUP COMPOSITION

	Levels				Asinh robustness
	Participation rate	Attendance days	Total hours	Total effort	Asinh total hours
Panel A. Leader effects relative to no leader					
Traditional leader					
Mixed	0.004	0.041	-0.308	-0.084	-0.007
	(0.027)	(0.166)	(0.628)	(0.552)	(0.113)
Homogeneous	0.060**	0.315*	1.378**	1.174**	0.235**
	(0.028)	(0.163)	(0.639)	(0.540)	(0.103)
Religious leader					
Mixed	0.058**	0.369**	0.864	1.047*	0.164
	(0.028)	(0.169)	(0.669)	(0.558)	(0.109)
Homogeneous	0.017	0.071	0.047	0.188	0.079
	(0.024)	(0.142)	(0.580)	(0.498)	(0.087)
Panel B. Treatment-effect contrasts					
Mixed: T − R	-0.054*	-0.328*	-1.173*	-1.131*	-0.171
	(0.030)	(0.179)	(0.688)	(0.592)	(0.118)
	[Holm p = 0.228]	[Holm p = 0.228]	[Holm p = 0.228]	[Holm p = 0.228]	[Holm p = 0.148]
Homogeneous: T − R	0.043	0.244	1.331**	0.986*	0.156
	(0.027)	(0.159)	(0.613)	(0.537)	(0.101)
	[Holm p = 0.226]	[Holm p = 0.226]	[Holm p = 0.122]	[Holm p = 0.202]	[Holm p = 0.123]
Δ_C	0.098**	0.572**	2.504***	2.116***	0.327**
	(0.041)	(0.242)	(0.928)	(0.806)	(0.156)
	[Holm p = 0.033]	[Holm p = 0.033]	[Holm p = 0.029]	[Holm p = 0.029]	[Holm p = 0.037]
Panel C. No-leader reference means and diagnostics					
Mixed	0.583	3.446	13.210	11.334	2.669
	(0.018)	(0.108)	(0.420)	(0.358)	(0.072)
Homogeneous	0.568	3.364	12.703	10.872	2.594
	(0.017)	(0.103)	(0.425)	(0.352)	(0.062)
Wald: homogeneous = mixed	p = 0.536	p = 0.585	p = 0.396	p = 0.358	p = 0.431
Observations	1,040	1,040	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336	336	336
Neighborhoods	28	28	28	28	28
Matched pretreatment outcome control	Yes	Yes	Yes	Yes	Yes
25m spillover control	Yes	Yes	Yes	Yes	Yes
Frozen participation LASSO controls	Yes	Yes	Yes	Yes	Yes
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes

Notes: The exposition of results in participation shares facilitates their interpretation, but underplays how much residents contributed. We therefore follow the same leader comparisons in labor days, hours and effort as well as the daily participation rate.
Panel A reports each leader-assignment effect relative to no leader. Panel B makes the direct traditional-minus-religious comparisons and reports Δ_C, how much larger the chief's advantage over the pastor is in homogeneous groups than in mixed groups. Panel C reports model-implied no-leader means and tests whether they are equal.
Residents who work once for a short time and those who work once for a long time have the same attendance count. Hours distinguish these contributions. We report the daily participation rate, count attendance in labor days, and sum hours and recorded effort over six days, strictly after Meeting 1 through and including the day of Meeting 2.
All specifications control for the resident's pretreatment measure, matched to the outcome in that column. The neighborhood-by-disclosure-by-composition fixed effects, 25-meter spillover indicator and 11 controls from the primary participation selection enter as well. The selection is still the primary participation selection, even for other measures and observation windows.
Meeting 1 does not contribute to these outcomes. A resident who did not attend has a zero, and that resident is also kept in the inverse-hyperbolic-sine hours regression; its sample is unchanged.
For a given contrast, the four outcomes in levels form one Holm family, while the single inverse-hyperbolic-sine hours outcome forms its own family. Brackets contain their adjusted p-values. The standard errors in parentheses allow correlation among residents assigned to the same group within a neighborhood. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C-3 – CASH AND IN-KIND CONTRIBUTIONS BY GROUP COMPOSITION

	Levels				Asinh robustness	
	Any cash contribution	Cash total	Any in-kind contribution	In-kind value	Asinh cash total	Asinh in-kind value
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	0.030*	14.971	-0.005	-53.318	0.221*	-0.067
	(0.017)	(27.229)	(0.032)	(101.388)	(0.125)	(0.248)
Homogeneous	0.013	-1.271	0.007	-1131.884*	0.086	-0.059
	(0.014)	(14.186)	(0.040)	(677.600)	(0.097)	(0.316)
Religious leader						
Mixed	0.006	-7.053	0.053	-6.317	0.046	0.419
	(0.016)	(20.885)	(0.034)	(99.060)	(0.115)	(0.266)
Homogeneous	0.001	-13.396	0.001	-957.190	-0.005	-0.051
	(0.013)	(13.200)	(0.041)	(684.773)	(0.094)	(0.329)
Panel B. Treatment-effect contrasts						
Mixed: T − R	0.024	22.024	-0.058	-47.001	0.175	-0.486*
	(0.020)	(27.104)	(0.037)	(112.707)	(0.147)	(0.288)
	[Holm p = 0.702]	[Holm p = 0.834]	[Holm p = 0.481]	[Holm p = 0.834]	[Holm p = 0.235]	[Holm p = 0.185]
Homogeneous: T − R	0.012	12.125	0.007	-174.694	0.090	-0.009
	(0.013)	(11.360)	(0.043)	(486.568)	(0.091)	(0.339)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.641]	[Holm p = 0.980]
Δ_C	-0.012	-9.898	0.064	-127.693	-0.084	0.477
	(0.024)	(29.392)	(0.057)	(499.382)	(0.173)	(0.445)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.626]	[Holm p = 0.567]
Panel C. No-leader reference means and diagnostics						
Mixed	0.022	38.315	0.266	543.478	0.161	2.122
	(0.009)	(15.616)	(0.021)	(61.742)	(0.063)	(0.160)
Homogeneous	0.016	21.739	0.288	1619.022	0.122	2.327
	(0.010)	(11.417)	(0.026)	(611.892)	(0.069)	(0.211)
Wald: homogeneous = mixed	p = 0.679	p = 0.392	p = 0.513	p = 0.080	p = 0.682	p = 0.437
Observations	1,040	1,040	1,040	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336	336	336	336
Neighborhoods	28	28	28	28	28	28
Matched pretreatment outcome control	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Labor is one form of contribution. Here we examine cash and goods.
Each leader assignment is compared with no leader in Panel A. Panel B reports the direct traditional-minus-religious contrasts and Δ_C, the traditional composition gradient minus the religious composition gradient. Panel C gives the model-implied means without a leader and tests whether they are equal. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Residents can give cash and goods as well as work. We record whether they gave any of each and the value of each in Congolese francs (CDF) over six days, strictly after Meeting 1 through and including the Meeting 2 day.
The resident's pretreatment outcome is held constant in each model, matched to the outcome being studied, along with neighborhood × disclosure × composition fixed effects. We cluster standard errors by group within neighborhood and put them in parentheses. For the same contrast, brackets report Holm-adjusted p-values across the four level outcomes, with a separate family of the two asinh outcomes.
Combined material contributions add cash and the recorded value of in-kind giving. Before estimating the models, we check the source identity combined material value = cash + in-kind. Residents with zero contributions stay in both inverse-hyperbolic-sine specifications. Meeting 1 itself is excluded.

Table C-4 – RANDOMIZATION INFERENCE AND ALTERNATIVE CLUSTERING BY GROUP COMPOSITION

	Randomization inference			Alternative clustering		
	Estimate	Clustered SE	RI p	Group	Leader cluster	Neighborhood
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	0.37	(2.74)	0.907	0.37	0.37	0.37
				(2.74)	(2.70)	(3.20)
Homogeneous	5.98**	(2.78)	0.049	5.98**	5.98**	5.98*
				(2.78)	(2.74)	(3.18)
Religious leader						
Mixed	5.80**	(2.80)	0.057	5.80**	5.80**	5.80*
				(2.80)	(2.74)	(3.28)
Homogeneous	1.66	(2.37)	0.587	1.66	1.66	1.66
				(2.37)	(2.48)	(2.65)
Panel B. Treatment-effect contrasts						
Mixed: T − R	-5.43*	(2.96)	0.090	-5.43*	-5.43*	-5.43*
				(2.96)	(2.78)	(3.09)
Homogeneous: T − R	4.32	(2.72)	0.173	4.32	4.32	4.32
				(2.72)	(3.02)	(3.78)
Δ_C: composition gradient	9.75**	(4.06)	0.030	9.75**	9.75**	9.75*
				(4.06)	(4.08)	(4.92)
Panel C. No-leader reference means and diagnostics						
Mixed	58.33	(1.79)		58.33	58.33	58.33
				(1.79)	(1.87)	(1.89)
Homogeneous	56.79	(1.73)		56.79	56.79	56.79
				(1.73)	(1.50)	(1.52)
Wald: all no-leader cells equal	p = 0.536			p = 0.536	p = 0.498	p = 0.503
Observations	1,040			1,040	1,040	1,040
Clusters				336	84	28
Neighborhoods	28			28	28	28
Neighborhood × disclosure × composition FEs				Yes	Yes	Yes
Pretreatment participation control	Yes			Yes	Yes	Yes
25m spillover control	Yes			Yes	Yes	Yes
LASSO-selected baseline controls	11			11	11	11
Group-neighborhood clusters	336					
NQC fixed effects	Yes					
Formation-stage permutations	5,000					

Notes: The observation is a resident, but the experiment formed teams and assigned the same leader to several teams. We examine inference from the formation stage and from allowing correlation across these teams. The outcome throughout is average salongo participation after the first meeting, in percentage points, using the same 1,040 residents. Six-day pretreatment participation, the 25-meter spillover indicator and the 11 baseline controls selected by LASSO for primary participation enter every regression, along with neighborhood × disclosure × composition fixed effects. The main specification and comparisons are unchanged. Panel A gives each leader-assignment effect relative to no leader in the same fixed-effect cell. Panel B gives the main table's leader contrasts; Δ_C is the difference between the Traditional and Religious homogeneous-minus-mixed gradients. Panel C gives model-implied no-leader reference means and clustered delta-method tests of equality among all displayed means.

Randomization inference (RI) repeats the group-formation procedure across 5,000 assignments while holding the realized sample fixed. That means retaining each wave's group sizes, random filling order, composition and fallback rules, gender × wealth targets, territorial diversity in mixed groups and treatment coding, without resampling residents or repeating the original best-balance selection. It also means using the same 11 selected controls rather than selecting them again; their coefficients are reestimated on each replay. We report two-sided RI p-values and standard errors clustered one-way by group–neighborhood. Holding exposure fixed requires a qualification. This regressor excludes households belonging to the same group, and its value could change when that group changes. Our draws need not preserve observed exposure, yet we use the observed indicator in their regressions. This conditional group-formation replay gives diagnostics, not exact randomization tests conditional on exposure or inference from the full rerandomization design. Stars use clustered asymptotic inference: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Finally, we allow errors to be correlated across groups assigned to the same leader. These alternative-clustering columns use 336 experimental group–neighborhood clusters, 84 neighborhood-wave × leader-arm clusters containing four traditional, religious or no-leader groups each, or 28 neighborhood-wave assignments for neighborhood clustering. Only this level changes: sample, outcome, all controls and fixed effects, leader effects, contrasts and no-leader comparisons remain the same. In Panels B and C we use the clustered delta method. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4 Alternative Explanations

C.4.1 Endogenous Leader Behavior

Table C-5 – ARRIVING TOGETHER WITH THE ASSIGNED LEADER

	Chain 30 sec. (1)	Chain 1 min. (main) (2)	Direct gap 1 min. (3)	Direct gap 2 min. (4)
I. Group composition (pooled across disclosure)				
Panel A. Adjusted levels				
Traditional leader				
Mixed	5.67	8.22	3.87	6.74
	(1.41)	(1.58)	(0.75)	(1.10)
Homogeneous	6.98	10.52	4.96	7.84
	(1.63)	(2.27)	(1.21)	(1.78)
Religious leader				
Mixed	4.74	5.80	2.91	5.64
	(1.04)	(1.09)	(0.62)	(0.97)
Homogeneous	3.97	5.24	2.75	4.09
	(1.35)	(1.57)	(0.65)	(1.09)
Panel B. Comparisons				
Traditional leader				
Homogeneous minus mixed	1.31	2.29	1.09	1.10
	(1.54)	(1.65)	(1.27)	(1.65)
Religious leader				
Homogeneous minus mixed	-0.77	-0.56	-0.17	-1.55
	(1.15)	(1.37)	(0.72)	(1.20)
II. Group composition by disclosure				
Panel C. Adjusted levels				
Traditional leader				
Mixed, without disclosure	7.16	9.41	4.86	8.43
	(2.04)	(2.25)	(1.27)	(1.61)
Mixed, with disclosure	4.18	7.03	2.87	5.05
	(1.23)	(1.55)	(0.81)	(1.26)
Homogeneous, without disclosure	5.63	9.35	3.85	6.99
	(1.60)	(2.29)	(1.23)	(1.90)
Homogeneous, with disclosure	8.33	11.68	6.06	8.69
	(2.02)	(2.63)	(1.51)	(2.18)
Religious leader				
Mixed, without disclosure	4.96	6.73	2.53	5.90
	(1.18)	(1.36)	(0.69)	(1.21)
Mixed, with disclosure	4.51	4.87	3.30	5.38
	(1.27)	(1.22)	(0.96)	(1.19)
Homogeneous, without disclosure	3.34	4.71	1.97	3.63
	(1.30)	(1.57)	(0.59)	(1.17)
Homogeneous, with disclosure	4.60	5.78	3.53	4.55
	(1.60)	(1.77)	(1.00)	(1.22)
Panel D. Comparisons				
Traditional leader				
Disclosure effect, mixed	-2.98	-2.38	-1.98	-3.37*
	(1.85)	(2.22)	(1.51)	(1.88)
Disclosure effect, homogeneous	2.69	2.33	2.21*	1.70
	(1.60)	(1.89)	(1.29)	(2.01)
Religious leader				
Disclosure effect, mixed	-0.45	-1.86	0.76	-0.52
	(1.30)	(1.39)	(1.12)	(1.40)
Disclosure effect, homogeneous	1.26	1.06	1.55	0.91
	(1.10)	(1.12)	(0.99)	(0.97)
Respondents, chiefs	336	336	336	336
Respondents, pastors	336	336	336	336
Neighborhoods per leader type	28	28	28	28

Notes: For the main-text peer comparison, each day's number of co-arriving ordinary groupmates is divided by the fixed number of ordinary groupmates, including nonattenders. The one-minute measure weights homogeneous and mixed groups equally and adjusts each partner measure for its own pretreatment value. In the joint model, the effect of composition and the slope on the pretreatment measure can each differ across leader types and partner measures. Standard errors cluster by neighborhood. The pooled chief–ordinary-groupmate difference has $p = 0.309$. Stars on differences use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Each column is the percentage of six post-meeting days meeting its co-arrival definition. Each measure is adjusted for its own value over the six pretreatment days in the regression. Both the resident and the partner must supply labor for the day to count. An absence is recorded as zero. Direct gap compares the absolute time between the resident's and leader's check-ins, with no transitive connections through other arrivals. Chains connect consecutive check-ins among all recorded attendees whenever the gap is no more than the indicated interval; the one-minute chain is the main measure. We estimate chiefs and pastors separately. Pooled regressions compare homogeneous with mixed groups across disclosure conditions; adding disclosure and its interaction with composition then compares disclosure with no disclosure in each composition. Predicted levels use the common baseline mean. Levels are percentages, differences percentage points, and parentheses contain neighborhood-clustered standard errors.

Table C-6 – INTERACTIONS WITH THE ASSIGNED LEADER

	Any interaction (1)	Discussed salongo (2)	Asked to participate (3)	Social consequences (4)	Divine consequences (5)
I. Group composition (pooled across disclosure)					
Panel A. Adjusted levels					
Traditional leader					
Mixed	50.30	39.61	13.78	2.11	3.33
	(3.93)	(5.01)	(3.38)	(1.13)	(1.90)
Homogeneous	46.00	35.33	12.65	2.00	2.67
	(4.86)	(4.51)	(2.83)	(1.12)	(1.23)
Religious leader					
Mixed	40.68	36.72	10.95	1.33	3.16
	(3.64)	(3.81)	(2.80)	(0.97)	(1.24)
Homogeneous	44.30	34.76	14.08	2.74	3.94
	(4.24)	(4.51)	(3.24)	(1.63)	(2.01)
Panel B. Comparisons					
Traditional leader					
Homogeneous minus mixed	-4.30	-4.28	-1.13	-0.11	-0.66
	(4.09)	(4.19)	(3.90)	(1.70)	(2.02)
Religious leader					
Homogeneous minus mixed	3.62	-1.96	3.13	1.41	0.78
	(4.28)	(5.07)	(2.74)	(1.37)	(2.29)
II. Group composition by disclosure					
Panel C. Adjusted levels					
Traditional leader					
Mixed, without disclosure	49.89	37.18	7.05	1.47	1.23
	(4.11)	(6.02)	(2.96)	(1.37)	(1.41)
Mixed, with disclosure	50.70	42.25	21.12	2.82	5.63
	(6.68)	(6.95)	(5.84)	(1.95)	(3.88)
Homogeneous, without disclosure	50.07	36.11	11.52	1.27	1.42
	(5.91)	(5.30)	(3.41)	(1.48)	(1.35)
Homogeneous, with disclosure	42.24	34.62	13.70	2.68	3.81
	(6.55)	(6.31)	(3.81)	(1.81)	(2.04)
Religious leader					
Mixed, without disclosure	45.39	42.74	14.69	2.66	5.36
	(5.82)	(6.03)	(4.39)	(1.87)	(2.54)
Mixed, with disclosure	36.19	30.98	7.40	0.07	1.07
	(5.43)	(5.34)	(2.79)	(0.15)	(1.16)
Homogeneous, without disclosure	43.10	32.17	14.33	4.12	3.67
	(5.58)	(5.05)	(3.62)	(2.32)	(2.19)
Homogeneous, with disclosure	45.51	37.38	13.81	1.33	4.22
	(6.58)	(6.51)	(4.33)	(1.32)	(2.29)
Panel D. Comparisons					
Traditional leader					
Disclosure effect, mixed	0.81	5.07	14.06**	1.35	4.41
	(7.67)	(8.24)	(6.15)	(2.46)	(4.30)
Disclosure effect, homogeneous	-7.83	-1.50	2.18	1.41	2.39
	(7.90)	(7.42)	(4.48)	(2.45)	(2.44)
Religious leader					
Disclosure effect, mixed	-9.21	-11.77	-7.29	-2.59	-4.28
	(8.61)	(8.49)	(4.71)	(1.76)	(3.00)
Disclosure effect, homogeneous	2.41	5.20	-0.52	-2.79	0.56
	(8.74)	(7.37)	(4.60)	(1.84)	(1.97)
Respondents, chiefs	299	299	299	299	299
Respondents, pastors	302	302	302	302	302
Neighborhoods per leader type	28	28	28	28	28

Notes: Ordinary respondents report whether they interacted with their assigned leader, discussed salongo, were asked to participate, or heard the leader mention social or divine consequences. We exclude reports about unknown persons. Each regression controls for the respondent's pretreatment participation.
The comparison between homogeneous and mixed groups pools disclosure, in a regression for each leader type. We also compare disclosure within each composition, adding disclosure and its interaction with composition. All predicted levels set baseline participation to the common mean. Levels are percentages and differences percentage points. Parentheses give neighborhood-clustered standard errors. Stars on differences use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4.2 Majority Membership

Table C-7 – SHARED TERRITORIAL ORIGIN WITH THE NEIGHBORHOOD CHIEF

	Baseline participation control			+ Spillover and LASSO		
	Different territory	Same territory	Difference	Different territory	Same territory	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-1.05	6.09	7.13	-1.28	5.26	6.54
	(3.35)	(7.31)	(8.56)	(3.35)	(6.65)	(8.00)
Homogeneous, pooled		4.85*			6.01**	
		(2.91)			(2.78)	
Religious leader						
Mixed	3.41	14.69**	11.27	3.51	12.47*	8.96
	(3.54)	(7.26)	(8.66)	(3.35)	(6.68)	(7.93)
Homogeneous, pooled		0.85			1.68	
		(2.53)			(2.38)	
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed	58.92	56.35	-2.57	58.92	56.35	-2.57
	(2.23)	(5.17)	(5.92)	(1.98)	(4.96)	(5.63)
Homogeneous, pooled		56.79			56.79	
		(1.76)			(1.73)	
Observations		1,040			1,040	
Group clusters		336			336	
Groups containing respondents in each category (first / second)						
Chief, mixed		56 / 34			56 / 34	
Chief homogeneous groups (pooled)		56			56	
Pastor, mixed		55 / 36			55 / 36	
Pastor homogeneous groups (pooled)		56			56	

Notes: In the specification with spillover and LASSO controls, we compare the pastor effect among the chief's coethnics in mixed groups with the pooled pastor effect in homogeneous groups. The former minus the latter is 10.79 percentage points (SE 7.09; $p = 0.129$), calculated using the full coefficient covariance matrix.

The table compares leader effects among residents who share the neighborhood chief's territorial origin and those who do not. We use the neighborhood chief as the reference person in all three arms, including when a pastor or no leader was assigned. All residents of a homogeneous group share the chief's territorial origin by design, leaving their effects pooled. Only in mixed groups do we estimate differences between coethnic categories; the difference column takes the second category's effect minus the first.

Panel B's reference is based on actual no-leader residents. We predict for them using their own observed covariates and fixed effects, averaging within composition and, for mixed groups, territorial-origin category. Both blocks include baseline participation and neighborhood-by-composition-by-disclosure fixed effects. The second adds spillover and the 11 LASSO-selected baseline controls described, together with their missing-value treatment, in Appendix Section C.5. Leader effects, interactions with the characteristic and any separate category levels pool disclosure within composition. Standard errors in parentheses cluster by group conditional on the categories. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 4.3 discusses the estimates.

C.4.3 Pastors' Territorial Origins

Table C-8 – LEADER EFFECTS UNDER THE RELIGIOUS-LEADER TERRITORY RESTRICTION

	Leader by composition		
	Baseline	+ Spillover	+ LASSO
Panel A. Leader effects relative to no leader			
Traditional leader			
Mixed	0.65	0.60	0.24
	(2.71)	(2.70)	(2.62)
Homogeneous	4.79*	4.89*	6.04**
	(2.88)	(2.87)	(2.76)
Religious leader			
Mixed	10.05***	10.12***	8.46**
	(3.52)	(3.51)	(3.45)
Homogeneous	-1.43	-1.50	-0.84
	(3.33)	(3.31)	(3.03)
Panel B. Treatment-effect contrasts			
Mixed: T − R	-9.40***	-9.52***	-8.22**
	(3.43)	(3.44)	(3.45)
Homogeneous: T − R	6.22*	6.38*	6.88**
	(3.74)	(3.71)	(3.38)
Δ_C: composition gradient	15.62***	15.91***	15.10***
	(5.08)	(5.07)	(4.84)
Panel C. No-leader reference means and diagnostics			
Mixed	58.33	58.33	58.33
	(1.92)	(1.90)	(1.77)
Homogeneous	56.79	56.79	56.79
	(1.76)	(1.76)	(1.73)
Wald: homogeneous = mixed	p = 0.555	p = 0.553	p = 0.534
Observations	908	908	908
Group-neighborhood clusters	292	292	292
Neighborhoods	28	28	28
Design fixed effects	Yes	Yes	Yes
Pretreatment participation control	Yes	Yes	Yes
25m spillover control	No	Yes	Yes
LASSO-selected baseline controls	No	No	Yes

Notes: Could pastors' limited influence in homogeneous groups reflect a lack of ethnic ties? We repeat the pooled scope specification after excluding 132 residents assigned a pastor in the 11 neighborhoods where the pastor and chief originate from different territories. Residents assigned a chief or no leader all remain. Panel A gives leader effects pooling disclosure, Panel B compares chief and pastor effects, and Panel C gives predicted no-leader means. The first column adjusts for participation before treatment and neighborhood-by-composition-by-disclosure fixed effects. Moving to the second column adds nearby spillover exposure; moving to the third adds the 11 baseline controls selected by LASSO, retaining the preceding adjustments. Parentheses give standard errors clustered by group. See Section 4.2 for the territorial comparison. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4.4 Leadership by Example

Table C-9 – LEADERSHIP BY EXAMPLE: LEADER PRETREATMENT PARTICIPATION

	Group-clustered SE			Neighborhood-clustered SE		
	Participation at/below leader-type median	Participation above leader-type median	Difference	Participation at/below leader-type median	Participation above leader-type median	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	0.91	0.51	-0.40	0.91	0.51	-0.40
	(3.59)	(4.00)	(5.05)	(4.26)	(5.50)	(6.50)
Homogeneous	7.96**	2.12	-5.85	7.96	2.12	-5.85
	(3.93)	(3.84)	(5.21)	(5.79)	(3.78)	(6.99)
Religious leader						
Mixed	4.91	8.09**	3.18	4.91	8.09*	3.18
	(3.84)	(4.00)	(5.21)	(4.75)	(4.68)	(6.16)
Homogeneous	3.65	-2.87	-6.52	3.65	-2.87	-6.52
	(3.40)	(3.35)	(4.53)	(4.19)	(4.45)	(6.23)
Panel B. Implied no-leader counterfactuals						
No leader, mixed		58.33			58.33	
		(1.99)			(2.21)	
No leader, homogeneous		56.79			56.79	
		(1.79)			(1.70)	
Observations			1,040			
Group clusters			336			

Notes: The table compares effects of leaders who worked more before the meetings with those who worked less. We use a separate median for chiefs and pastors, place ties in the lower category, and report the effect above the median minus the effect at or below it. The rate covers the common week before any meeting: four to six observed eligible days, excluding Sunday. Medians are 0.80 for chiefs and 0.83 for pastors.
Since a no-leader team has no such rate, we let the participation category enter only through interactions with leader assignment and composition and include no standalone category term. Panel B gives both categories one pooled no-leader reference in each composition and block. It averages predictions for actual no-leader residents, retaining their observed covariates and fixed effects. All 1,040 participants enter one regression pooling disclosure within composition, with baseline participation and neighborhood-by-composition-by-disclosure fixed effects.
The two blocks fit the same regression and differ only in clustering, using 336 groups or 28 neighborhoods. The latter lets outcomes be dependent across groups assigned the same leader. Section 4.2 discusses the interpretation. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.4.5 Pretreatment Participation and Baseline Adjustment

Table C-10 – PRETREATMENT DYNAMICS: POOLED ACROSS DISCLOSURE

	All six pretreatment days			Excluding day 1		
Panel A. Linear trends						
	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value
Traditional leader						
Mixed	-0.50	(0.92)	0.587	-0.60	(1.06)	0.576
Homogeneous	-1.48	(0.86)	0.087	-0.87	(1.15)	0.453
Religious leader						
Mixed	1.10	(0.84)	0.190	1.01	(1.03)	0.325
Homogeneous	-2.21	(0.90)	0.014	-1.76	(1.20)	0.145
Panel B. Early versus late pretreatment						
	Estimate	Standard error	*p*-value	Estimate	Standard error	*p*-value
Traditional leader						
Mixed	-0.33	(2.93)	0.911	0.06	(2.97)	0.985
Homogeneous	-4.23	(2.99)	0.158	-2.40	(3.41)	0.483
Religious leader						
Mixed	2.05	(2.69)	0.446	1.05	(2.79)	0.707
Homogeneous	-4.82	(3.19)	0.131	-2.50	(3.57)	0.485
Panel C. Saturated day-specific paths						
	F	Restrictions	*p*-value	*F*	Restrictions	*p*-value
Traditional leader						
Mixed	0.97	5	0.435	1.19	4	0.316
Homogeneous	1.66	5	0.145	1.15	4	0.335
Religious leader						
Mixed	1.17	5	0.323	1.26	4	0.286
Homogeneous	2.75	5	0.019	1.97	4	0.099
Respondents		1,040			1,040	
Group–neighborhood clusters		336			336	
Neighborhoods		28			28	

Notes: Panels A and B report percentage-point estimates, clustered standard errors, and *p*-values for each of the four leader-by-composition paths, pooling disclosure conditions. Panel A estimates the change in the leader-minus-no-leader gap per campaign day. Panel B compares days 1–3 with days 4–6 in the full window and days 2–3 with days 4–6 after excluding day 1. Panel C jointly tests the day-specific coefficients for each path, taking campaign day 6 as the reference. We report the F statistic, number of restrictions, and p-value for this test. All regressions include respondent fixed effects and neighborhood × information × composition × campaign-day fixed effects. Standard errors are clustered one-way by assigned group within neighborhood.

Table C-11 – BASELINE IMBALANCE AND TREATMENT-SPECIFIC BASELINE ADJUSTMENT

	Leader by composition				Environment specific		
	Baseline adjustment	+ Spillover	+ LASSO		Baseline adjustment	+ Spillover	+ LASSO
Panel A. Leader effects relative to no leader							
Traditional leader							
Mixed	0.73	0.68	0.23	Without disclosure	-0.20	-0.23	-1.34
	(2.82)	(2.83)	(2.68)		(3.47)	(3.47)	(3.40)
				With disclosure	1.56	1.49	1.55
					(4.40)	(4.39)	(4.03)
Homogeneous	4.13	4.20	5.40**	Without disclosure	-0.50	-0.53	1.12
	(2.83)	(2.83)	(2.73)		(4.50)	(4.49)	(4.31)
				With disclosure	8.87***	9.06***	9.79***
					(3.26)	(3.27)	(3.25)
Religious leader							
Mixed	6.01**	6.00**	5.38*	Without disclosure	8.74**	8.72**	8.43**
	(2.93)	(2.94)	(2.81)		(3.98)	(3.99)	(3.85)
				With disclosure	3.29	3.28	2.20
					(4.29)	(4.32)	(4.00)
Homogeneous	1.10	1.10	1.78	Without disclosure	1.39	1.39	1.38
	(2.57)	(2.57)	(2.42)		(3.93)	(3.92)	(3.67)
				With disclosure	0.84	0.83	2.23
					(3.26)	(3.26)	(3.07)
Panel B. Treatment-effect contrasts							
Mixed: T − R	-5.29*	-5.32*	-5.15*	Mixed, with: T − R	-1.73	-1.79	-0.66
	(2.94)	(2.96)	(2.89)		(4.36)	(4.38)	(4.14)
				Mixed, without: T − R	-8.95**	-8.95**	-9.77**
					(3.88)	(3.90)	(4.02)
Homogeneous: T − R	3.03	3.11	3.62	Homogeneous, with: T − R	8.02**	8.23**	7.57**
	(2.85)	(2.85)	(2.68)		(3.73)	(3.72)	(3.46)
				Homogeneous, without: T − R	-1.89	-1.92	-0.26
					(4.18)	(4.16)	(3.98)
Δ_C: composition gradient	8.32**	8.43**	8.77**	Δ_E: H, with − M, without	16.97***	17.18***	17.34***
	(4.10)	(4.11)	(3.98)		(5.39)	(5.42)	(5.31)
Panel C. No-leader reference means and diagnostics							
Mixed	58.33	58.33	58.33	Without disclosure	58.88	58.88	58.88
	(1.98)	(1.97)	(1.78)		(2.45)	(2.44)	(2.23)
				With disclosure	57.79	57.79	57.79
					(3.11)	(3.11)	(2.80)
Homogeneous	56.79	56.79	56.79	Without disclosure	57.79	57.79	57.79
	(1.73)	(1.73)	(1.72)		(2.91)	(2.91)	(2.77)
				With disclosure	55.80	55.80	55.80
					(1.84)	(1.84)	(2.01)
Wald: all no-leader cells equal	p = 0.557	p = 0.557	p = 0.534		p = 0.770	p = 0.770	p = 0.773
Panel D. Joint tests							
All composition coefficients zero	0.135	0.133	0.084	All environment coefficients zero	0.062	0.059	0.038
Observations	1,040	1,040	1,040		1,040	1,040	1,040
Group–neighborhood clusters	336	336	336		336	336	336
Neighborhoods	28	28	28		28	28	28
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes		Yes	Yes	Yes
Assignment-cell baseline slopes	Yes	Yes	Yes		Yes	Yes	Yes
25m spillover control	No	Yes	Yes		No	Yes	Yes
LASSO-selected baseline controls	No	No	Yes		No	No	Yes

Notes: Each observation is a respondent's mean salongo participation after the first meeting, in percentage points, within the assigned treatment cell. We allow the relationship between pretreatment and subsequent participation to differ across all 12 assignment cells. Every specification includes neighborhood × information × composition fixed effects, and pretreatment participation is centered at the full-sample mean. Panels A–D report, respectively, leader effects, leader contrasts, model-implied no-leader means and equality tests, and joint tests of the displayed leader effects. $\Delta_C = (\tau^T_H - \tau^T_M) - (\tau^R_H - \tau^R_M)$ and $\Delta_E = (\tau^T_{HI} - \tau^T_{MN}) - (\tau^R_{HI} - \tau^R_{MN})$. Spillover columns add 25m exposure; LASSO columns additionally include 11 LASSO-selected baseline controls. We use the delta method for inference, with standard errors clustered by group within neighborhood. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

C.5 Resident and Group Characteristics

The characteristics in Tables C-12–C-16 and C-19–C-20 allow us to examine who responds to leaders. We estimate Equation (2), pooling disclosure within composition for both leader effects and heterogeneity. For characteristics measured in every arm, the regression allows their association with participation to vary by composition. Residents assigned no leader cannot assess an assigned leader. We therefore use only interactions with leader assignment for those assessments. For shared territory with the chief, we estimate heterogeneity only in mixed groups, since every resident of a homogeneous group shares the chief's territory.

At baseline, residents placed six identities in a forced ranking. Own-territory identification receives 1 when it is least important and 6 when it is most important. The group score averages those of the three ordinary members where available, including each respondent's own score when observed and excluding the leader. Of the groups, 330 have all three scores and six have two; the sample still has all 1,040 respondents. We call identification higher only above the group median of 10/3, placing ties in the lower category. We neither control for individual identification separately nor include separate interactions for it. The comparison describes differences across groups with different identification and cannot isolate a causal peer effect.

Table C-13 further divides each group-identification category in mixed groups by whether residents share the neighborhood chief's territory. The regression interacts both characteristics with each leader type and includes the corresponding no-leader category terms; homogeneous-group effects retain the group-identification split. No-leader levels average fitted predictions over actual no-leader respondents within each joint identity and coethnicity category.

For both authority and opinion, the estimates include 1,037 observations in 335 groups. Using meeting membership records, we match each rating with the assigned leader. Missing ratings lead to the loss of three residents in one group, where no mean rating is available. For another 33 residents we can use the mean observed among the group's ordinary members, and we condition the standard errors on the categories after this replacement. Ratings of an assigned leader are undefined in no-leader groups. We average fitted predictions for those actual no-leader residents at their observed covariates and fixed effects within composition, giving both assessment categories the same pooled no-leader reference. The higher-versus-lower comparison can reflect pre-existing willingness to participate in addition to a different response to the leader. For ethnic identification at group level, the reference averages fitted predictions for actual no-leader respondents in the observed composition and identity category, at their own covariates and fixed effects. For shared territory with the chief, reference levels average fitted predictions over actual no-leader respondents within composition and, in mixed groups, territorial-origin category, retaining their observed covariates and fixed effects. For the other characteristics observed in every arm, we predict no-leader levels for the same set of chief- and pastor-arm residents within each composition and category.

Following Allport and Ross (1967) and the revised measurement of Gorsuch and McPherson (1989), we distinguish intrinsic from extrinsic religiosity. The survey adapts nine items. Intrinsic religiosity averages items 1, 3 reversed, 4, 5, and 9 reversed. Extrinsic religiosity averages items 2, 6, 7, and 8: one personal-comfort item and three social-contact items, each with equal weight. Each answer is between 1 and 5. To reverse an item, we subtract the answer from 6. We calculate a scale when at least three answers are valid. Six residents do not have a usable baseline scale; each analysis therefore uses 1,034 residents in 336 groups. Five have four observed intrinsic items, and five have three extrinsic items.

We divide residents using a cutpoint chosen from the distribution of the scale alone. We seek the closest possible balance between the two category sizes and put scores strictly above the cutpoint in the higher category. Intrinsic religiosity has a cutpoint of 3.2, with 498 residents below or at it and 536 above. Extrinsic religiosity has a cutpoint of 2.75, with 493 lower and 541 higher. We estimate a separate regression for each scale, following the specifications and no-leader reference convention for the other resident characteristics.

The first specification includes baseline participation and the design fixed effects. The second adds the corresponding 25-meter spillover measure and the 11 baseline controls described in Appendix Section C.2. Tests in this exploratory heterogeneity family are unadjusted for multiple comparisons.

Table C-12 – GROUP ETHNIC IDENTIFICATION AND LEADER EFFECTIVENESS

	Baseline participation control			+ Spillover and LASSO		
	Low identity	High identity	Difference	Low identity	High identity	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-2.14	5.59	7.72	-2.47	5.22	7.69
	(3.58)	(5.82)	(7.35)	(3.51)	(5.32)	(6.81)
Homogeneous	3.65	6.67	3.02	3.42	9.68**	6.26
	(3.91)	(4.39)	(6.13)	(3.85)	(3.96)	(5.73)
Religious leader						
Mixed	0.07	14.85***	14.78**	0.13	13.18***	13.05**
	(3.90)	(4.60)	(6.15)	(3.78)	(4.20)	(5.78)
Homogeneous	7.58**	-7.33	-14.91**	6.77**	-4.35	-11.13*
	(3.41)	(4.94)	(6.79)	(3.27)	(4.56)	(6.36)
Panel B. Predicted no-leader participation by identity category						
No-leader respondents in each category						
Mixed	61.88	52.42	-9.47**	61.88	52.42	-9.47***
	(2.29)	(3.51)	(4.19)	(2.21)	(2.86)	(3.62)
Homogeneous	57.48	55.84	-1.63	57.48	55.84	-1.63
	(2.05)	(2.87)	(3.52)	(2.14)	(2.70)	(3.44)
Observations	1,040			1,040		
Group clusters	336			336		
Groups containing respondents in each category (first / second)						
Chief, mixed	36 / 20			36 / 20		
Chief, homogeneous	33 / 23			33 / 23		
Pastor, mixed	29 / 27			29 / 27		
Pastor, homogeneous	29 / 27			29 / 27		

Notes: In Panel B the no-leader reference belongs to actual no-leader residents in each observed composition and identity category. Their own covariates and fixed effects enter the predictions we average, with intervals for uncertainty in the fitted model.
Does the ethnic identification of the team matter for how residents respond to leaders? Residents placed six identities in a forced ranking at baseline. Own territorial origin receives 1 when least important and 6 when most important. We average the available scores in the three ordinary-member seats, including the respondent's own when observed and excluding the leader. There are three scores in 330 groups and two in six; the resulting measure is assigned to all 1,040 outcome respondents in 336 groups. A group is higher only strictly above the group median of 10/3, with ties lower. There are 603 lower- and 437 higher-category residents. We subtract the effect in the lower category from that in the higher one. Individual identification has no separate control or interaction. Within each composition we pool disclosure for leader effects, their interactions with the group characteristic and category levels. Baseline participation and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. The second also includes spillover and 11 LASSO-selected baseline controls, with selection and missing values described in Appendix Section C.5. Parentheses give standard errors clustered by group, conditioning on category membership. Stars use unadjusted exploratory tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 4.3 and Figure 5B discuss the estimates.

Table C-13 – GROUP ETHNIC IDENTIFICATION AND COETHNICITY IN MIXED GROUPS

	Baseline participation control			+ Spillover and LASSO		
	Different territory	Same territory	Difference	Different territory	Same territory	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Lower group identification	-2.67	-1.20	1.47	-2.80	-2.11	0.69
	(4.06)	(9.82)	(11.20)	(4.03)	(9.35)	(10.74)
Higher group identification	1.60	15.62	14.02	1.15	15.08	13.93
	(6.64)	(11.86)	(13.84)	(6.52)	(10.54)	(12.99)
Panel B. Predicted no-leader participation by identity and coethnicity						
No-leader respondents in each category						
Lower group identification	61.30	64.00	2.70	61.30	64.00	2.70
	(2.77)	(6.23)	(7.30)	(2.52)	(6.42)	(7.23)
Higher group identification	54.81	45.10	-9.71	54.81	45.10	-9.71
	(3.71)	(9.27)	(10.37)	(3.13)	(8.26)	(9.39)
Observations	1,040			1,040		
Group clusters	336			336		
Mixed groups containing respondents in each category (different / same territory)						
Chief, lower identification	36 / 19			36 / 19		
Chief, higher identification	20 / 15			20 / 15		
No leader, lower identification	35 / 19			35 / 19		
No leader, higher identification	21 / 16			21 / 16		

Notes: We ask whether the comparison between the chief's coethnics and other residents changes with the group's ethnic identification. We compare residents of mixed groups who share the neighborhood chief's territorial origin with residents who come from other origins, separately by the group's ethnic identification. The difference subtracts effects for other-origin residents from effects for the chief's coethnics. We use the chief's origin to define coethnicity even when a group is assigned a pastor or no leader. Group identification is the average of ordinary members' available baseline own-territory scores, ranging from 1–6 with higher scores giving the identity more importance. Groups strictly above the fixed group median, 10/3, are higher; groups at the median or below are lower.
We retain all 1,040 residents in 336 groups in the regression, including pastor groups, although the table displays only traditional-leader effects in mixed groups. In those groups, leader assignment, group identification, and coethnicity are fully interacted, with the corresponding category terms for no-leader residents. We also estimate homogeneous-group effects separately by group identification, leaving them out of this table. All effects pool disclosure. Panel B averages predictions over actual no-leader residents in each observed identity-by-coethnicity category, using their own covariates and fixed effects; its standard errors reflect model uncertainty. Higher-identification coethnics number 19 residents in 15 chief groups and 17 residents in 16 no-leader groups.
Both blocks include baseline participation and neighborhood-by-composition-by-disclosure fixed effects. Spillover and the same 11 LASSO-selected baseline controls enter the second block, following the selection and missing-value treatment in Appendix Section C.5. Parentheses give group-clustered standard errors. Stars use unadjusted exploratory tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 4.3 discusses this breakdown of Figure 5A.

Table C-14 – OPINION OF THE ASSIGNED LEADER

	Baseline participation control			+ Spillover and LASSO		
	Index at/below median	Index above median	Difference	Index at/below median	Index above median	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-4.05	8.36**	12.41**	-3.51	6.79*	10.30**
	(3.56)	(4.00)	(5.23)	(3.38)	(3.56)	(4.51)
Homogeneous	0.95	8.12**	7.16	2.68	8.71**	6.03
	(4.18)	(3.44)	(5.05)	(4.04)	(3.37)	(4.98)
Religious leader						
Mixed	4.90	7.70*	2.80	4.07	7.56**	3.49
	(3.51)	(3.94)	(4.73)	(3.39)	(3.73)	(4.54)
Homogeneous	2.92	-1.23	-4.14	1.64	1.56	-0.08
	(3.92)	(3.18)	(5.04)	(3.67)	(3.12)	(4.89)
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed, pooled		58.33			58.33	
		(1.93)			(1.74)	
Homogeneous, pooled		56.79			56.79	
		(1.76)			(1.72)	
Observations		1,037			1,037	
Group clusters		335			335	
Groups containing respondents in each category (first / second)						
Chief, mixed		45 / 39			45 / 39	
Chief, homogeneous		46 / 48			46 / 48	
Pastor, mixed		48 / 46			48 / 46	
Pastor, homogeneous		44 / 47			44 / 47	

Notes: Are leaders more influential among residents who view them more favorably? The pre-discussion opinion index has four items: trustworthiness, hard work, morality and inspiration. Scores above the median of 0.20195 enter the second category and those at or below it, including ties, enter the first. The difference subtracts the first category's effect from the second. We verify the assigned leader being rated against meeting membership.
No-leader residents have no assigned leader to assess, so both categories share the same reference within composition in Panel B. It averages fitted predictions for actual no-leader respondents, with their own observed covariates and fixed effects. Leader effects, interactions with the characteristic and any separate category levels pool disclosure within each composition. Missing assessments lead to the loss of three residents where no group mean is available. For another 33 residents we can use their group's observed mean, and inference conditions on the categories obtained after this replacement.
Baseline participation and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. The second also includes spillover and 11 LASSO-selected baseline controls, with selection and missing values described in Appendix Section C.5. Parentheses give standard errors clustered by group, conditioning on category membership. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 4.3 discusses the estimates.

Table C-15 – PERCEIVED SECULAR AUTHORITY OF THE ASSIGNED LEADER

	Baseline participation control			+ Spillover and LASSO		
	Score ≤ 4	Score > 4	Difference	Score ≤ 4	Score > 4	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-2.33	8.25*	10.58*	-2.10	6.76*	8.86*
	(3.44)	(4.75)	(6.02)	(3.31)	(3.95)	(4.97)
Homogeneous	2.90	7.85**	4.95	4.25	8.62**	4.36
	(3.65)	(3.92)	(4.98)	(3.35)	(4.01)	(4.82)
Religious leader						
Mixed	4.68	9.14**	4.45	5.40*	6.61*	1.20
	(3.38)	(3.79)	(4.24)	(3.26)	(3.64)	(4.11)
Homogeneous	1.31	-0.08	-1.39	1.69	1.52	-0.17
	(3.29)	(3.54)	(4.69)	(3.20)	(3.57)	(4.89)
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed, pooled		58.33			58.33	
		(1.91)			(1.72)	
Homogeneous, pooled		56.79			56.79	
		(1.75)			(1.72)	
Observations		1,037			1,037	
Group clusters		335			335	
Groups containing respondents in each category (first / second)						
Chief, mixed		47 / 34			47 / 34	
Chief, homogeneous		53 / 43			53 / 43	
Pastor, mixed		52 / 38			52 / 38	
Pastor, homogeneous		51 / 39			51 / 39	

Notes: Are leaders more influential among residents who see them as having greater secular authority? Residents rated the assigned leader before discussion, and we compare scores above 4 with scores at or below 4. The difference subtracts the first category's effect from the second. We verify the assigned leader being rated against meeting membership. Missing assessments lead to the loss of three residents where no group mean is available. For another 33 residents we can use their group's observed mean, and inference conditions on the categories obtained after this replacement.
No-leader residents have no assigned leader to assess, so both categories share the same reference within composition in Panel B. It averages fitted predictions for actual no-leader respondents, with their own observed covariates and fixed effects. Leader effects, interactions with the characteristic and any separate category levels pool disclosure within each composition.
Baseline participation and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. The second also includes spillover and 11 LASSO-selected baseline controls, with selection and missing values described in Appendix Section C.5. Parentheses give standard errors clustered by group, conditioning on category membership. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 4.3 and Figure 6A discuss the estimates.

Table C-16 – COETHNICITY WITH THE LEADER

	Baseline participation control			+ Spillover and LASSO		
	Non-coethnics	Coethnics	Difference	Non-coethnics	Coethnics	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-1.05	6.09	7.13	-1.28	5.26	6.54
	(3.35)	(7.31)	(8.56)	(3.35)	(6.65)	(8.00)
Homogeneous, pooled		4.85*			6.01**	
		(2.91)			(2.78)	
Religious leader						
Mixed	5.19	9.82	4.63	4.50	9.68	5.18
	(3.55)	(7.23)	(8.60)	(3.36)	(6.78)	(8.03)
Homogeneous	1.65	0.19	-1.46	2.56	0.93	-1.62
	(3.80)	(3.39)	(5.11)	(3.72)	(3.12)	(4.87)
Panel B. Implied no-leader counterfactuals						
Traditional leader						
Mixed	58.92	56.35	-2.57	58.92	56.35	-2.57
	(2.23)	(5.17)	(5.92)	(1.98)	(4.96)	(5.63)
Homogeneous, pooled		56.79			56.79	
		(1.76)			(1.73)	
Religious leader						
Mixed	59.71	54.07	-5.64	59.71	54.07	-5.64
	(2.32)	(5.06)	(5.92)	(2.06)	(4.89)	(5.65)
Homogeneous	59.39	55.51	-3.88	59.39	55.51	-3.88
	(3.32)	(1.98)	(3.86)	(3.23)	(1.95)	(3.75)
Observations: chief / pastor models		1,040 / 1,039			1,040 / 1,039	
Group clusters (both models)		336			336	
Groups containing respondents in each category (non-coethnics / coethnics)						
Chief, mixed		56 / 34			56 / 34	
Chief, homogeneous (pooled)		56			56	
Pastor, mixed		55 / 39			55 / 39	
Pastor, homogeneous		23 / 34			23 / 34	

Notes: Coethnicity is defined by the same recorded origin as the relevant leader: the neighborhood chief for chief effects and the neighborhood pastor for pastor effects. We identify both leaders from the final assignment registries. Unlike Table C-7, the pastor rows use coethnicity with the pastor. Each leader-specific comparison comes from a separate full-sample regression that includes both leader treatments and their interactions with that coethnicity measure. The difference subtracts non-coethnics from coethnics; estimates are in percentage points.
Chief effects in homogeneous groups remain pooled, following the composition assignment. Pastor effects are split in both compositions, allowing the association between coethnicity and participation without a leader to differ by composition. One no-leader resident lacks a recorded origin, leaving 1,039 observations for the pastor comparison. In homogeneous groups, pastor coethnicity varies mainly across neighborhoods. Panel B averages fitted predictions for actual no-leader residents in each composition and relevant coethnicity category, retaining their observed covariates and fixed effects; the homogeneous chief reference is pooled. Both blocks include baseline participation and neighborhood-by-composition-by-disclosure fixed effects. The second adds the 25-meter spillover indicator and the fixed 11 LASSO-selected baseline controls (Appendix Section C.2). Leader effects, coethnicity interactions and separate category levels pool disclosure within composition. Standard errors in parentheses cluster by accountability group. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. The estimates do not establish that the coethnicity gradients differ between chiefs and pastors.

Table C-17 – INDIVIDUAL ETHNIC IDENTIFICATION AND RELIGIOSITY

	Homogeneous	Mixed
Intrinsic religiosity	+0.069 ($p = .150$)	+0.011 ($p = .785$)
Extrinsic religiosity	+0.067 ($p = .123$)	−0.023 ($p = .636$)
Christian identity, adjusted ranks	+0.173 ($p < .001$)	+0.025 ($p = .591$)
Respondents	516	518
Group clusters	168	168

Notes: Entries are Pearson correlations between respondents' own baseline ethnic identification and the religious measure in each row, with two-sided p-values in parentheses. We use complete pairs within each composition; all treatment arms are included. Intrinsic and extrinsic religiosity use the baseline scales described in Appendix Section C.5. Ethnic identification scores own territorial origin from 1 (least important) to 6 (most important) in the six-identity ranking. For the Christian-identity row, we remove Christian identity before scoring ethnicity and remove ethnic identity before scoring Christianity; the resulting scores run from 1 to 5. This removes direct competition between the two identities for a rank, but shared remaining alternatives can still induce dependence. Standard errors cluster by accountability group in standardized bivariate regressions, without additional controls or fixed effects. Tests are unadjusted for multiple comparisons. These are descriptive associations.

Table C-18 – GROUP ETHNIC IDENTIFICATION AND INDIVIDUAL RELIGIOSITY

	Homogeneous	Mixed
Intrinsic religiosity	−0.015 ($p = .756$)	−0.036 ($p = .373$)
Extrinsic religiosity	+0.029 ($p = .521$)	−0.076 ($p = .123$)
Christian identity, adjusted ranks	+0.115 ($p = .011$)	+0.024 ($p = .584$)
Respondents	516	518
Group clusters	168	168

Notes: Entries correlate respondents' religious measures with the ethnic identification of their assigned group. The group score averages available baseline scores among the three ordinary members, excluding the leader; 330 groups have three scores and six have two. The intrinsic and extrinsic rows use the original group ethnic score. For the Christian-identity row, we first remove Christian identity from each member's ranking, score ethnicity among the remaining five options, and average those scores over the same ordinary members. We pair this group score with the respondent's own Christian score after removing ethnic identity. The adjustment and remaining dependence concern are as in Table C-17. The unit of observation remains the respondent, with complete pairs in each composition and all treatment arms included. Parentheses contain two-sided, group-clustered p-values from standardized bivariate regressions without additional controls or fixed effects, unadjusted for multiple comparisons.

Table C-19 – INTRINSIC RELIGIOSITY AND LEADER EFFECTIVENESS

	Baseline participation control			+ Spillover and LASSO		
	Score ≤ 3.2	Score > 3.2	Difference	Score ≤ 3.2	Score > 3.2	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-1.00	0.76	1.76	-2.27	1.92	4.19
	(4.83)	(4.27)	(7.04)	(4.59)	(4.14)	(6.68)
Homogeneous	3.70	3.79	0.09	5.21	4.92	-0.30
	(4.47)	(4.52)	(6.83)	(4.34)	(4.12)	(6.32)
Religious leader						
Mixed	-0.58	11.75**	12.33*	0.71	9.89**	9.18
	(4.88)	(4.61)	(7.44)	(5.01)	(4.27)	(7.36)
Homogeneous	-4.01	5.34	9.35	-2.18	5.28	7.46
	(4.43)	(3.61)	(6.29)	(4.20)	(3.46)	(6.00)
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed	59.21	55.27	-3.94	58.06	56.62	-1.44
	(3.50)	(3.02)	(5.09)	(3.52)	(2.70)	(5.00)
Homogeneous	57.19	52.17	-5.02	55.47	52.27	-3.20
	(3.16)	(2.75)	(4.70)	(3.08)	(2.48)	(4.32)
Observations	1,034			1,034		
Group clusters	336			336		
Groups containing respondents in each category (first / second)						
Chief, mixed	48 / 49			48 / 49		
Chief, homogeneous	50 / 44			50 / 44		
Pastor, mixed	49 / 50			49 / 50		
Pastor, homogeneous	48 / 49			48 / 49		

Notes: Do pastors influence residents more when religion guides their conduct? Intrinsic religiosity averages baseline items 1, 3 (reversed), 4, 5, and 9 (reversed), using at least three valid answers on the 1–5 response scale. The nine-item survey adaptation follows Allport and Ross (1967)'s intrinsic/extrinsic distinction and Gorsuch and McPherson (1989)'s revised measurement. We split residents at 3.2, the discrete cutpoint that most nearly balances category sizes using only the scale distribution. There are 498 residents at or below it and 536 above. Six have no usable baseline items, leaving 1,034 residents in 336 groups. The difference column subtracts the lower category's effect from the higher category's effect.
Panel B gives common no-leader predictions within category. Leader effects, interactions with religiosity, and category levels pool disclosure within each composition. Baseline participation and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. The second also includes spillover and 11 LASSO-selected baseline controls, with selection and missing values described in Appendix Section C.5. Parentheses give standard errors clustered by group, conditioning on category membership. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. See Section 4.3 and Figure 6B.

Table C-20 – EXTRINSIC RELIGIOSITY AND LEADER EFFECTIVENESS

	Baseline participation control			+ Spillover and LASSO		
	Score $\leq$ 2.75	Score $>$ 2.75	Difference	Score $\leq$ 2.75	Score $>$ 2.75	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-6.41	6.21	12.62*	-4.30	3.46	7.76
	(5.00)	(4.22)	(7.19)	(4.52)	(3.84)	(6.28)
Homogeneous	3.76	4.16	0.40	5.96	4.77	-1.19
	(4.00)	(4.64)	(6.48)	(3.91)	(4.38)	(6.20)
Religious leader						
Mixed	7.01	5.14	-1.87	7.73*	3.63	-4.10
	(4.51)	(3.98)	(6.17)	(4.35)	(3.86)	(6.05)
Homogeneous	4.07	-2.18	-6.25	4.49	-0.95	-5.44
	(3.17)	(4.44)	(5.82)	(3.20)	(4.25)	(5.80)
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed	60.30	54.16	-6.14	60.95	54.24	-6.71
	(3.50)	(2.65)	(4.65)	(3.29)	(2.33)	(4.35)
Homogeneous	52.69	56.13	3.44	52.41	54.86	2.46
	(2.44)	(3.03)	(4.17)	(2.48)	(2.94)	(4.17)
Observations	1,034			1,034		
Group clusters	336			336		
Groups containing respondents in each category (first / second)						
Chief, mixed	46 / 45			46 / 45		
Chief, homogeneous	44 / 46			44 / 46		
Pastor, mixed	42 / 47			42 / 47		
Pastor, homogeneous	46 / 44			46 / 44		

Notes: Are leaders more influential among residents who see religion as a source of personal comfort and social contact? Our nine-item baseline survey adapts the intrinsic/extrinsic distinction in Allport and Ross (1967) and revised measurement in Gorsuch and McPherson (1989). Items 2, 6, 7, and 8 measure extrinsic religiosity. Answers are on a 1–5 scale. A resident needs at least three valid answers for us to take the equally weighted item average. We lose six residents without usable baseline items, leaving 1,034 in 336 groups. We use the scale distribution alone to choose the discrete split closest to equal category sizes: 2.75, with 493 residents at or below and 541 above. The reported difference subtracts the lower category effect from the higher one.
Panel B gives common no-leader predictions within category. Leader effects, interactions with religiosity, and category levels pool disclosure within each composition. Baseline participation and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. The second also includes spillover and 11 LASSO-selected baseline controls, with selection and missing values described in Appendix Section C.5. Parentheses give standard errors clustered by group, conditioning on category membership. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 4.3 discusses the estimates.

C.6 Which Leaders Can Exercise Their Authority?

We ask whether the assigned leader is equipped to exercise the relevant form of authority. Tables C-21–C-23 estimate Equation (2), pooling both leader effects and their interactions with leader characteristics across disclosure within each composition. The first specification controls for baseline participation and neighborhood-by-composition-by-disclosure fixed effects. The second adds spillover exposure and 11 LASSO-selected baseline controls. We include leader attributes through their interactions with leader assignment and omit a separate level term, since they are undefined for residents without an assigned leader. Standard errors are clustered by randomized group; the tables also report neighborhood-clustered standard errors and tests of the higher-minus-lower contrasts. These comparisons draw on only 28 leaders of each type, and we do not adjust the exploratory tests for multiple comparisons.

Chiefs benefit from local roots; pastors from religious conviction. If chiefs' authority comes from being embedded in local ethnic networks, chiefs who have lived in the neighborhood longer should be better able to sustain participation. A longer time living in the neighborhood gives them more knowledge of residents, stronger local ties, and greater standing in the community. These are exactly the types of relationships through which chiefs would be able to enforce participation in local collective action. Pastors, on the other hand, should depend less on local tenure because their authority should not rely as strongly on neighborhood networks or local standing. Instead, their effectiveness should depend more on how strongly they themselves believe that God's oversight rewards cooperation. Thus, we expect pastor effects to be stronger among pastors with stronger beliefs in divine enforcement.

Table C-21 shows how chief effects vary with avenue tenure, comparing chiefs who have lived on the avenue for 26 years or fewer with chiefs with longer tenure. In homogeneous groups, chiefs with longer tenure raise participation by 10.3 percentage points (SE 4.5), compared with 0.1 points for chiefs with shorter tenure (SE 3.2). The 10.2-point difference is suggestive (SE 5.2; $p = 0.051$); it remains similar with additional controls, although neighborhood-clustered inference is less precise ($p = 0.143$ without additional controls and $p = 0.152$ with them; Table C-21). The chief-tenure difference in mixed groups is close to zero. Pastor effects do not increase with their own avenue tenure: in homogeneous groups, their effect falls from 7.1 to -5.4 points, a -12.6-point difference (SE 4.4; $p = 0.005$). Local rootedness therefore has different associations with the two leaders' effectiveness.

Pastors' influence may also depend on their own conviction in the obligation they invoke. Table C-22 compares leaders with lower and higher conviction in divine moral enforcement, defined within each leader type. We combine the questions from Lang et al. (2019) asking whether God rewards good people and punishes bad people into one measure. In mixed groups, pastors in the higher-conviction category increase participation by 12.2 percentage points (SE 3.7), while those in the lower category increase it by 0.6 points (SE 4.1). The difference is 11.6 points (SE 5.1; $p = 0.024$). The adjusted contrast remains positive with neighborhood-clustered inference ($p = 0.025$). In homogeneous groups, the difference is 7.4 points and is less precise (SE 4.6; $p = 0.107$). We do not find significant variation in chief effects with chiefs' conviction (Table C-22). The comparisons support an interpretation in which pastors' conviction in religious obligation matters for their reach.

Leadership roles across institutions. Some leaders who participated in the campaign were primarily recognized for one role but also held a position in the other institution. Thus, some chiefs also reported serving as pastors, and some pastors reported serving as chiefs. Table C-23 compares leaders who head both institutions with those who do not. The cross-institution sample is small: five chiefs also head a religious institution, and seven pastors also head a traditional institution. In mixed groups, the chief effect is 7.4 points among those who head both institutions (SE 5.4), compared with -0.7 points among the other chiefs (SE 3.2). The 8.1-point difference is imprecise (SE 6.0; $p = 0.178$). Pastors who also head a traditional institution do not have larger effects in homogeneous groups; their higher-minus-lower difference is -4.8 points (SE 4.4; $p = 0.276$). The directions are consistent with religious roles extending chiefs' reach across territorial boundaries, but these pooled estimates do not establish an asymmetry between the two types of leader.

Table C-21 – LEADER EFFECTS BY YEARS OF RESIDENCE ON THE AVENUE

	Baseline participation control			+ Spillover and LASSO		
	Chief: $\leq$ 26 years Pastor: $\leq$ 27 years	Chief: $>$ 26 years Pastor: $>$ 27 years	Difference	Chief: $\leq$ 26 years Pastor: $\leq$ 27 years	Chief: $>$ 26 years Pastor: $>$ 27 years	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	0.53	0.97	0.43	0.17	0.59	0.42
	(3.87)	(3.71)	(5.02)	(3.83)	(3.42)	(4.81)
	[5.06]	[4.71]	[6.35]	[4.59]	[3.59]	[5.29]
Homogeneous	0.11	10.32**	10.21*	1.95	10.64**	8.69*
	(3.20)	(4.47)	(5.21)	(3.12)	(4.19)	(4.91)
	[3.40]	[5.82]	[6.76]	[2.80]	[5.44]	[5.89]
Religious leader						
Mixed	5.21	7.38*	2.18	5.35	6.29*	0.94
	(4.01)	(3.84)	(5.21)	(3.91)	(3.62)	(5.05)
	[4.96]	[4.53]	[6.13]	[4.80]	[3.83]	[5.68]
Homogeneous	7.14**	-5.42	-12.56***	7.50**	-4.13	-11.63***
	(3.32)	(3.30)	(4.44)	(2.94)	(3.25)	(4.09)
	[3.92]	[4.34]	[5.95]	[3.25]	[4.05]	[5.20]
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed	58.33	58.33	–	58.33	58.33	–
	(1.99)	(1.99)		(1.79)	(1.79)	
Homogeneous	56.79	56.79	–	56.79	56.79	–
	(1.77)	(1.77)		(1.75)	(1.75)	
Observations	1,040			1,040		
Group clusters	336			336		
Neighborhood clusters	28			28		
Pretreatment participation control	Yes			Yes		
Spillover and LASSO controls	No			Yes		
Traditional, mixed: neighborhood p	0.946			0.937		
Traditional, homogeneous: neighborhood p	0.143			0.152		
Religious, mixed: neighborhood p	0.726			0.869		
Religious, homogeneous: neighborhood p	0.044			0.034		

Notes: We compare leader effects according to how long the leader has lived on the avenue, in years. The first category contains chiefs with at most 26 years and pastors with at most 27 years. There are 15 chiefs in this category and 13 in the second; pastors are divided 14 in each category.
To compare these leaders, we interact their characteristic with leader assignment, without a separate level term for the characteristic, pooling disclosure within each composition for both the leader effects and the interactions. The difference reported is the second category's effect less the first, in percentage points. Participation before treatment and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. Spillover exposure and the same 11 LASSO-selected baseline controls enter the second block. Both categories of leaders are compared with the same residents assigned no leader: those residents cannot be split by an assigned leader's years on the avenue, since no leader was assigned to them. For each composition, Panel B averages their predicted participation, retaining the covariates and fixed effects observed for each resident. For each estimate we report group-clustered standard errors in parentheses and neighborhood-clustered standard errors in square brackets. The category tests at the bottom use neighborhood clustering. See Section C.6 for the results. Stars use unadjusted group-clustered tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C-22 – LEADER EFFECTS BY BELIEF IN DIVINE REWARDS AND PUNISHMENTS

	Baseline participation control			+ Spillover and LASSO		
	Chief: score ≤ 3.5 Pastor: score ≤ 4.5	Chief: score > 3.5 Pastor: score > 4.5	Difference	Chief: score ≤ 3.5 Pastor: score ≤ 4.5	Chief: score > 3.5 Pastor: score > 4.5	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	3.00	-1.38	-4.38	3.12	-2.26	-5.38
	(3.60)	(3.86)	(4.93)	(3.66)	(3.54)	(4.82)
	[4.11]	[5.10]	[5.71]	[3.66]	[4.05]	[4.73]
Homogeneous	6.64**	3.10	-3.53	7.21**	4.76	-2.45
	(3.30)	(4.44)	(5.27)	(3.37)	(4.06)	(4.98)
	[3.54]	[5.91]	[7.06]	[3.95]	[5.00]	[6.40]
Religious leader						
Mixed	0.59	12.17***	11.58**	-0.15	11.92***	12.06**
	(4.09)	(3.66)	(5.12)	(3.91)	(3.51)	(4.97)
	[5.06]	[4.18]	[5.73]	[4.30]	[3.95]	[5.10]
Homogeneous	-2.79	4.57	7.37	-0.41	3.72	4.13
	(3.27)	(3.47)	(4.56)	(3.04)	(3.31)	(4.25)
	[3.54]	[4.67]	[6.13]	[3.13]	[4.31]	[5.47]
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed	58.33	58.33	–	58.33	58.33	–
	(2.01)	(2.01)		(1.78)	(1.78)	
Homogeneous	56.79	56.79	–	56.79	56.79	–
	(1.74)	(1.74)		(1.72)	(1.72)	
Observations	1,040			1,040		
Group clusters	336			336		
Neighborhood clusters	28			28		
Pretreatment participation control	Yes			Yes		
Spillover and LASSO controls	No			Yes		
Traditional, mixed: neighborhood p	0.450			0.265		
Traditional, homogeneous: neighborhood p	0.621			0.705		
Religious, mixed: neighborhood p	0.053			0.025		
Religious, homogeneous: neighborhood p	0.240			0.457		

Notes: We compare leader effects according to the leader's baseline beliefs that God rewards good behavior and punishes improper behavior, combined in an index. The cutpoint is 3.5 for chiefs and 4.5 for pastors. There are 14 leaders of each type on each side of the cutpoint, with scores equal to it entering the first category.

The characteristic enters only through its interaction with leader assignment. We pool disclosure within each composition for both the leader effects and these interactions. The difference reported is the second category's effect less the first, in percentage points. Participation before treatment and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. We add spillover exposure and the same 11 LASSO-selected baseline controls in the second block. Residents assigned no leader have no leader characteristic on which to divide them. For that reason, the reference mean in Panel B is common to both categories. We calculate it within each composition from the no-leader residents, predicting participation with each resident's observed covariates and fixed effects and then averaging these predictions. For each estimate we report group-clustered standard errors in parentheses and neighborhood-clustered standard errors in square brackets. The category tests at the bottom use neighborhood clustering. See Section C.6 for the results. Stars use unadjusted group-clustered tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table C-23 – LEADER EFFECTS BY JOINT CHIEF AND PASTOR ROLES

	Baseline participation control			+ Spillover and LASSO		
	Does not hold both roles	Chief and pastor roles	Difference	Does not hold both roles	Chief and pastor roles	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed	-0.74	7.37	8.11	-1.02	6.63	7.65
	(3.17)	(5.41)	(6.01)	(3.09)	(5.46)	(6.17)
	[4.18]	[5.96]	[6.63]	[3.71]	[3.92]	[4.82]
Homogeneous	6.23*	-1.58	-7.81	7.55**	-1.15	-8.69
	(3.32)	(3.99)	(4.84)	(3.12)	(4.67)	(5.34)
	[3.88]	[5.25]	[6.24]	[3.46]	[6.40]	[6.90]
Religious leader						
Mixed	5.43*	8.82	3.39	4.68	9.06*	4.38
	(3.23)	(5.64)	(6.18)	(3.12)	(5.39)	(6.02)
	[4.51]	[2.84]	[4.73]	[3.97]	[4.05]	[5.28]
Homogeneous	2.05	-2.76	-4.81	3.32	-3.25	-6.58
	(2.94)	(3.73)	(4.41)	(2.74)	(3.46)	(4.03)
	[3.59]	[3.83]	[5.06]	[3.26]	[3.48]	[4.63]
Panel B. Implied no-leader counterfactuals						
Common no-leader comparison						
Mixed	58.33	58.33	–	58.33	58.33	–
	(2.01)	(2.01)		(1.81)	(1.81)	
Homogeneous	56.79	56.79	–	56.79	56.79	–
	(1.78)	(1.78)		(1.76)	(1.76)	
Observations		1,040			1,040	
Group clusters		336			336	
Neighborhood clusters		28			28	
Pretreatment participation control		Yes			Yes	
Spillover and LASSO controls		No			Yes	
Traditional, mixed: neighborhood p		0.232			0.124	
Traditional, homogeneous: neighborhood p		0.222			0.218	
Religious, mixed: neighborhood p		0.480			0.413	
Religious, homogeneous: neighborhood p		0.350			0.167	

Notes: We compare leaders who hold positions in both institutions with the remaining leaders. The second category contains leaders recorded as both an avenue/customary chief and a pastor, cure, or reverend: 5 chiefs and 7 pastors. The first category contains the remaining 23 chiefs and 21 pastors.
To compare these leaders, we interact their characteristic with leader assignment, without a separate level term for the characteristic, pooling disclosure within each composition for both the leader effects and the interactions. The difference reported is the second category's effect less the first, in percentage points. Participation before treatment and neighborhood-by-composition-by-disclosure fixed effects enter both blocks. Spillover exposure and the same 11 LASSO-selected baseline controls enter the second block. Both categories of leaders are compared with the same residents assigned no leader: those residents cannot be split by the offices held by an assigned leader, since no leader was assigned to them. For each composition, Panel B averages their predicted participation, retaining the covariates and fixed effects observed for each resident. For each estimate we report group-clustered standard errors in parentheses and neighborhood-clustered standard errors in square brackets. The category tests at the bottom use neighborhood clustering. See Section C.6 for the results. Stars use unadjusted group-clustered tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Denominational ties may extend religious authority. Collected data also allow us to establish whether residents share the neighborhood pastor's religious denomination. This matters for whether mobilization needs to occur through a familiar set of dogmas or can hold through a broad religious framing of obligation. Table C-24 shows that pastors raise participation by 7.6 points among mixed-group residents who do not share their denomination (SE 3.1; $p = 0.015$). Within-denomination links are therefore not necessary for this effect. In homogeneous groups, the pastor effect among residents who share his denomination is positive, at 12.3 points, but imprecise (SE 8.3; $p = 0.142$); its difference from the effect among other residents is also imprecise ($p = 0.155$). Only 42 religious-arm groups contain a resident who shares the pastor's denomination, so there is suggestive evidence that parochial ties extend pastors' reach into homogeneous groups.

Table C-24 – DENOMINATIONAL TIES AND LEADER EFFECTIVENESS

	Denomination relative to the neighborhood pastor		
	Different denomination	Same denomination	Difference
Panel A. Leader effects relative to no leader			
Traditional leader			
Mixed	1.39	-6.70	-8.09
	(3.31)	(9.46)	(10.66)
Homogeneous	3.14	11.99	8.85
	(3.34)	(9.15)	(10.20)
Religious leader			
Mixed	7.61**	-3.69	-11.30
	(3.11)	(10.14)	(10.85)
Homogeneous	-0.96	12.26	13.23
	(2.90)	(8.33)	(9.29)
Panel B. Implied no-leader counterfactuals			
No leader, mixed	56.93	55.66	-1.26
	(2.22)	(5.87)	(6.55)
No leader, homogeneous	54.89	50.26	-4.64
	(1.95)	(5.51)	(5.95)
Observations		1,032	
Group clusters		336	

Notes: The difference column reports the effect in the second category minus the effect in the first. Panel B reports model-implied no-leader participation in percentage points. Predictions set leader terms to zero and baseline participation to its estimation-sample mean, averaging the fixed effects and category levels over the combined chief and pastor participants within each composition and category. Both categories compare the respondent's denomination with the neighborhood pastor's denomination. The comparison is defined for residents in every assignment arm, including no leader. Effects and interactions are pooled across disclosure within each composition; the model also includes separate composition-specific category levels, baseline participation, and neighborhood-by-composition-by-disclosure fixed effects. The sample includes 1,032 residents in 336 groups. Forty-two pastor-arm groups contain at least one resident who shares the pastor's denomination. Standard errors in parentheses are clustered by accountability group. Appendix Section C.6 discusses these estimates. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D Source of Enforcement

D.1 Main Effects, Alternative Outcomes, and Inference

Table D-1 – LEADER EFFECTS BY COMPOSITION AND DISCLOSURE

	Composition and disclosure		
	Baseline	+ Spillover	+ LASSO
Panel A. Leader effects relative to no leader			
Traditional leader			
Mixed, without disclosure	-0.29	-0.31	-1.11
	(3.48)	(3.48)	(3.48)
Mixed, with disclosure	1.75	1.67	1.75
	(4.53)	(4.52)	(4.16)
Homogeneous, without disclosure	-0.31	-0.34	1.37
	(4.46)	(4.45)	(4.27)
Homogeneous, with disclosure	9.95***	10.13***	10.62***
	(3.63)	(3.63)	(3.47)
Religious leader			
Mixed, without disclosure	9.18**	9.16**	9.02**
	(3.97)	(3.98)	(3.83)
Mixed, with disclosure	3.41	3.40	2.48
	(4.30)	(4.32)	(3.99)
Homogeneous, without disclosure	1.66	1.67	1.64
	(3.81)	(3.80)	(3.56)
Homogeneous, with disclosure	0.01	-0.01	1.71
	(3.27)	(3.26)	(3.07)
Panel B. Treatment-effect contrasts			
Mixed, with: T − R	-1.66	-1.73	-0.73
	(4.40)	(4.41)	(4.15)
Mixed, without: T − R	-9.47**	-9.47**	-10.13**
	(3.98)	(4.01)	(4.17)
Homogeneous, with: T − R	9.94**	10.13**	8.91**
	(4.03)	(4.00)	(3.67)
Homogeneous, without: T − R	-1.97	-2.01	-0.27
	(4.13)	(4.12)	(3.90)
Δ_E: H, with − M, without	19.41***	19.60***	19.04***
	(5.67)	(5.69)	(5.57)
Panel C. No-leader reference means and diagnostics			
Mixed, without disclosure	58.88	58.88	58.88
	(2.46)	(2.44)	(2.24)
Mixed, with disclosure	57.79	57.79	57.79
	(3.14)	(3.14)	(2.81)
Homogeneous, without disclosure	57.79	57.79	57.79
	(2.91)	(2.91)	(2.76)
Homogeneous, with disclosure	55.80	55.80	55.80
	(1.96)	(1.98)	(2.05)
Wald: all no-leader cells equal	p = 0.789	p = 0.790	p = 0.781
Observations	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336
Neighborhoods	28	28	28
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes
Pretreatment participation control	Yes	Yes	Yes
25m spillover control	No	Yes	Yes
LASSO-selected baseline controls	No	No	Yes

Notes: Panel A reports leader effects relative to no leader; Panel B reports direct leader contrasts; Panel C reports no-leader reference means. Leader effects vary by composition and disclosure; Δ_E compares the traditional-minus-religious gap in homogeneous groups with disclosure with the gap in mixed groups without disclosure. The outcome is average participation on six eligible days after Meeting 1 through the Meeting 2 day, in percentage points. Every column controls for pretreatment participation and neighborhood-by-composition-by-disclosure fixed effects. The second adds the 25-meter spillover indicator; the third also adds 11 LASSO-selected baseline controls. Group-clustered standard errors appear in parentheses. In the preferred specification, disclosure changes the chief effect in homogeneous groups by 9.24 percentage points (SE 5.52; $p = 0.095$) and the pastor effect in mixed groups by −6.55 points (SE 5.49; $p = 0.233$); both contrasts use the full coefficient covariance matrix. Section 5 and Figure 7 discuss these estimates. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D-2 – LABOR CONTRIBUTION OUTCOMES BY COMPOSITION AND DISCLOSURE

	Levels				Asinh robustness
	Participation rate	Attendance days	Total hours	Total effort	Asinh total hours
Panel A. Leader effects relative to no leader					
Traditional leader					
Mixed, without disclosure	-0.011	-0.027	-0.669	-0.256	-0.033
	(0.035)	(0.217)	(0.855)	(0.737)	(0.163)
Mixed, with disclosure	0.017	0.103	0.027	0.074	0.015
	(0.042)	(0.246)	(0.906)	(0.816)	(0.155)
Homogeneous, without disclosure	0.014	0.049	0.663	0.386	0.031
	(0.043)	(0.248)	(0.906)	(0.812)	(0.151)
Homogeneous, with disclosure	0.106***	0.583***	2.098**	1.968***	0.440***
	(0.035)	(0.207)	(0.887)	(0.703)	(0.134)
Religious leader					
Mixed, without disclosure	0.090**	0.580**	1.701*	1.522**	0.309**
	(0.038)	(0.235)	(0.937)	(0.756)	(0.151)
Mixed, with disclosure	0.025	0.151	0.001	0.555	0.015
	(0.040)	(0.238)	(0.931)	(0.809)	(0.154)
Homogeneous, without disclosure	0.016	0.035	-0.032	-0.137	0.029
	(0.036)	(0.210)	(0.806)	(0.706)	(0.129)
Homogeneous, with disclosure	0.017	0.108	0.130	0.519	0.131
	(0.031)	(0.186)	(0.816)	(0.690)	(0.113)
Panel B. Treatment-effect contrasts					
Mixed, without: T − R	-0.101**	-0.607**	-2.369**	-1.778**	-0.342**
	(0.042)	(0.256)	(0.965)	(0.843)	(0.174)
	[Holm p = 0.058]	[Holm p = 0.058]	[Holm p = 0.058]	[Holm p = 0.058]	[Holm p = 0.050]
Mixed, with: T − R	-0.007	-0.048	0.025	-0.481	0.000
	(0.042)	(0.249)	(0.961)	(0.834)	(0.160)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]
Homogeneous, without: T − R	-0.003	0.015	0.695	0.523	0.003
	(0.039)	(0.229)	(0.821)	(0.767)	(0.145)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.986]
Homogeneous, with: T − R	0.089**	0.474**	1.967**	1.449*	0.309**
	(0.037)	(0.217)	(0.903)	(0.749)	(0.136)
	[Holm p = 0.063]	[Holm p = 0.088]	[Holm p = 0.088]	[Holm p = 0.088]	[Holm p = 0.024]
Δ_E	0.190***	1.081***	4.337***	3.226***	0.651***
	(0.056)	(0.336)	(1.324)	(1.129)	(0.221)
	[Holm p = 0.003]	[Holm p = 0.004]	[Holm p = 0.004]	[Holm p = 0.005]	[Holm p = 0.003]
Panel C. No-leader reference means and diagnostics					
Mixed, without disclosure	0.589	3.478	13.438	11.571	2.683
	(0.022)	(0.140)	(0.578)	(0.460)	(0.099)
Mixed, with disclosure	0.578	3.413	12.983	11.098	2.654
	(0.028)	(0.165)	(0.615)	(0.553)	(0.105)
Homogeneous, without disclosure	0.578	3.457	12.877	11.272	2.622
	(0.028)	(0.161)	(0.626)	(0.527)	(0.095)
Homogeneous, with disclosure	0.558	3.272	12.529	10.473	2.565
	(0.021)	(0.126)	(0.573)	(0.461)	(0.078)
Wald: four cells equal	p = 0.781	p = 0.690	p = 0.736	p = 0.391	p = 0.798
Observations	1,040	1,040	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336	336	336
Neighborhoods	28	28	28	28	28
Matched pretreatment outcome control	Yes	Yes	Yes	Yes	Yes
25m spillover control	Yes	Yes	Yes	Yes	Yes
Frozen participation LASSO controls	Yes	Yes	Yes	Yes	Yes
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes

Notes: The exposition of results in participation shares facilitates their interpretation, but underplays how much residents contributed. We therefore follow the same leader comparisons in labor days, hours and effort as well as the daily participation rate.
Panel A reports each leader-assignment effect relative to no leader. Panel B makes the direct traditional-minus-religious comparisons and reports Δ_E, the traditional-minus-religious contrast in homogeneous groups with disclosure less that in mixed groups without disclosure. Panel C reports model-implied no-leader means and tests whether they are equal.
Residents who work once for a short time and those who work once for a long time have the same attendance count. Hours distinguish these contributions. We report the daily participation rate, count attendance in labor days, and sum hours and recorded effort over six days, strictly after Meeting 1 through and including the day of Meeting 2.
All specifications control for the resident's pretreatment measure, matched to the outcome in that column. The neighborhood-by-disclosure-by-composition fixed effects, 25-meter spillover indicator and 11 controls from the primary participation selection enter as well. The selection is still the primary participation selection, even for other measures and observation windows.
Meeting 1 does not contribute to these outcomes. A resident who did not attend has a zero, and that resident is also kept in the inverse-hyperbolic-sine hours regression; its sample is unchanged.
For any given contrast, the four level outcomes enter one Holm adjustment and the asinh-hours outcome enters another. Brackets give these adjusted p-values. We cluster standard errors by group within neighborhood and put them in parentheses. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D-3 – CASH AND IN-KIND CONTRIBUTIONS BY COMPOSITION AND DISCLOSURE

	Levels				Asinh robustness	
	Any cash contribution	Cash total	Any in-kind contribution	In-kind value	Asinh cash total	Asinh in-kind value
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed, without disclosure	0.014	31.786	-0.074*	-258.738**	0.137	-0.653**
	(0.022)	(44.980)	(0.041)	(127.396)	(0.173)	(0.318)
Mixed, with disclosure	0.046*	-1.842	0.065	149.692	0.305*	0.518
	(0.025)	(30.774)	(0.047)	(154.389)	(0.182)	(0.366)
Homogeneous, without disclosure	0.012	6.736	-0.001	-404.699*	0.087	-0.180
	(0.018)	(8.137)	(0.058)	(216.039)	(0.119)	(0.459)
Homogeneous, with disclosure	0.013	-9.239	0.015	-1861.438	0.084	0.061
	(0.020)	(27.015)	(0.054)	(1332.055)	(0.150)	(0.428)
Religious leader						
Mixed, without disclosure	0.002	5.595	-0.011	-96.757	0.034	-0.112
	(0.026)	(29.524)	(0.039)	(138.257)	(0.178)	(0.300)
Mixed, with disclosure	0.010	-19.699	0.117**	81.611	0.059	0.950**
	(0.020)	(29.577)	(0.054)	(140.508)	(0.148)	(0.427)
Homogeneous, without disclosure	0.012	3.165	0.067	-0.268	0.076	0.486
	(0.020)	(8.133)	(0.059)	(237.941)	(0.133)	(0.472)
Homogeneous, with disclosure	-0.010	-29.877	-0.065	-1919.618	-0.086	-0.588
	(0.018)	(24.873)	(0.055)	(1336.608)	(0.131)	(0.441)
Panel B. Treatment-effect contrasts						
Mixed, without: T − R	0.012	26.190	-0.064	-161.981	0.103	-0.540
	(0.031)	(48.356)	(0.046)	(122.633)	(0.232)	(0.347)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.666]	[Holm p = 0.666]	[Holm p = 0.658]	[Holm p = 0.241]
Mixed, with: T − R	0.036	17.857	-0.052	68.082	0.247	-0.432
	(0.025)	(24.618)	(0.057)	(187.360)	(0.182)	(0.455)
	[Holm p = 0.647]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.350]	[Holm p = 0.350]
Homogeneous, without: T − R	-0.000	3.571	-0.067	-404.431*	0.011	-0.666
	(0.017)	(6.854)	(0.062)	(208.376)	(0.111)	(0.492)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.830]	[Holm p = 0.212]	[Holm p = 0.922]	[Holm p = 0.353]
Homogeneous, with: T − R	0.023	20.638	0.081	58.181	0.170	0.649
	(0.020)	(21.559)	(0.058)	(940.972)	(0.144)	(0.449)
	[Holm p = 0.735]	[Holm p = 0.735]	[Holm p = 0.647]	[Holm p = 0.951]	[Holm p = 0.299]	[Holm p = 0.299]
Δ_E	0.011	-5.552	0.144*	220.161	0.067	1.189**
	(0.037)	(52.956)	(0.074)	(948.659)	(0.273)	(0.568)
	[Holm p = 1.000]	[Holm p = 1.000]	[Holm p = 0.203]	[Holm p = 1.000]	[Holm p = 0.805]	[Holm p = 0.074]
Panel C. No-leader reference means and diagnostics						
Mixed, without disclosure	0.033	38.587	0.293	572.826	0.225	2.334
	(0.010)	(17.678)	(0.025)	(100.143)	(0.068)	(0.194)
Mixed, with disclosure	0.011	38.043	0.239	514.130	0.096	1.909
	(0.014)	(25.631)	(0.031)	(67.275)	(0.108)	(0.235)
Homogeneous, without disclosure	0.011	5.435	0.261	747.826	0.075	2.123
	(0.014)	(6.318)	(0.037)	(172.380)	(0.096)	(0.296)
Homogeneous, with disclosure	0.022	38.043	0.315	2490.217	0.170	2.531
	(0.013)	(21.843)	(0.037)	(1201.373)	(0.099)	(0.299)
Wald: four cells equal	p = 0.526	p = 0.122	p = 0.354	p = 0.236	p = 0.558	p = 0.349
Observations	1,040	1,040	1,040	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336	336	336	336
Neighborhoods	28	28	28	28	28	28
Matched pretreatment outcome control	Yes	Yes	Yes	Yes	Yes	Yes
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Labor is one form of contribution. Here we examine cash and goods.
Each leader assignment is compared with no leader in Panel A. Panel B reports the direct traditional-minus-religious contrasts and Δ_E, the traditional-minus-religious contrast in homogeneous groups with disclosure less that in mixed groups without disclosure. Panel C gives the model-implied means without a leader and tests whether they are equal. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.
Residents can give cash and goods as well as work. We record whether they gave any of each and the value of each in Congolese francs (CDF) over six days, strictly after Meeting 1 through and including the Meeting 2 day.
The resident's pretreatment outcome is held constant in each model, matched to the outcome being studied, along with neighborhood × disclosure × composition fixed effects. We cluster standard errors by group within neighborhood and put them in parentheses. For the same contrast, brackets report Holm-adjusted p-values across the four level outcomes, with a separate family of the two asinh outcomes.
Meeting 1 itself is excluded. Residents with zero contributions stay in both inverse-hyperbolic-sine specifications. Combined material contributions add cash and the recorded value of in-kind giving. Before estimating the models, we check the source identity combined material value = cash + in-kind.

Table D-4 – RANDOMIZATION INFERENCE AND ALTERNATIVE CLUSTERING BY COMPOSITION AND DISCLOSURE

	Randomization inference			Alternative clustering		
	Estimate	Clustered SE	RI p	Group	Leader cluster	Neighborhood
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed, without disclosure	-1.11	(3.48)	0.807	-1.11	-1.11	-1.11
				(3.48)	(3.43)	(3.59)
Mixed, with disclosure	1.75	(4.16)	0.680	1.75	1.75	1.75
				(4.16)	(4.15)	(4.93)
Homogeneous, without disclosure	1.37	(4.27)	0.747	1.37	1.37	1.37
				(4.27)	(4.27)	(5.36)
Homogeneous, with disclosure	10.62***	(3.47)	0.016	10.62***	10.62***	10.62**
				(3.47)	(3.50)	(4.16)
Religious leader						
Mixed, without disclosure	9.02**	(3.83)	0.036	9.02**	9.02**	9.02*
				(3.83)	(3.79)	(4.48)
Mixed, with disclosure	2.48	(3.99)	0.567	2.48	2.48	2.48
				(3.99)	(4.05)	(4.74)
Homogeneous, without disclosure	1.64	(3.56)	0.704	1.64	1.64	1.64
				(3.56)	(3.60)	(3.87)
Homogeneous, with disclosure	1.71	(3.07)	0.686	1.71	1.71	1.71
				(3.07)	(3.03)	(3.24)
Panel B. Treatment-effect contrasts						
Mixed, with: T − R	-0.73	(4.15)	0.869	-0.73	-0.73	-0.73
				(4.15)	(4.22)	(4.69)
Mixed, without: T − R	-10.13**	(4.17)	0.025	-10.13**	-10.13**	-10.13*
				(4.17)	(4.19)	(5.07)
Homogeneous, with: T − R	8.91**	(3.67)	0.051	8.91**	8.91**	8.91*
				(3.67)	(3.66)	(4.47)
Homogeneous, without: T − R	-0.27	(3.90)	0.954	-0.27	-0.27	-0.27
				(3.90)	(3.95)	(4.43)
Δ_E: H, with − M, without	19.04***	(5.57)	0.002	19.04***	19.04***	19.04***
				(5.57)	(5.55)	(6.83)
Panel C. No-leader reference means and diagnostics						
Mixed, without disclosure	58.88	(2.24)		58.88	58.88	58.88
				(2.24)	(2.25)	(2.27)
Mixed, with disclosure	57.79	(2.81)		57.79	57.79	57.79
				(2.81)	(2.82)	(2.86)
Homogeneous, without disclosure	57.79	(2.76)		57.79	57.79	57.79
				(2.76)	(2.77)	(2.81)
Homogeneous, with disclosure	55.80	(2.05)		55.80	55.80	55.80
				(2.05)	(2.06)	(2.09)
Wald: all no-leader cells equal	p = 0.781			p = 0.781	p = 0.650	p = 0.659
Observations	1,040			1,040	1,040	1,040
Clusters				336	84	28
Neighborhoods	28			28	28	28
Neighborhood × disclosure × composition FEs				Yes	Yes	Yes
Pretreatment participation control	Yes			Yes	Yes	Yes
25m spillover control	Yes			Yes	Yes	Yes
LASSO-selected baseline controls	11			11	11	11
Group-neighborhood clusters	336					
NQC fixed effects	Yes					
Formation-stage permutations	5,000					

Notes: The observation is a resident, but the experiment formed teams and assigned the same leader to several teams. We examine inference from the formation stage and from allowing correlation across these teams. The outcome throughout is average salongo participation after the first meeting, in percentage points, using the same 1,040 residents. Six-day pretreatment participation, the 25-meter spillover indicator and the 11 baseline controls selected by LASSO for primary participation enter every regression, along with neighborhood × disclosure × composition fixed effects. The main specification and comparisons are unchanged. Panel A gives each leader-assignment effect relative to no leader in the same fixed-effect cell. Panel B gives the main table's leader contrasts; Δ_E subtracts the Traditional-minus-Religious contrast in mixed groups without disclosure from that in homogeneous groups with disclosure. Panel C gives model-implied no-leader reference means and clustered delta-method tests of equality among all displayed means.

Randomization inference (RI) repeats the group-formation procedure across 5,000 assignments while holding the realized sample fixed. That means retaining each wave's group sizes, random filling order, composition and fallback rules, gender × wealth targets, territorial diversity in mixed groups and treatment coding, without resampling residents or repeating the original best-balance selection. It also means using the same 11 selected controls rather than selecting them again; their coefficients are reestimated on each replay. We report two-sided RI p-values and standard errors clustered one-way by group–neighborhood. Holding exposure fixed requires a qualification. This regressor excludes households belonging to the same group, and its value could change when that group changes. Our draws need not preserve observed exposure, yet we use the observed indicator in their regressions. This conditional group-formation replay gives diagnostics, not exact randomization tests conditional on exposure or inference from the full rerandomization design. Stars use clustered asymptotic inference: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Finally, we allow errors to be correlated across groups assigned to the same leader. These alternative-clustering columns use 336 experimental group–neighborhood clusters, 84 neighborhood-wave × leader-arm clusters containing four traditional, religious or no-leader groups each, or 28 neighborhood-wave assignments for neighborhood clustering. Only this level changes: sample, outcome, all controls and fixed effects, leader effects, contrasts and no-leader comparisons remain the same. In Panels B and C we use the clustered delta method. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.2 Additional Outcome Windows

Our main outcome measures participation over six administrative days, starting strictly after the first meeting and including the second-meeting day. Does the comparison between leaders depend on this window? Figure D-1 locates the main and alternative periods on the implementation timeline, and Table D-5 gives estimates for the main window, the intervening middle period, the final week of each wave, and all eligible post-treatment periods. In making these comparisons, we retain the same residents, fixed effects, common pretreatment participation control, 25-meter spillover indicator, 11 baseline controls selected for primary participation, and clustering level. The controls are selected once, rather than again for each window.

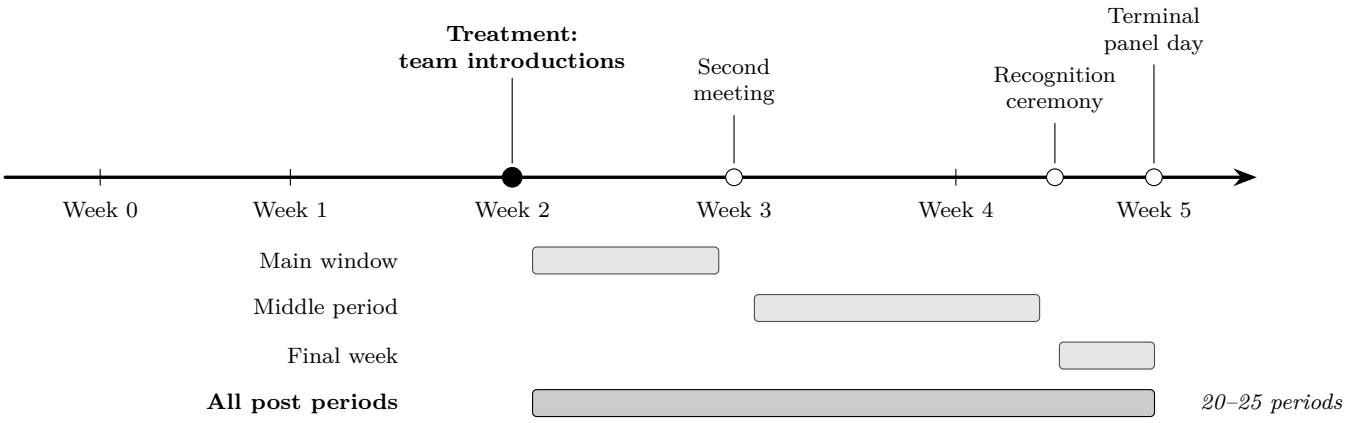


Figure D-1: Outcome windows used in the timing robustness analysis.

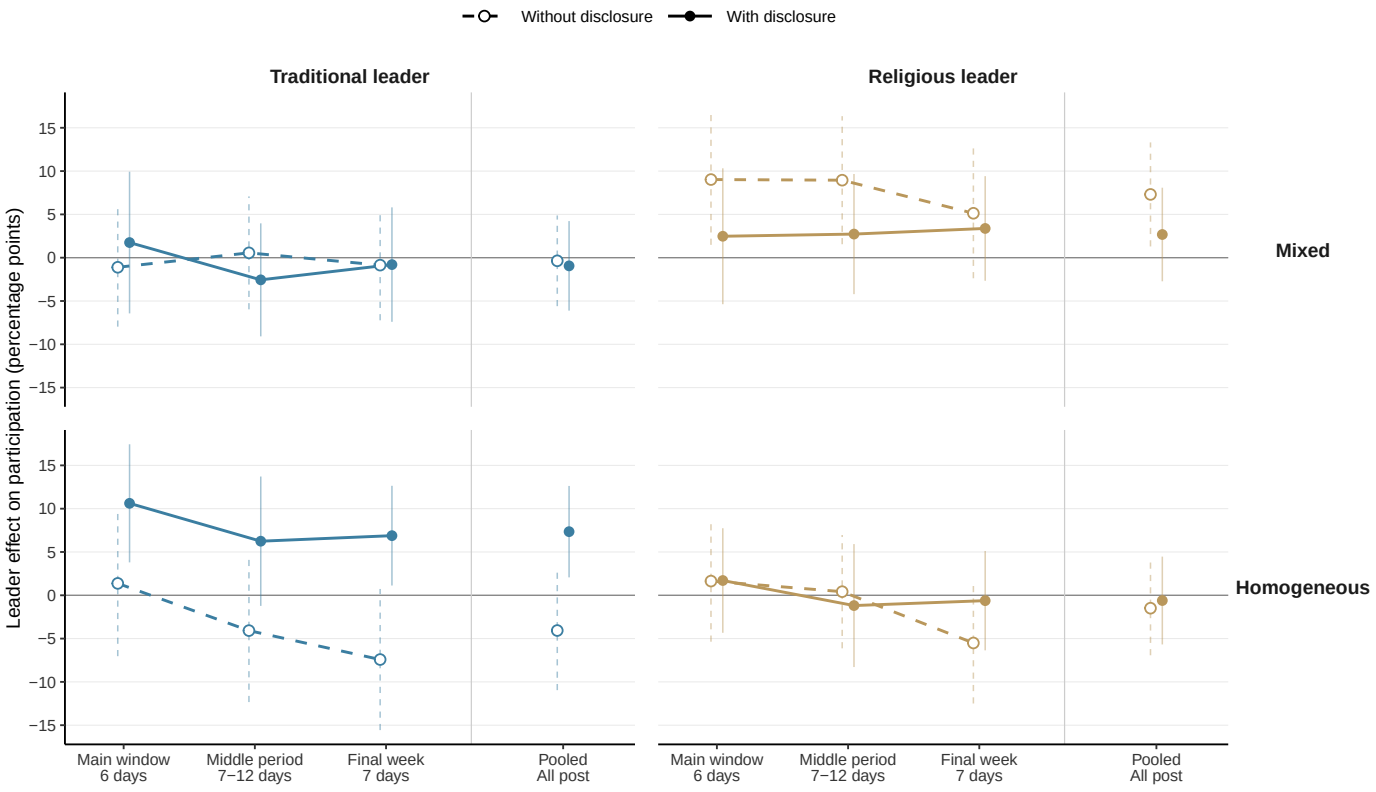


Figure D-2: LEADER EFFECTS ACROSS POST-TREATMENT OUTCOME WINDOWS

Notes: Points report each leader's effect on participation relative to no leader within the indicated composition and disclosure condition. Lines connect the main window, middle period, and final week; the separate points to the right pool all eligible post-treatment days. Vertical bars are 95 percent confidence intervals clustered at the group–neighborhood level. Estimates reproduce Appendix Table D-5, using 1,040 respondents and 336 clusters, neighborhood × disclosure × composition fixed effects, and pretreatment participation. All specifications also include the 25-meter spillover measure and the same 11 LASSO-selected baseline controls as the main specification. Appendix Figure D-1 describes the outcome windows.

Table D-5 – ROBUSTNESS TO ALTERNATIVE POST-OUTCOME WINDOWS

	Leader by composition					Environment specific			
	Main window	Middle period	Final week	All post		Main window	Middle period	Final week	All post
Panel A. Leader effects relative to no leader									
Traditional leader									
Mixed	0.37	-0.99	-0.81	-0.63	Without disclosure	-1.11	0.56	-0.86	-0.36
	(2.74)	(2.38)	(2.33)	(1.89)		(3.48)	(3.32)	(3.24)	(2.66)
					With disclosure	1.75	-2.56	-0.80	-0.94
						(4.16)	(3.32)	(3.36)	(2.63)
Homogeneous	5.98**	1.06	-0.29	1.62	Without disclosure	1.37	-4.08	-7.42*	-4.07
	(2.78)	(2.89)	(2.65)	(2.28)		(4.27)	(4.20)	(4.14)	(3.50)
					With disclosure	10.62***	6.24	6.88**	7.35***
						(3.47)	(3.80)	(2.93)	(2.68)
Religious leader									
Mixed	5.80**	5.86**	4.28*	5.02**	Without disclosure	9.02**	8.95**	5.13	7.31**
	(2.80)	(2.59)	(2.44)	(2.06)		(3.83)	(3.76)	(3.82)	(3.05)
					With disclosure	2.48	2.73	3.38	2.68
						(3.99)	(3.53)	(3.07)	(2.75)
Homogeneous	1.66	-0.40	-3.08	-1.06	Without disclosure	1.64	0.42	-5.51	-1.49
	(2.37)	(2.48)	(2.35)	(1.93)		(3.56)	(3.33)	(3.55)	(2.76)
					With disclosure	1.71	-1.19	-0.62	-0.60
						(3.07)	(3.61)	(2.92)	(2.58)
Panel B. Treatment-effect contrasts									
Mixed: T − R	-5.43*	-6.85**	-5.09**	-5.65***	Mixed, with: T − R	-0.73	-5.29	-4.17	-3.62
	(2.96)	(2.71)	(2.47)	(2.13)		(4.15)	(3.93)	(3.27)	(2.92)
					Mixed, without: T − R	-10.13**	-8.39**	-5.99	-7.67**
						(4.17)	(3.77)	(3.68)	(3.09)
Homogeneous: T − R	4.32	1.46	2.79	2.68	Homogeneous, with: T − R	8.91**	7.43**	7.50**	7.95***
	(2.72)	(2.66)	(2.43)	(2.09)		(3.67)	(3.47)	(3.04)	(2.66)
					Homogeneous, without: T − R	-0.27	-4.50	-1.91	-2.58
						(3.90)	(3.88)	(3.61)	(3.06)
Δ_C: composition gradient	9.75**	8.31**	7.88**	8.33***	Δ_E: H, with − M, without	19.04***	15.82***	13.49***	15.62***
	(4.06)	(3.82)	(3.48)	(3.01)		(5.57)	(5.16)	(4.80)	(4.09)
Panel C. No-leader reference means and diagnostics									
Mixed	58.33	42.81	26.63	39.65	Without disclosure	58.88	42.38	27.64	39.90
	(1.79)	(1.54)	(1.62)	(1.26)		(2.24)	(2.32)	(2.34)	(1.84)
					With disclosure	57.79	43.24	25.62	39.40
						(2.81)	(1.99)	(2.24)	(1.70)
Homogeneous	56.79	43.86	27.95	40.30	Without disclosure	57.79	45.53	32.45	42.75
	(1.73)	(1.89)	(1.80)	(1.50)		(2.76)	(2.55)	(2.87)	(2.26)
					With disclosure	55.80	42.18	23.45	37.85
						(2.05)	(2.78)	(1.99)	(1.89)
Wald: all no-leader cells equal	p = 0.536	p = 0.668	p = 0.585	p = 0.738		p = 0.781	p = 0.781	p = 0.071	p = 0.421
Observations	1,040	1,040	1,040	1,040		1,040	1,040	1,040	1,040
Group-neighborhood clusters	336	336	336	336		336	336	336	336
Neighborhoods	28	28	28	28		28	28	28	28
Post periods: mean	6.0	9.5	7.0	22.5		6.0	9.5	7.0	22.5
Post periods: range	6–6	7–12	7–7	20–25		6–6	7–12	7–7	20–25
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
Pretreatment participation control	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
25m spillover control	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes
11 LASSO-selected baseline controls	Yes	Yes	Yes	Yes		Yes	Yes	Yes	Yes

Notes: We compare leaders over four periods of the work record. The main six days run from the day after the assigned first meeting through the assigned second-meeting day, which is included. The middle period starts at the assigned second-meeting cutoff and stops the day before the final week. The final week after recognition runs from $T-6$ to T, the panel's last day. Finally, we average all non-treatment days after the assigned first meeting.
In making these comparisons we retain the common pre-meeting participation baseline, the 25-meter exposure indicator, the same 11 LASSO-selected baseline controls as the main specification, and neighborhood × disclosure × composition fixed effects. Panel A compares each leader assignment with no leader in its fixed-effect cell. To compare chiefs and pastors directly, Panel B reports traditional-minus-religious effects and their gradients across composition and environment. Panel C reports the model-implied means without a leader and tests whether they are equal. For both panels we obtain inference from the clustered delta method. Standard errors use one-way group–neighborhood clustering. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.3 Alternative Explanations Under Disclosure

Panels C–D of Tables B-9, C-5, and C-6 extend the speech, co-arrival, and follow-up comparisons to disclosure. We also ask whether residents report more contact with leaders or ordinary teammates than residents in no-leader groups.

Do leaders follow up with residents outside the meetings? At the start of Meeting 2, residents reported whether, during the preceding week, they had interacted with their assigned leader, discussed *salongo*, been asked to participate, or heard about social or divine consequences. Table C-6 compares these reports across composition and disclosure conditions. We use one report per respondent about the assigned leader, leaving out respondent-leaders, ordinary-peer reports, and no-leader groups. The sample has 299 respondents assigned a chief and 302 assigned a pastor, all with observed pretreatment participation.

We run separate regressions for the two leader types. Panels A–B pool disclosure and report adjusted levels and composition differences from a regression on a homogeneous-group indicator and mean-centered respondent pretreatment participation. Panels C–D add disclosure and its interaction with composition, giving adjusted levels and disclosure differences. All adjusted levels use the same pooled leader-report baseline mean. Standard errors cluster by neighborhood so that groups assigned to the same leader may be correlated. Composition differences come from the pooled regression and need not equal a 50/50 average of the disclosure-specific levels.

For the analysis of outcomes directed toward specific group members, let j index the referenced person and let $J_{jg} = 1$ indicate that this person is the assigned leader. The general extension of equation (1) is

$$Y_{ij} = \lambda_{n(i),q(i),c(i)} + \rho Y_{ij0} + \sum_{e \in \mathcal{E}} \Big[\left(\tau_e^T + \kappa_e^T J_{jg}\right) T_{g(i)} + \left(\tau_e^R + \kappa_e^R J_{jg}\right) R_{g(i)} \Big] \mathbf{1}\{e(i) = e\} + X_i'\beta + \varepsilon_{ij}. \quad (5)$$

The outcome Y_{ij} concerns a particular person; Y_{ij0} is its pretreatment value when we include one. For an ordinary participant as the target, the leader-assignment effect is τ_e^L. For the assigned leader as the target, it is $\tau_e^L + \kappa_e^L$. The contact regressions in Table D-6 use this treatment structure separately in each environment, with neighborhood fixed effects. They control for the respondent's pretreatment participation in place of the dyad-specific baseline outcome and additional controls. We cluster standard errors by randomized group within neighborhood. These reports establish whether contact occurred, but cannot tell us who initiated it or how intensive the follow-up was.

Leader contact relative to no-leader groups. Appendix Table D-6 presents results for several types of recorded interactions. Reported contact with leaders is not greater in the environments where they induce more participation. In homogeneous groups with disclosure, reported contact with the assigned chief is only 5.0 percentage points higher (SE 6.5) than the rates of contact between ordinary members of no-leader groups. In mixed groups with disclosure, where the chief's participation effect is small, reported contact with the assigned chief is 20.8 points higher (SE 6.6) than contact between ordinary members of no-leader groups. Pastors show a similar mismatch. In mixed groups without disclosure, where the participation effect is largest, respondents report nearly the same rate of contact with the assigned pastor as between ordinary members of no-leader groups, a difference of 6.5 percentage points (SE 5.4). Requests, references to consequences, and interactions among ordinary participants show no offsetting pattern. If anything, respondents report more contact with chiefs where their effects on participation are weaker. The recorded contacts therefore provide little support for differences in follow-up effort as an explanation of where leaders mobilize participation.

Table D-6 – INTERACTIONS ABOUT PARTICIPATION AND ITS CONSEQUENCES

	Environment-specific effects									
	Any interaction		Discussed salongo		Asked respondent to participate		Mentioned social consequences		Mentioned divine consequences	
	Non-leader	Leader	Non-leader	Leader	Non-leader	Leader	Non-leader	Leader	Non-leader	Leader
Panel A. Leader effects relative to no leader, by environment										
Traditional leader										
Mixed, without disclosure	-0.017	0.120**	-0.060	0.074	-0.030	-0.027	-0.004	0.000	-0.012	-0.003
	(0.057)	(0.052)	(0.059)	(0.065)	(0.026)	(0.025)	(0.008)	(0.015)	(0.014)	(0.016)
Mixed, with disclosure	0.015	0.208***	0.028	0.185***	-0.026	0.137**	0.014	0.025	-0.014	0.033
	(0.051)	(0.066)	(0.048)	(0.068)	(0.029)	(0.054)	(0.010)	(0.019)	(0.015)	(0.037)
Homogeneous, without disclosure	-0.059	0.145**	-0.054	0.064	-0.054**	0.023	-0.005	0.004	-0.009	0.002
	(0.053)	(0.060)	(0.046)	(0.047)	(0.027)	(0.033)	(0.011)	(0.017)	(0.015)	(0.014)
Homogeneous, with disclosure	-0.043	0.050	0.015	0.079	-0.036	0.054	0.014*	0.014	-0.017	0.022
	(0.056)	(0.065)	(0.044)	(0.057)	(0.028)	(0.043)	(0.008)	(0.013)	(0.020)	(0.022)
Religious leader										
Mixed, without disclosure	-0.032	0.065	0.002	0.122*	-0.024	0.041	-0.013	0.009	0.017	0.024
	(0.052)	(0.054)	(0.060)	(0.064)	(0.030)	(0.045)	(0.011)	(0.019)	(0.021)	(0.024)
Mixed, with disclosure	0.114**	0.073	0.107**	0.076	-0.021	-0.008	0.008	-0.005	-0.026**	-0.015
	(0.053)	(0.072)	(0.052)	(0.065)	(0.029)	(0.027)	(0.008)	(0.006)	(0.012)	(0.016)
Homogeneous, without disclosure	-0.028	0.097·	-0.075*	0.051	-0.033	0.049	-0.004	0.032	0.006	0.025
	(0.055)	(0.059)	(0.041)	(0.052)	(0.026)	(0.037)	(0.011)	(0.024)	(0.013)	(0.022)
Homogeneous, with disclosure	-0.085·	0.106·	-0.052	0.133*	-0.030	0.040	0.002	0.011	-0.021	0.017
	(0.053)	(0.073)	(0.040)	(0.071)	(0.029)	(0.045)	(0.009)	(0.011)	(0.015)	(0.022)
Panel B. No-leader environment reference means and diagnostics										
No leader, mixed, without disclosure	0.399		0.310		0.105		0.019		0.027	
	(0.040)		(0.037)		(0.026)		(0.008)		(0.013)	
No leader, mixed, with disclosure	0.313		0.257		0.104		0.004		0.044	
	(0.036)		(0.034)		(0.029)		(0.004)		(0.016)	
No leader, homogeneous, without disclosure	0.356		0.299		0.096		0.011		0.019	
	(0.037)		(0.033)		(0.021)		(0.007)		(0.007)	
No leader, homogeneous, with disclosure	0.400		0.278		0.106		0.008		0.031	
	(0.033)		(0.028)		(0.018)		(0.006)		(0.013)	
Wald: four no-leader cells equal	p = 0.259		p = 0.712		p = 0.986		p = 0.346		p = 0.515	
Observations	2,844	2,844	2,844	2,844	2,844	2,844	2,844	2,844	2,844	2,844
Groups	336	336	336	336	336	336	336	336	336	336
Neighborhoods	28	28	28	28	28	28	28	28	28	28
Baseline participation control	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: The last three outcomes record whether the referenced person asked the respondent to participate, mentioned social consequences, or mentioned divine consequences. The sample excludes respondent-leaders; each observation is a non-leader respondent evaluating a referenced person. Only composition-by-disclosure environment estimates are shown. Within each outcome block, the non-leader column reports the leader-assignment effect for participant-to-participant dyads; the leader column reports the total effect for participant-to-leader dyads. All effects are relative to dyads in no-leader groups in the same environment. Panel B averages fitted predictions over actual no-leader dyads in each environment, retaining their observed baseline participation and neighborhood composition; the equality test uses their full covariance. All regressions control for respondent i's average observed *salongo* participation before the first accountability meeting held in the neighborhood. All regressions include neighborhood fixed effects. Standard errors are clustered by randomized group within neighborhood. Significance: · $p < 0.15$, * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table D-7 – PASTOR EFFECTS BY CHIEF COETHNICITY AND DISCLOSURE

	Mixed groups		
	Baseline	+ Spillover	+ LASSO
Panel A. Pastor effects relative to no leader (percentage points)			
Not coethnic with chief			
Without disclosure	5.67	5.69	6.08
	(4.93)	(4.93)	(4.74)
With disclosure	1.15	0.95	0.79
	(4.98)	(5.00)	(4.60)
With minus without disclosure	-4.52	-4.74	-5.29
	(7.00)	(7.03)	(6.57)
Chief's coethnics			
Without disclosure	20.12*	20.00*	18.24**
	(10.38)	(10.44)	(8.86)
With disclosure	9.05	9.63	6.67
	(10.29)	(10.30)	(10.25)
With minus without disclosure	-11.07	-10.37	-11.56
	(14.59)	(14.68)	(13.52)
Difference in disclosure contrasts (coethnic minus non-coethnic)	-6.55	-5.63	-6.27
	(17.36)	(17.47)	(15.99)
Panel B. Predicted no-leader participation (percent)			
Not coethnic with chief			
Without disclosure	61.74	61.74	61.74
	(3.01)	(2.98)	(2.77)
With disclosure	56.10	56.10	56.10
	(3.26)	(3.27)	(2.83)
Chief's coethnics			
Without disclosure	49.21	49.21	49.21
	(7.04)	(7.12)	(6.35)
With disclosure	63.49	63.49	63.49
	(7.56)	(7.53)	(7.73)
Baseline participation control	Yes	Yes	Yes
Neighborhood × composition × disclosure FE	Yes	Yes	Yes
Spillover control	No	Yes	Yes
LASSO controls	No	No	Yes
Observations	1,040	1,040	1,040
Group clusters	336	336	336

Notes: The table compares pastor effects in mixed groups among residents who share the neighborhood chief's territorial origin and those who do not, separately with and without disclosure. Coethnicity refers to the chief in every assignment arm. Disclosure contrasts subtract without from with disclosure; the difference in contrasts subtracts non-coethnics from coethnics. Panel B reports no-leader levels equal to the observed subgroup means, with standard errors for uncertainty in the model. For each observed coethnicity-by-disclosure category, these levels average fitted predictions for actual no-leader respondents, retaining their own covariates and fixed effects. The regression behind these predictions and effects includes all 1,040 residents. In mixed groups, chief and pastor assignments interact fully with disclosure and coethnicity, with corresponding no-leader category terms. Homogeneous-group leader effects vary by disclosure. We display only the pastor effects within mixed groups. Among non-coethnics, the pastor and no-leader samples contain, respectively, 27/28 groups without disclosure and 28/28 with disclosure; for coethnics it is 18/18 and 18/17 groups, respectively. The regression also retains mixed groups assigned a chief: non-coethnics occur in 28 groups in each disclosure arm, and coethnics in 15 without and 19 with disclosure. We keep baseline participation and the stated fixed effects in each column, adding spillover in the second and the same 11 LASSO-selected baseline controls in the third. Parentheses give standard errors clustered by group. Stars use unadjusted exploratory tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.3.1 Pastors' Territorial Origins

Table D-8 – LEADER EFFECTS UNDER THE RELIGIOUS-LEADER TERRITORY RESTRICTION: DISCLOSURE SPECIFICATION

	Environment specific		
	Baseline	+ Spillover	+ LASSO
Panel A. Leader effects relative to no leader			
Traditional leader			
Mixed, without disclosure	-0.36	-0.38	-1.25
	(3.27)	(3.26)	(3.32)
Mixed, with disclosure	1.67	1.57	1.67
	(4.33)	(4.31)	(3.95)
Homogeneous, without disclosure	-0.24	-0.28	1.48
	(4.42)	(4.40)	(4.24)
Homogeneous, with disclosure	9.83***	10.08***	10.62***
	(3.61)	(3.60)	(3.44)
Religious leader			
Mixed, without disclosure	9.56*	9.52*	8.76*
	(4.93)	(4.90)	(4.94)
Mixed, with disclosure	10.53**	10.73**	8.17*
	(5.04)	(5.04)	(4.78)
Homogeneous, without disclosure	-0.50	-0.45	-1.20
	(4.75)	(4.73)	(4.45)
Homogeneous, with disclosure	-2.35	-2.55	-0.46
	(4.59)	(4.54)	(4.08)
Panel B. Treatment-effect contrasts			
Mixed, without disclosure: T − R	-9.93**	-9.90**	-10.01*
	(4.89)	(4.89)	(5.26)
Mixed, with disclosure: T − R	-8.86*	-9.17*	-6.51
	(4.85)	(4.85)	(4.49)
Homogeneous, without disclosure: T − R	0.26	0.17	2.68
	(5.03)	(5.01)	(4.63)
Homogeneous, with disclosure: T − R	12.18**	12.63**	11.08**
	(5.41)	(5.35)	(4.82)
Δ_E: H, with − M, without	22.11***	22.54***	21.09***
	(7.31)	(7.28)	(7.16)
Panel C. No-leader reference means and diagnostics			
Mixed, without disclosure	58.88	58.88	58.88
	(2.29)	(2.26)	(2.13)
Mixed, with disclosure	57.79	57.79	57.79
	(3.08)	(3.08)	(2.84)
Homogeneous, without disclosure	57.79	57.79	57.79
	(2.92)	(2.91)	(2.83)
Homogeneous, with disclosure	55.80	55.80	55.80
	(1.93)	(1.94)	(1.95)
Wald: four no-leader cells equal	p = 0.770	p = 0.767	p = 0.755
Observations	908	908	908
Group-neighborhood clusters	292	292	292
Neighborhoods	28	28	28
Design fixed effects	Yes	Yes	Yes
Pretreatment participation control	Yes	Yes	Yes
25m spillover control	No	Yes	Yes
LASSO-selected baseline controls	No	No	Yes

Notes: Could ethnic mismatch explain pastors' effects under either disclosure condition? We repeat the disclosure specification after excluding 132 residents assigned a pastor in the 11 neighborhoods where the pastor and chief originate from different territories. Residents assigned a chief or no leader all remain. Panel A gives leader effects separately for each composition and disclosure condition, Panel B compares chief and pastor effects, and Panel C gives predicted no-leader means. The first column adjusts for participation before treatment and neighborhood-by-composition-by-disclosure fixed effects. Moving to the second column adds nearby spillover exposure; moving to the third adds the 11 baseline controls selected by LASSO, retaining the preceding adjustments. Parentheses give standard errors clustered by group. See Section 5 for the territorial comparison. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.3.2 Leadership by Example

Table D-9 – LEADERSHIP BY EXAMPLE: DISCLOSURE SPECIFICATION

	Group-clustered SE			Neighborhood-clustered SE		
	Participation at/below leader-type median	Participation above leader-type median	Difference	Participation at/below leader-type median	Participation above leader-type median	Difference
Panel A. Leader effects relative to no leader						
Traditional leader						
Mixed, without disclosure	0.41	-0.96	-1.37	0.41	-0.96	-1.37
	(5.15)	(4.43)	(6.53)	(5.62)	(4.85)	(7.45)
Mixed, with disclosure	1.41	2.00	0.59	1.41	2.00	0.59
	(4.97)	(6.67)	(7.68)	(5.80)	(8.00)	(8.92)
Homogeneous, without disclosure	2.63	-2.80	-5.43	2.63	-2.80	-5.43
	(5.14)	(6.29)	(7.39)	(6.23)	(7.93)	(8.98)
Homogeneous, with disclosure	13.32**	7.05	-6.27	13.32*	7.05	-6.27
	(5.74)	(4.31)	(7.07)	(6.70)	(5.39)	(8.84)
Religious leader						
Mixed, without disclosure	8.07	10.60*	2.53	8.07	10.60	2.53
	(5.00)	(5.88)	(7.37)	(5.87)	(7.60)	(9.30)
Mixed, with disclosure	1.74	5.60	3.86	1.74	5.60	3.86
	(5.80)	(5.37)	(7.28)	(6.80)	(5.96)	(8.02)
Homogeneous, without disclosure	5.10	-2.84	-7.94	5.10	-2.84	-7.94
	(5.07)	(4.88)	(6.33)	(5.39)	(5.46)	(6.50)
Homogeneous, with disclosure	2.19	-2.86	-5.05	2.19	-2.86	-5.05
	(4.39)	(4.69)	(6.37)	(5.02)	(5.08)	(7.47)
Panel B. Implied no-leader counterfactuals						
No leader, mixed, without disclosure		55.52			55.52	
		(2.47)			(2.46)	
No leader, mixed, with disclosure		57.52			57.52	
		(3.13)			(3.20)	
No leader, homogeneous, without disclosure		55.43			55.43	
		(3.12)			(3.11)	
No leader, homogeneous, with disclosure		52.73			52.73	
		(1.88)			(1.94)	
Observations			1,040			
Group clusters			336			
Neighborhood clusters			28			

Notes: We divide chiefs and pastors by their participation during the common week before any meeting, then compare their effects within each disclosure environment. Rates exclude Sunday and use four to six observed eligible days. The median is 0.80 for chiefs and 0.83 for pastors. The comparison puts leaders at the median in the lower category and subtracts that category's effect from the one above it. We estimate a single disclosure-specific regression for all 1,040 participants in each block, including baseline participation and neighborhood-by-composition-by-disclosure fixed effects. Leader participation has no value for a no-leader group, so its category enters only through interactions with leader assignment, composition and disclosure; there is no standalone term. For the same reason, Panel B does not divide no-leader groups into participation categories. It uses a common no-leader reference in percentage points for both categories in each block. We predict it over the same chief- and pastor-arm residents in each composition-by-disclosure environment, setting baseline participation at the estimation-sample mean, and average.
Only the clustering changes across blocks, from 336 groups to 28 neighborhoods. The latter lets outcomes be dependent across groups assigned the same leader. Section 5 discusses the interpretation. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.3.3 Pretreatment Participation

Table D-10 – PRETREATMENT DYNAMICS

	All six pretreatment days			Excluding day 1		
Panel A. Linear trends						
	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Traditional leader						
Mixed, without disclosure	-1.83	(1.32)	0.165	-1.29	(1.60)	0.421
Mixed, with disclosure	0.83	(1.26)	0.510	0.10	(1.40)	0.945
Homogeneous, without disclosure	-0.84	(1.35)	0.533	-0.00	(1.91)	0.999
Homogeneous, with disclosure	-2.12	(1.06)	0.047	-1.73	(1.27)	0.175
Religious leader						
Mixed, without disclosure	1.20	(1.21)	0.321	1.33	(1.58)	0.400
Mixed, with disclosure	1.00	(1.15)	0.385	0.69	(1.30)	0.596
Homogeneous, without disclosure	-2.75	(1.41)	0.052	-2.03	(1.81)	0.263
Homogeneous, with disclosure	-1.68	(1.10)	0.126	-1.49	(1.59)	0.349
Panel B. Early versus late pretreatment						
	Estimate	Standard error	p-value	Estimate	Standard error	p-value
Traditional leader						
Mixed, without disclosure	-3.13	(4.19)	0.456	-0.72	(4.31)	0.868
Mixed, with disclosure	2.47	(4.06)	0.543	0.83	(4.09)	0.839
Homogeneous, without disclosure	-2.20	(4.79)	0.647	-0.29	(5.52)	0.958
Homogeneous, with disclosure	-6.25	(3.53)	0.078	-4.51	(3.99)	0.260
Religious leader						
Mixed, without disclosure	2.83	(3.79)	0.457	2.26	(4.11)	0.583
Mixed, with disclosure	1.28	(3.80)	0.736	-0.16	(3.77)	0.967
Homogeneous, without disclosure	-6.96	(4.71)	0.141	-4.06	(5.00)	0.418
Homogeneous, with disclosure	-2.68	(4.25)	0.529	-0.94	(5.08)	0.854
Panel C. Saturated day-specific paths						
	F	Restrictions	p-value	F	Restrictions	p-value
Traditional leader						
Mixed, without disclosure	2.97	5	0.012	3.39	4	0.010
Mixed, with disclosure	0.51	5	0.768	0.50	4	0.737
Homogeneous, without disclosure	0.47	5	0.801	0.20	4	0.936
Homogeneous, with disclosure	1.83	5	0.106	1.76	4	0.137
Religious leader						
Mixed, without disclosure	1.56	5	0.170	1.94	4	0.104
Mixed, with disclosure	0.42	5	0.832	0.34	4	0.850
Homogeneous, without disclosure	2.72	5	0.020	2.17	4	0.072
Homogeneous, with disclosure	1.29	5	0.270	1.35	4	0.253
Respondents		1,040			1,040	
Group–neighborhood clusters		336			336	
Neighborhoods		28			28	

Notes: Panels A and B report percentage-point estimates, clustered standard errors, and p-values for each of the eight leader-by-information-by-composition paths. Panel A estimates the change in the leader-minus-no-leader gap per campaign day. Panel B compares days 1–3 with days 4–6 in the full window and days 2–3 with days 4–6 after excluding day 1. Panel C jointly tests the day-specific coefficients for each path, taking campaign day 6 as the reference. We report the F statistic, number of restrictions, and p-value for this test. All regressions include respondent fixed effects and neighborhood × information × composition × campaign-day fixed effects. Standard errors are clustered one-way by assigned group within neighborhood.

Table D-11 – LEADER EFFECTS EXCLUDING REPLACEMENT RESPONDENTS

	Leader by composition			Environment specific	
	Full sample	No replacements		Full sample	No replacements
Panel A. Leader effects relative to no leader					
Traditional leader					
Mixed	0.23	-0.66	Without disclosure	-1.34	0.40
	(2.68)	(2.83)		(3.40)	(3.88)
			With disclosure	1.55	-1.80
				(4.03)	(3.95)
Homogeneous	5.40**	5.65*	Without disclosure	1.12	-0.47
	(2.73)	(3.15)		(4.31)	(4.90)
			With disclosure	9.79***	12.19***
				(3.25)	(3.66)
Religious leader					
Mixed	5.38*	4.72	Without disclosure	8.43**	11.87***
	(2.81)	(3.18)		(3.85)	(4.44)
			With disclosure	2.20	-1.90
				(4.00)	(4.38)
Homogeneous	1.78	4.22	Without disclosure	1.38	3.01
	(2.42)	(2.77)		(3.67)	(4.15)
			With disclosure	2.23	5.74*
				(3.07)	(3.41)
Panel B. Treatment-effect contrasts					
Mixed: T − R	-5.15*	-5.38	Mixed, with: T − R	-0.66	0.10
	(2.89)	(3.27)		(4.14)	(4.62)
			Mixed, without: T − R	-9.77**	-11.48**
				(4.02)	(4.64)
Homogeneous: T − R	3.62	1.42	Homogeneous, with: T − R	7.57**	6.46*
	(2.68)	(3.16)		(3.46)	(3.73)
			Homogeneous, without: T − R	-0.26	-3.48
				(3.98)	(4.85)
Δ_C: composition gradient	8.77**	6.80	Δ_E: H, with − M, without	17.34***	17.93***
	(3.98)	(4.58)		(5.31)	(5.94)
Panel C. No-leader reference means and diagnostics					
Mixed	58.33	57.71	Without disclosure	58.88	58.33
	(1.78)	(1.75)		(2.23)	(2.45)
			With disclosure	57.79	57.11
				(2.80)	(2.41)
Homogeneous	56.79	56.01	Without disclosure	57.79	56.58
	(1.72)	(1.79)		(2.77)	(2.69)
			With disclosure	55.80	55.40
				(2.01)	(2.22)
Wald: all no-leader cells equal	p = 0.534	p = 0.497		p = 0.773	p = 0.847
Panel D. Joint tests					
All composition coefficients zero	0.084	0.141	All environment coefficients zero	0.038	0.011
Observations	1,040	809		1,040	809
Group–neighborhood clusters	336	331		336	331
Neighborhoods	28	28		28	28
Neighborhood × disclosure × composition FEs	Yes	Yes		Yes	Yes
Assignment-cell baseline slopes	Yes	Yes		Yes	Yes
25m spillover control	Yes	Yes		Yes	Yes
11 fixed baseline controls	Yes	Yes		Yes	Yes

Notes: Do the leader effects depend on residents who entered through replacement queues? Each block compares all 1,040 residents with the 809 initially assigned residents, excluding the 231 whose final roster identifier differs from the original assignment. Panels A–D report leader effects in percentage points, chief-minus-pastor contrasts, no-leader means and equality tests, and joint tests of the leader effects. The full-sample columns reproduce the preferred specification in Table C-11. All columns retain the same 11 baseline controls selected for primary participation, their full-sample missing-value replacements, and the 25-meter spillover indicator; controls are not reselected. Pretreatment participation is centered at each estimation sample's mean and interacted with all 12 assignment cells. Neighborhood-by-disclosure-by-composition fixed effects are included. No-leader means average fitted predictions over actual no-leader residents within each sample and category. $\Delta_C = (\tau^T_H - \tau^T_M) - (\tau^R_H - \tau^R_M)$ and $\Delta_E = (\tau^T_{HI} - \tau^T_{MN}) - (\tau^R_{HI} - \tau^R_{MN})$. Standard errors in parentheses are clustered by accountability group within neighborhood. Section 5 discusses the comparison. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.4 Responses to Revealed Participation Histories

Table D-12 – LEADER EFFECTS BY DAYS OF PRETREATMENT PARTICIPATION

	0–1 days	2–6 days	Difference
Panel A. Leader effects relative to no leader			
Traditional leader			
Mixed, without disclosure	-7.04	4.38	11.42
	(5.93)	(5.26)	(8.91)
Mixed, with disclosure	7.84	-5.28	-13.12
	(7.01)	(5.05)	(9.46)
Homogeneous, without disclosure	-5.82	7.87	13.69
	(7.35)	(5.36)	(9.37)
Homogeneous, with disclosure	19.03***	1.72	-17.31**
	(6.06)	(4.15)	(7.93)
Religious leader			
Mixed, without disclosure	6.31	9.43	3.12
	(5.89)	(5.85)	(8.93)
Mixed, with disclosure	4.75	-2.27	-7.02
	(7.29)	(4.55)	(9.07)
Homogeneous, without disclosure	-2.45	6.14	8.59
	(6.83)	(4.99)	(9.43)
Homogeneous, with disclosure	0.81	4.20	3.39
	(5.19)	(4.31)	(7.07)
	0–1 days	2–6 days	Difference
Panel B. Implied no-leader counterfactuals			
Traditional leader			
Mixed, without disclosure	38.75	74.36	35.61***
	(3.80)	(3.74)	(6.05)
Mixed, with disclosure	37.76	80.62	42.87***
	(4.79)	(3.30)	(6.23)
Homogeneous, without disclosure	39.47	70.68	31.21***
	(5.39)	(3.65)	(6.98)
Homogeneous, with disclosure	25.88	76.85	50.97***
	(3.67)	(2.68)	(4.66)
Religious leader			
Mixed, without disclosure	38.75	74.36	35.61***
	(3.80)	(3.74)	(6.05)
Mixed, with disclosure	37.76	80.62	42.87***
	(4.79)	(3.30)	(6.23)
Homogeneous, without disclosure	39.47	70.68	31.21***
	(5.39)	(3.65)	(6.98)
Homogeneous, with disclosure	25.88	76.85	50.97***
	(3.67)	(2.68)	(4.66)
Observations		1,040	
Group clusters		336	

Notes: In mixed groups, the disclosure change in the implied no-leader comparison is -0.99 points (SE 6.11) for residents in the 0–1-day category and 6.26 (SE 4.99) for those in the 2–6-day category. The pastor effect changes by -1.56 (SE 9.21) and -11.71 (SE 7.42), respectively.
We distinguish prior nonparticipants and infrequent participants, who worked on 0–1 eligible days before any meeting, from stronger contributors, who worked on 2–6. In comparing their responses to a leader, we take the stronger contributors' effect minus the weaker contributors' effect. These are the estimates in Figure 8, discussed in Section 5.3.
The no-leader level is a standardized prediction for the residents who support each displayed leader effect. We set centered baseline participation, spillover and LASSO controls to zero, then average fixed-effect and category components over those treated residents. There are 1,040 residents in 336 groups. We include baseline participation and neighborhood-by-composition-by-disclosure fixed effects, together with 25-meter spillover exposure and the 11 LASSO-selected baseline controls. Each environment has separate pretreatment-category levels and interactions with leader assignment. Standard errors cluster by group. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

D.5 Salongo Contributions and Leadership Alignment

Recorded contributions. We measure labor from recorded arrival and departure times, summing valid intervals without double-counting overlaps for all recorded attendees in the 28 treatment neighborhoods, Waves 1–7 (August 2025–April 2026), including nonexperimental residents. The 70,250 hours equal 8,781 eight-hour days. Cash contributions total 154,250 CDF. We value in-kind contributions, including materials and refreshments, at their recorded unit prices, giving 3,324,417 CDF. We do not count cash purchases again as in-kind contributions. To value labor, we use the historical STEP II wage in Kananga, \$3/day (Agence Congolaise de Presse, 2021), with the eight-hour day used in the earlier STEP evaluation (Brandily et al., 2025). The resulting \$0.375/hour values labor at \$26,344. We convert cash and goods at 2,250 CDF/USD, bringing contributions to \$27,890. We compare this total with the 205,682,923 CDF in provincial property-tax receipts reported for 2025 (Direction Générale des Recettes du Kasaï Central, 2026), or \$91,415, without adjusting for inflation across the different periods.

Worksites. We observe GPS locations in 20 neighborhoods during Waves 1–5. We keep separately recorded sites on the same day distinct and pool observations from different days within a neighborhood only when their maximum pairwise distance is no more than five meters. We obtain 596 GPS worksite histories (Figure D-3). We sum labor across visits, sharing session labor equally over the sites separately recorded for that session. The count for Waves 6–7 requires an imputation, since there are no GPS data. Each neighborhood contributes one worksite per non-Sunday wave workday: 30 days for each of four Wave 6 neighborhoods and 29 for each of four Wave 7 neighborhoods. The addition is $4 \times 30 + 4 \times 29 = 236$, taking 596 histories to an estimated 832 worksites. This includes four neighborhood-days for which no attendance was recorded. The count combines mapped histories and imputed neighborhood-workday counts; it does not establish how many distinct projects were completed.

We calculate the workload benchmark from positive-effort person-hours in the GPS histories, using all recorded attendees. We accept check-in/check-out intervals in 2025–2026 that run from zero to 12 hours, counting overlapping intervals once and excluding spells rated as no effort. We remove 23 histories last observed on their neighborhood's final GPS-recorded day, five with unusable or unlinked labor, and two with no positive labor. The 566 remaining histories sum to 54,735.3375 positive-effort person-hours:

$$\frac{54{,}735.3375}{566} \approx 96.7 \quad \text{person-hours per history.}$$

We express additional hours in worksite workload equivalents using this observed mean, rather than treating it as the hours needed to complete a project. Table D-13 reapplies the construction at 1, 5, 10, 20 and 50 meters, with policy estimates and projected additional hours fixed.

Alignment projection. What would happen if homogeneous teams had a chief and public disclosure, and mixed teams a pastor without disclosure? For this projection, we predict six-day hours for all 1,040 residents (520 per composition), retaining their own covariates and using the fixed effect that corresponds to the assigned disclosure condition. The estimated regression includes matched baseline hours, neighborhood-by-composition-by-disclosure fixed effects, spillover exposure and the same 11 LASSO controls chosen for participation. We extend these predictions using the ratio of mean full-period no-leader hours to mean six-day no-leader hours within each composition–disclosure cell. Multiplying by this ratio assumes that the effects operate throughout the 33–34 days of a wave, including the days before the first meeting. We leave unobserved blackout days unimputed.

The gain from alignment is 7,251 hours: 53,240 projected with aligned leaders less 45,989 without leaders, keeping the disclosure rules the same. Dividing by the unrounded mean above gives 75.0 worksite workload equivalents. We obtain a conditional 95% interval of 2,424–12,078 additional hours with clustering by group. This interval does not include uncertainty about policy selection, persistence of effects or conversion into worksite workloads. The projection also omits leader supply, additional costs and changes in spillovers.

Table D-13 – Worksite Workload Benchmarks by Distance Cutoff

Distance cutoff (m)	Selected histories	Allocated person-hours	Mean person-hours	Workload equivalents
1	666	55,564.47	83.43	86.9
5	566	54,735.34	96.71	75.0
10	460	53,506.31	116.32	62.3
20	353	50,006.21	141.66	51.2
50	243	44,159.90	181.73	39.9

Notes: We repeat the spatial grouping within the same 20 neighborhoods in Waves 1–5, changing only the maximum distance between observations assigned to a worksite history. Complete-linkage grouping, separation of same-day sites, equal allocation of session labor across recorded sites, and the positive-labor and final-GPS-day eligibility rules are unchanged. We reapply eligibility at each cutoff, so the retained histories and allocated person-hours can change. Workload equivalents divide the same projected gain of 7,251.124 person-hours by each unrounded mean. These are sensitivities of the descriptive conversion; they do not count verified completed projects or change the policy treatment-effect estimates.

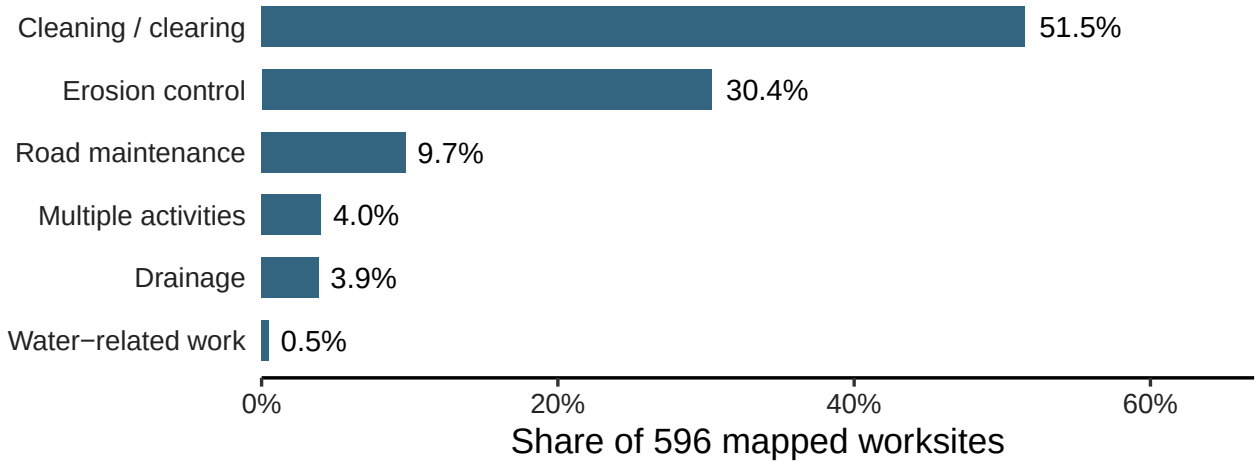


Figure D-3 – ACTIVITIES AT OBSERVED CAMPAIGN WORKSITES

Notes: Shares of the 596 location-based worksite histories in the 20 neighborhoods with GPS records (Waves 1–5). Revisits are pooled using a maximum five-meter distance, as described above. Multiple activities denotes histories with more than one recorded activity type. The counts describe observed worksites, not independently verified completed projects.

E Beyond Participation

E.1 Cooperation Beliefs, Identity, and Governance

Table E-1 – LEADER EFFECTS ON COOPERATION NORMS AND DIVINE MONITORING BELIEFS

Panel A. Evolution in the two environments where leaders increased participation

	Traditional leader: homogeneous, with disclosure				Religious leader: mixed, without disclosure			
Outcome	Meeting 1	Meeting 2	Endline	Pooled	Meeting 1	Meeting 2	Endline	Pooled
Active participants are well regarded	0.01	-0.11	-0.01	-0.03	-0.06	-0.10	-0.01	-0.06
	(0.08)	(0.10)	(0.10)	(0.07)	(0.09)	(0.10)	(0.12)	(0.08)
Contributing is a moral obligation	-0.09	-0.11	0.01	-0.06	-0.09	-0.00	0.04	-0.02
	(0.07)	(0.08)	(0.10)	(0.07)	(0.08)	(0.09)	(0.09)	(0.06)
God rewards community support	-0.03	-0.01	0.03	0.00	-0.09	-0.05	0.02	-0.04
	(0.06)	(0.08)	(0.07)	(0.06)	(0.06)	(0.07)	(0.08)	(0.05)
God punishes neglect of community obligations	-0.14	-0.09	-0.15	-0.13	-0.21	-0.05	-0.06	-0.11
	(0.15)	(0.14)	(0.10)	(0.09)	(0.15)	(0.13)	(0.13)	(0.10)
Family p-value	0.889	0.591	0.573	0.677	0.431	1.000	1.000	1.000

Panel B. Pooled effects in every leader environment

Traditional leader	Well regarded	Moral obligation	God rewards	God punishes	Family p-value
Mixed, without disclosure	0.04	-0.05	-0.04	-0.13	0.610
	(0.08)	(0.07)	(0.05)	(0.09)	
Mixed, with disclosure	0.14*	0.12**	0.08	0.21**	0.149
	(0.08)	(0.06)	(0.05)	(0.10)	
Homogeneous, without disclosure	-0.04	0.05	-0.05	0.02	1.000
	(0.07)	(0.08)	(0.05)	(0.08)	
Homogeneous, with disclosure	-0.03	-0.06	0.00	-0.13	0.677
	(0.07)	(0.07)	(0.06)	(0.09)	
Religious leader					
Mixed, without disclosure	-0.06	-0.02	-0.04	-0.11	1.000
	(0.08)	(0.06)	(0.05)	(0.10)	
Mixed, with disclosure	-0.06	0.04	-0.03	0.13	1.000
	(0.09)	(0.07)	(0.05)	(0.12)	
Homogeneous, without disclosure	-0.07	0.10	0.00	0.07	0.565
	(0.06)	(0.07)	(0.05)	(0.09)	
Homogeneous, with disclosure	0.01	0.00	0.02	-0.02	1.000
	(0.08)	(0.07)	(0.06)	(0.10)	

Notes: The four outcomes concern norms of cooperation, with answers from 1 (strongly disagree) to 5 (strongly agree). We code non-substantive answers as missing. Each leader effect is measured against no leader in the same composition-by-disclosure environment. Residents answered at Meeting 1, immediately after the first accountability meeting, at Meeting 2, and at Endline after recognition. Counterfactual means for both panels appear on the continuation page.
Separate regressions for each survey control for that outcome's baseline response, gender, and assignment wealth stratum, with neighborhood × disclosure × composition fixed effects. We also pool the three surveys. The pooled model includes survey-wave × neighborhood × disclosure × composition fixed effects and lets baseline-response, gender, and wealth-stratum slopes differ by wave. Parentheses contain standard errors clustered by group-neighborhood, with the main paper's finite-sample correction. Family p-values use the Bonferroni minimum-p test over the four displayed outcomes within each leader-environment and survey wave. Item availability and don't-know answers give 983 observations at Meeting 1, 943–946 at Meeting 2, 965 at Endline, and 2891–2894 in the pooled regressions. Significance: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$ (unadjusted).

Table E-1 – LEADER EFFECTS ON COOPERATION NORMS AND DIVINE MONITORING BELIEFS (continued)

Panel C. No-leader counterfactual means for the evolution in Panel A

	Homogeneous, with disclosure				Mixed, without disclosure			
Outcome	Meeting 1	Meeting 2	Endline	Pooled	Meeting 1	Meeting 2	Endline	Pooled
Well regarded	4.45	4.48	4.45	4.46	4.43	4.42	4.41	4.42
	(0.06)	(0.06)	(0.07)	(0.05)	(0.06)	(0.06)	(0.07)	(0.04)
Moral obligation	4.59	4.52	4.46	4.52	4.41	4.51	4.56	4.49
	(0.05)	(0.05)	(0.07)	(0.05)	(0.06)	(0.06)	(0.07)	(0.05)
God rewards	4.64	4.55	4.63	4.61	4.57	4.55	4.57	4.56
	(0.04)	(0.06)	(0.05)	(0.04)	(0.04)	(0.05)	(0.06)	(0.04)
God punishes	4.14	4.05	4.22	4.14	4.19	4.29	4.30	4.26
	(0.11)	(0.10)	(0.07)	(0.07)	(0.08)	(0.09)	(0.08)	(0.06)

Panel D. Pooled no-leader counterfactual means for Panel B

Environment	Well regarded	Moral obligation	God rewards	God punishes
Mixed, without disclosure	4.42	4.49	4.56	4.26
	(0.04)	(0.05)	(0.04)	(0.06)
Mixed, with disclosure	4.39	4.41	4.58	4.05
	(0.06)	(0.05)	(0.04)	(0.08)
Homogeneous, without disclosure	4.52	4.35	4.54	4.13
	(0.04)	(0.05)	(0.03)	(0.05)
Homogeneous, with disclosure	4.46	4.52	4.61	4.14
	(0.05)	(0.05)	(0.04)	(0.07)

Notes: The means use the same outcomes, estimation samples, controls, and fixed effects as the corresponding treatment estimates on the preceding page. They average model predictions over no-leader observations in each displayed environment and survey window, retaining those observations' covariates. The pooled means use the stacked survey sample. Both leader comparisons in Panel B use the same environment-specific reference. Parentheses report group-clustered standard errors.

Table E-2 – LEADER EFFECTS ON IDENTITY SALIENCE

Panel A. Evolution in the two environments where leaders increased participation

	Traditional leader: homogeneous, with disclosure				Religious leader: mixed, without disclosure			
Identity rank	Meeting 1	Meeting 2	Endline	Pooled	Meeting 1	Meeting 2	Endline	Pooled
Congolese	0.45*	0.20	0.05	0.23	-0.09	-0.26	0.06	-0.10
	(0.25)	(0.23)	(0.18)	(0.14)	(0.19)	(0.23)	(0.22)	(0.18)
Kasaien	0.09	-0.20	0.02	-0.03	0.06	0.18	0.21	0.15
	(0.23)	(0.20)	(0.21)	(0.14)	(0.19)	(0.24)	(0.20)	(0.15)
Christian	0.02	0.44**	-0.05	0.13	0.28	0.28	0.59**	0.38*
	(0.24)	(0.22)	(0.24)	(0.18)	(0.26)	(0.26)	(0.26)	(0.20)
Own territory	0.26	-0.05	0.45***	0.23*	-0.12	-0.15	-0.25	-0.17
	(0.18)	(0.20)	(0.17)	(0.13)	(0.23)	(0.22)	(0.21)	(0.16)
Own denomination	-0.31	-0.26	-0.14	-0.23	-0.27	-0.30	-0.20	-0.26*
	(0.24)	(0.24)	(0.22)	(0.19)	(0.21)	(0.20)	(0.18)	(0.14)
Own tribe	-0.49*	-0.12	-0.34	-0.32*	0.13	0.25	-0.42*	-0.01
	(0.26)	(0.23)	(0.21)	(0.18)	(0.19)	(0.21)	(0.23)	(0.16)
Family p-value	0.354	0.274	0.045	0.487	1.000	0.870	0.140	0.331

Panel B. Pooled effects in every leader environment

	Congolese	Kasaien	Christian	Own territory	Own denomination	Own tribe	Family p-value
Traditional leader							
Mixed, without disclosure	0.00	0.14	-0.11	-0.21	0.11	0.07	0.872
	(0.16)	(0.15)	(0.16)	(0.14)	(0.15)	(0.15)	
Mixed, with disclosure	0.17	0.20	-0.19	-0.02	-0.31**	0.16	0.258
	(0.12)	(0.14)	(0.15)	(0.16)	(0.15)	(0.13)	
Homogeneous, without disclosure	0.11	-0.05	0.07	-0.03	-0.19	0.13	1.000
	(0.16)	(0.11)	(0.18)	(0.12)	(0.15)	(0.13)	
Homogeneous, with disclosure	0.23	-0.03	0.13	0.23*	-0.23	-0.32*	0.487
	(0.14)	(0.14)	(0.18)	(0.13)	(0.19)	(0.18)	
Religious leader							
Mixed, without disclosure	-0.10	0.15	0.38*	-0.17	-0.26*	-0.01	0.331
	(0.18)	(0.15)	(0.20)	(0.16)	(0.14)	(0.16)	
Mixed, with disclosure	-0.05	-0.11	-0.02	0.18	0.02	-0.03	1.000
	(0.13)	(0.15)	(0.15)	(0.14)	(0.16)	(0.14)	
Homogeneous, without disclosure	0.06	0.04	0.15	-0.21*	-0.33**	0.25**	0.208
	(0.14)	(0.12)	(0.17)	(0.12)	(0.15)	(0.12)	
Homogeneous, with disclosure	0.07	0.10	0.30	0.06	-0.19	-0.34**	0.242
	(0.15)	(0.14)	(0.20)	(0.14)	(0.17)	(0.17)	

Notes: The identity questions require a ranking of six identities, so the ranks necessarily add to 21. We report each separately and no aggregate index, reversing ranks to make higher values indicate greater salience. We estimate the effect of leader assignment relative to no leader within each composition-by-disclosure environment at three points: immediately following the first accountability meeting (Meeting 1), at Meeting 2, and following recognition (Endline). The continuation page reports counterfactual means for both panels.
Separate-survey estimates adjust for the same identity's baseline rank, gender, and assignment wealth stratum, with neighborhood × disclosure × composition fixed effects. When pooling all three surveys, we let these slopes vary by survey wave and use survey-wave × neighborhood × disclosure × composition fixed effects. Standard errors in parentheses cluster at group–neighborhood, using the main paper's finite-sample correction. Bonferroni's minimum-p family test covers the six ranks within a displayed leader-environment and survey wave. There are 983 residents at Meeting 1, 947 at Meeting 2, and 961 at Endline, or 2,891 pooled observations. Stars use unadjusted tests: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table E-2 – LEADER EFFECTS ON IDENTITY SALIENCE (continued)

Panel C. No-leader counterfactual means for the evolution in Panel A

	Homogeneous, with disclosure				Mixed, without disclosure			
Outcome	Meeting 1	Meeting 2	Endline	Pooled	Meeting 1	Meeting 2	Endline	Pooled
Congolese	3.80	4.25	4.21	4.09	4.27	4.38	4.35	4.33
	(0.17)	(0.15)	(0.13)	(0.10)	(0.14)	(0.15)	(0.14)	(0.12)
Kasaien	3.25	3.89	3.89	3.68	3.53	3.65	3.93	3.70
	(0.13)	(0.13)	(0.13)	(0.09)	(0.13)	(0.13)	(0.14)	(0.11)
Christian	4.04	3.58	3.66	3.75	3.74	3.71	3.28	3.58
	(0.17)	(0.16)	(0.16)	(0.14)	(0.16)	(0.18)	(0.17)	(0.12)
Own territory	3.19	3.18	2.90	3.09	3.18	3.50	3.57	3.41
	(0.11)	(0.12)	(0.12)	(0.08)	(0.14)	(0.15)	(0.14)	(0.11)
Own denomination	3.08	2.95	2.93	2.99	3.12	2.84	2.64	2.87
	(0.17)	(0.17)	(0.14)	(0.13)	(0.16)	(0.14)	(0.12)	(0.10)
Own tribe	3.64	3.15	3.40	3.40	3.15	2.92	3.23	3.10
	(0.18)	(0.14)	(0.16)	(0.13)	(0.13)	(0.12)	(0.15)	(0.10)

Panel D. Pooled no-leader counterfactual means for Panel B

Environment	Congolese	Kasaien	Christian	Own territory	Own denomination	Own tribe
Mixed, without disclosure	4.33	3.70	3.58	3.41	2.87	3.10
	(0.12)	(0.11)	(0.12)	(0.11)	(0.10)	(0.10)
Mixed, with disclosure	4.23	3.71	3.90	3.25	2.82	3.10
	(0.08)	(0.09)	(0.11)	(0.10)	(0.11)	(0.09)
Homogeneous, without disclosure	4.29	3.78	3.76	3.47	2.86	2.83
	(0.09)	(0.07)	(0.13)	(0.07)	(0.11)	(0.09)
Homogeneous, with disclosure	4.09	3.68	3.75	3.09	2.99	3.40
	(0.10)	(0.09)	(0.14)	(0.08)	(0.13)	(0.13)

Notes: The means use the same outcomes, estimation samples, controls, and fixed effects as the corresponding treatment estimates on the preceding page. They average model predictions over no-leader observations in each displayed environment and survey window, retaining those observations' covariates. The pooled means use the stacked survey sample. Both leader comparisons in Panel B use the same environment-specific reference. Parentheses report group-clustered standard errors.

Table E-3 – GOVERNANCE AND INSTITUTIONAL OUTCOMES: TESTS BY OUTCOME FAMILY

		Smallest adjusted q in the two mobilizing environments		Rejections at $q < 0.10$	
Outcome family	Outcomes	Chief: homogeneous, disclosure	Pastor: mixed, no disclosure	All leader effects	All contrasts
Government service performance	5	0.858	0.858	0/60	0/40
Government accountability and knowledge	4	0.849	0.608	0/48	0/32
Performance: public and community authorities	4	0.901	0.901	0/48	0/32
Performance: other authorities	4	0.638	0.638	0/48	0/32
Expected assistance from the locality chief	3	0.812	0.686	0/36	0/24
Local chief service performance	5	0.561	0.999	0/60	0/40
Local chief accountability and capture	2	0.664	0.664	0/24	0/16
Trust in public and local authorities	5	0.304	0.304	0/60	0/40
Trust in religious and civil-society actors	4	0.641	0.641	0/48	0/32
All governance families	36	0.304	0.304	0/432	0/288

Notes: For the 36 outcomes measured at baseline and endline, we keep nine outcome families. There are 720 tests in this battery, and none survives adjustment at $q < 0.10$. We report the smallest Benjamini–Hochberg adjusted q-value in each mobilizing treatment cell and outcome family. All leader effects enter the adjustment, as do all treatment-effect contrasts, but the nine thematic families remain separate and effects are adjusted separately from contrasts. In the last two columns we count adjusted rejections and divide by tests across all leader effects and all contrasts, respectively.
For every ANCOVA, the baseline value of the outcome is retained along with gender and wealth stratum. We absorb neighborhood × disclosure × composition fixed effects and use the main paper's finite-sample correction when clustering standard errors by randomized group–neighborhood.
We recode appeal frequency monotonically and make higher institutional-performance answers mean better performance. Refusals to answer questions about knowledge of names are treated as missing. The questions on predetermined ties to chiefs are left out, as are mechanically complementary allocations in the diversion questions and provider categories that cannot be ordered.

E.2 Relationships with Peers and Leaders

Table E-4 – ENDPOINT CHANGES IN OPINIONS ABOUT ORDINARY PEERS

	Inspires respondent	Moral example	Enforcement authority	Trusts responsibility	Would help target	Expected help from target
Panel A. Leader effects by composition						
Traditional leader						
Mixed	0.04	0.11	0.09	0.14	0.12	0.02
	(0.09)	(0.09)	(0.08)	(0.10)	(0.09)	(0.09)
Homogeneous	0.14	0.08	0.04	-0.09	0.00	0.11
	(0.08)	(0.08)	(0.08)	(0.09)	(0.07)	(0.09)
Religious leader						
Mixed	0.20*	0.05	0.08	0.12	0.24**	0.05
	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)	(0.10)
Homogeneous	0.05	0.08	0.09	-0.07	0.05	-0.00
	(0.09)	(0.08)	(0.08)	(0.10)	(0.07)	(0.08)
Panel B. Environment-specific leader effects						
Traditional leader						
Mixed, without disclosure	0.14	0.27	0.14	0.08	0.18	0.13
	(0.12)	(0.12)	(0.10)	(0.14)	(0.14)	(0.12)
Mixed, with disclosure	-0.07	-0.06	0.04	0.20	0.04	-0.10
	(0.12)	(0.12)	(0.14)	(0.14)	(0.12)	(0.13)
Homogeneous, without disclosure	0.10	0.05	0.07	-0.16	-0.21	-0.11
	(0.12)	(0.11)	(0.12)	(0.13)	(0.10)	(0.12)
Homogeneous, with disclosure	0.17	0.10	0.01	-0.02	0.21**	0.32**
	(0.11)	(0.11)	(0.10)	(0.12)	(0.08)	(0.13)
Religious leader						
Mixed, without disclosure	0.25	0.06	0.16	0.20	0.43**	0.15
	(0.12)	(0.10)	(0.10)	(0.13)	(0.15)	(0.14)
Mixed, with disclosure	0.15	0.04	0.01	0.03	0.06	-0.04
	(0.13)	(0.13)	(0.14)	(0.13)	(0.09)	(0.14)
Homogeneous, without disclosure	0.13	0.11	0.12	-0.11	-0.02	-0.13
	(0.12)	(0.11)	(0.11)	(0.14)	(0.10)	(0.12)
Homogeneous, with disclosure	-0.04	0.05	0.06	-0.03	0.11	0.13
	(0.12)	(0.11)	(0.12)	(0.14)	(0.10)	(0.11)
Panel C. No-leader counterfactual means						
Mixed, pooled across disclosure	0.21	0.21	0.01	0.00	-0.14	-0.08
	(0.06)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)
Homogeneous, pooled across disclosure	0.11	0.13	-0.04	-0.02	-0.09	-0.08
	(0.05)	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)
Mixed, without disclosure	0.12	0.13	0.03	-0.06	-0.21	-0.13
	(0.07)	(0.06)	(0.06)	(0.08)	(0.11)	(0.09)
Mixed, with disclosure	0.29	0.29	-0.02	0.07	-0.06	-0.02
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.09)
Homogeneous, without disclosure	0.08	0.12	-0.07	0.03	-0.01	-0.00
	(0.07)	(0.06)	(0.07)	(0.08)	(0.06)	(0.07)
Homogeneous, with disclosure	0.15	0.15	-0.01	-0.07	-0.18	-0.17
	(0.07)	(0.06)	(0.07)	(0.08)	(0.06)	(0.08)
Respondent–target pairs	2,632	2,631	2,633	2,633	2,633	2,633
Ordinary-target pairs	2,071	2,071	2,072	2,072	2,072	2,072
Assigned-leader-target pairs	561	560	561	561	561	561
Respondents	885	885	885	885	885	885
Group-neighborhood clusters	323	323	323	323	323	323
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes	Yes
Baseline outcome control	No	No	No	No	No	No

Notes: We follow opinions about the same ordinary teammate across the two meetings. Residents first answered before discussion at the start of Meeting 1, then again at the start of Meeting 2. We subtract the earlier response from the later response for each respondent–teammate pair and divide by the baseline standard deviation. For both help amounts, we apply the inverse hyperbolic sine before taking this change and standardizing by the baseline standard deviation. The CDF magnitudes in the main text invert that transformation at the baseline mean; they do not estimate the average treatment effect in CDF.

Panel A pools disclosure and compares mixed with homogeneous groups. Panel B estimates the four environments separately. For Figure 10, we pool the other three environments separately for each leader type, keeping the environment where that leader increased participation separate. This differs from Panel B and gives different coefficients. Opinions of ordinary teammates assigned no leader provide the comparison throughout, under the same experimental conditions. Panel C gives predicted changes for these ordinary teammates in baseline-standard-deviation units, averaging predictions over matched respondent–teammate pairs in each composition or environment. The reference is the same for ordinary-peer and assigned-leader effects; Table E-5 uses this model to report opinions of assigned leaders.

Parentheses give standard errors clustered by randomized group. Willingness-to-help regressions include exposure within 25 meters; outcome-specific LASSO selected no further baseline controls. Stars use Benjamini–Hochberg adjusted p-values across the six ordinary-peer outcomes for each treatment comparison: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. See Section 6; Figure 10 highlights the two help outcomes.

Table E-5 – ENDPOINT CHANGES IN OPINIONS ABOUT ASSIGNED LEADERS

	Inspires respondent	Moral example	Enforcement authority	Trusts responsibility	Would help target	Expected help from target
Panel A. Leader effects by composition						
Traditional leader						
Mixed	0.05	0.02	-0.02	0.19	0.08	0.07
	(0.11)	(0.10)	(0.08)	(0.10)	(0.12)	(0.10)
Homogeneous	0.03	-0.02	-0.04	0.12	0.16	0.07
	(0.08)	(0.08)	(0.08)	(0.10)	(0.08)	(0.08)
Religious leader						
Mixed	0.07	0.08	0.09	0.23*	0.32***	0.17
	(0.09)	(0.08)	(0.09)	(0.10)	(0.10)	(0.10)
Homogeneous	0.11	0.00	0.08	0.16	0.04	0.02
	(0.09)	(0.08)	(0.09)	(0.11)	(0.06)	(0.08)
Panel B. Environment-specific leader effects						
Traditional leader						
Mixed, without disclosure	0.18	0.09	0.03	0.32	0.28	0.12
	(0.15)	(0.16)	(0.12)	(0.15)	(0.16)	(0.14)
Mixed, with disclosure	-0.09	-0.05	-0.09	0.06	-0.14	0.01
	(0.14)	(0.13)	(0.10)	(0.13)	(0.17)	(0.14)
Homogeneous, without disclosure	0.15	-0.04	0.13	0.14	0.18	-0.02
	(0.11)	(0.11)	(0.11)	(0.14)	(0.14)	(0.11)
Homogeneous, with disclosure	-0.08	-0.01	-0.19	0.11	0.15	0.16
	(0.11)	(0.11)	(0.12)	(0.15)	(0.10)	(0.13)
Religious leader						
Mixed, without disclosure	0.15	0.10	0.14	0.22	0.42**	0.21
	(0.11)	(0.09)	(0.13)	(0.13)	(0.16)	(0.15)
Mixed, with disclosure	-0.01	0.06	0.03	0.24	0.22	0.14
	(0.15)	(0.13)	(0.13)	(0.16)	(0.12)	(0.13)
Homogeneous, without disclosure	0.16	0.05	0.14	0.03	-0.06	-0.10
	(0.12)	(0.10)	(0.14)	(0.15)	(0.10)	(0.10)
Homogeneous, with disclosure	0.06	-0.04	0.01	0.29	0.15	0.14
	(0.13)	(0.12)	(0.13)	(0.15)	(0.07)	(0.14)
Panel C. No-leader counterfactual means						
Mixed, pooled across disclosure	0.21	0.21	0.01	0.00	-0.14	-0.08
	(0.06)	(0.05)	(0.05)	(0.05)	(0.07)	(0.06)
Homogeneous, pooled across disclosure	0.11	0.13	-0.04	-0.02	-0.09	-0.08
	(0.05)	(0.04)	(0.05)	(0.06)	(0.04)	(0.05)
Mixed, without disclosure	0.12	0.13	0.03	-0.06	-0.21	-0.13
	(0.07)	(0.06)	(0.06)	(0.08)	(0.11)	(0.09)
Mixed, with disclosure	0.29	0.29	-0.02	0.07	-0.06	-0.02
	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.09)
Homogeneous, without disclosure	0.08	0.12	-0.07	0.03	-0.01	-0.00
	(0.07)	(0.06)	(0.07)	(0.08)	(0.06)	(0.07)
Homogeneous, with disclosure	0.15	0.15	-0.01	-0.07	-0.18	-0.17
	(0.07)	(0.06)	(0.07)	(0.08)	(0.06)	(0.08)
Respondent–target pairs	2,632	2,631	2,633	2,633	2,633	2,633
Ordinary-target pairs	2,071	2,071	2,072	2,072	2,072	2,072
Assigned-leader-target pairs	561	560	561	561	561	561
Respondents	885	885	885	885	885	885
Group-neighborhood clusters	323	323	323	323	323	323
Neighborhood × disclosure × composition FEs	Yes	Yes	Yes	Yes	Yes	Yes
Baseline outcome control	No	No	No	No	No	No

Notes: The leader can also be the person being evaluated. We compare a resident's opinion of the assigned chief or pastor at the start of Meeting 1, before discussion, with the opinion at the start of Meeting 2. The outcome is the later answer minus the earlier answer, expressed in baseline standard deviations. For both help amounts, we first apply the inverse hyperbolic sine, then calculate the change and standardize by the baseline standard deviation. The main text illustrates CDF magnitudes by applying the inverse transformation at the baseline mean; this does not estimate the average treatment effect in CDF.

Panel A compares mixed with homogeneous groups, pooling disclosure, while Panel B reports each of the four environments. Figure 10 pools the other three environments separately for each leader type, explaining why its coefficients differ. The reference for every effect is an opinion of an ordinary teammate in a no-leader group under the same experimental conditions. To make opinions of peers and leaders comparable, both use the ordinary teammates assigned no leader as their reference. Panel C reports the predicted change for these teammates in baseline standard deviations: for each composition or environment, we average the predictions for matched respondent–teammate pairs. Table E-4 reports the ordinary-teammate changes from the same model.

Standard errors in parentheses are clustered by randomized group. Willingness-to-help regressions include the 25-meter spillover indicator; outcome-specific LASSO selected no additional baseline controls. Stars use Benjamini–Hochberg adjusted p-values over the six assigned-leader outcomes for each treatment comparison: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. Section 6 discusses the table, and Figure 10 highlights the two help outcomes.