

# Coethnic Bias and Wartime Informing

---

**Jason Lyall**, Yale University

**Yuki Shiraito**, Princeton University

**Kosuke Imai**, Princeton University

Information about insurgent groups is a central resource in civil wars: counterinsurgents seek it, insurgents safeguard it, and civilians often trade it. Yet despite its essential role in civil war dynamics, the act of informing is still poorly understood, due mostly to the classified nature of informant “tips.” As an alternative research strategy, we use an original 2,700 respondent survey experiment in 100 villages to examine attitudes toward the Guardians of Peace program, a widespread campaign by the International Security Assistance Force in Afghanistan to recruit local informants. We find that coethnic bias—the systematic tendency to favor cooperation with coethnics—shapes attitudes about informing and beliefs about retaliation, especially among Tajik respondents. This bias persists even after adjusting for additional explanations and potential confounding variables, suggesting that identity considerations such as coethnicity may influence attitudes toward high-risk behavior in wartime settings.

**W**hat explains the willingness of individuals to inform on insurgents during civil wars? To date, our theories of civil war violence have emphasized the potential role of civilians as sources of information about insurgent identities and activities. Indeed, a basic precept of counterinsurgency warfare is that state authorities and external interveners struggle to correctly identify insurgents who blend within the broader populace (Kalyvas 2006, 179; 2012; Leites and Wolf 1970; Lyall and Wilson 2009). Yet despite the central role of wartime informing, we still lack an individual-level empirical investigation into the different motives that might drive civilians to consider engaging in such risky wartime behavior.

At first glance, it appears likely that identity considerations such as shared ethnicity would influence wartime decision making. Scholars have often used ethnicity to explain

civil war onset (Cederman, Wimmer, and Min 2010; Horowitz 1985; Sambanis 2001; but see Fearon and Laitin 2003) as well as the patterns of violence (Lyall 2010; Petersen 2002; Posen 1993; Weidmann 2011) and combatant support (Lyall, Blair, and Imai 2013) that unfold once the fighting has begun. More generally, laboratory and field experiments have demonstrated that coethnics typically possess higher rates of collective goods provision than noncoethnics (Habyarimana et al. 2009). This suggests that information flows hinge on the identities of individuals involved.<sup>1</sup> Yet, the persistence and importance of these ethnic ties in wartime environments remain an open question, especially when potentially life-and-death activities such as informing are considered.

Recent studies of wartime dynamics have also suggested multiple motives for informing that stem more from bat-

---

Jason Lyall (jason.lyall@yale.edu) is associate professor, Department of Political Science, Yale University, New Haven, CT 06520. Yuki Shiraito (shiraito@princeton.edu) is a PhD candidate, Department of Politics, Princeton University, Princeton, NJ 08544. Kosuke Imai (kimai@princeton.edu) is professor, Department of Politics, Princeton University, Princeton, NJ 08544.

This research was approved by Yale University’s Human Subjects Committee under Institutional Review Board protocol 1105008575. Financial support for the survey from Yale’s Institution for Social and Policy Studies’ Field Experiment Initiative and the MacMillan Center for International and Area Studies is gratefully acknowledged. Additional support from the Air Force Office of Scientific Research (Lyall; grant FA 9550-09-1-0314) and the National Science Foundation (Imai; grant SES-0849715) is also acknowledged. Data and supporting materials necessary to reproduce the numerical results in the article are available in the *JOP* Dataverse (<https://dataverse.harvard.edu/dataverse/jop>). An online appendix containing supplemental analyses is available at <http://dx.doi.org/10.1086/681590>.

1. Following Chandra and Wilkinson (2008, 517), *ethnicity* is defined as an identity category in which descent-based attributes are necessary for membership.

tlefield dynamics than shared identities. Scholars have alternatively cast civilian attitudes and behavior as products of the relative distribution of combatant control (Bhavnani, Miodownik, and Choi 2011; Kalyvas 2006, 2012; Kalyvas and Kocher 2009); the nature of one's exposure to violence, as civilians punish guilty parties by withholding information (Bennett 2008; Condra and Shapiro 2012; Kocher, Pepsinsky, and Kalyvas 2011; Stoll 1993); and efforts to win "hearts and minds" through the provision of economic assistance, services, or monetary rewards that convince recipients to repay in the currency of useful tips, especially if they are poor or unemployed (Akerlof and Yellen 1994; Beath, Christia, and Enikolopov 2011; Berman, Shapiro, and Felter 2011; Department of the Army 2007; Findley and Young 2007; Vanden Eynde 2013). There may, in fact, be little distinctively "ethnic" about many of the motives behind wartime behavior, including informing.

Sorting out the relative contributions of these accounts is difficult, not least because such a task demands individual-level data from populations experiencing violence (or its threat) in civil war settings. Unsurprisingly, this requirement raises a host of methodological, ethical, and logistical challenges. These include obtaining access to these populations, ensuring enumerator and respondent safety, and mitigating biases inherent in answering questions about sensitive topics in dangerous environments. Twin obstacles—the classified nature of informant data and ethical considerations in tracking such risky wartime behavior—typically preclude the use of behavioral indicators to measure informing.

As an alternative research strategy, we use an indirect survey methodology known as endorsement experiments to confront these various challenges among a random sample of 2,700 Pashtun and Tajik respondents in 100 villages in Afghanistan (see Blair et al. [2013], Bullock, Imai, and Shapiro [2011], and Lyall et al. [2013], for the methodology and applications of endorsement experiments). We use the Guardians of Peace (GP) program, launched in 2010 by the International Security Assistance Force (ISAF) and Afghan National Security Forces (ANSF), as a means of estimating the presence and magnitude of coethnic bias toward attitudes about informing. The GP program, heavily promoted across the Afghan media landscape, was designed to encourage civilians to provide tips to local authorities via telephone lines and "walk-up" visits to nearby bases. We randomly assign the ethnicity of the program's endorsers across Tajik, Pashtun, and ISAF spokesmen to measure changes in support across a battery of questions about willingness to participate in the GP program.

We analyze the endorsement experiment using a multi-level model, which efficiently combines survey responses to several related questions and incorporates individual and

village characteristics (Bullock et al. 2011). Our analysis suggests that coethnic bias plays a key role in shaping attitudes toward informing, even in high-risk wartime settings marked by (threatened) violence against civilians. Pashtun and especially Tajik respondents prefer to collaborate with coethnics than risk crossing ethnic lines to inform on local insurgents. This coethnic bias persists even after we adjust for confounding factors such as the distribution of territorial control, prior violence, and receipt of economic assistance. These findings illustrate how identity considerations such as coethnicity can influence attitudes toward high-risk wartime behavior.

## EXPLAINING INFORMING

We define *informing* as the transfer of sensitive and timely information from civilians and potentially disgruntled insurgents to counterinsurgent forces about the identities or activities of armed groups during wartime. Informing helps counterinsurgents offset insurgent informational advantages. Indeed, obtaining tips from locals can facilitate the selective targeting of insurgent leaders, the disruption of military operations (e.g., ambushes), and possible insurgent defection.<sup>2</sup>

Public knowledge that the counterinsurgent has penetrated the village may undermine the willingness of individuals to collaborate with insurgents for fear of discovery, further complicating insurgent recruitment and operations. In turn, the rise of informants can force insurgents to devote a greater share of their resources into hardening their organization against defection and information leaks, introducing new inefficiencies that may cripple their effectiveness. It is little wonder, then, that counterinsurgents typically build large and intrusive intelligence-gathering operations (Galula 2006, 84, 87–88; Kalyvas 2006, 105–7; Leites and Wolf 1970, 136–37; Thompson 1966, 84).

## Existing approaches

To date, the twin difficulties of access and measurement have largely forced scholars to "black box" civilian attitudes and wartime behavior in our theories of civil war dynamics. Indeed, theories from different analytical approaches—including political science, behavioral economics, and agent-based modeling—rest on a remarkably similar set of (untested) assumptions about civilian wartime motivations.

In these accounts, civilians are viewed as rational individuals who are typically politically neutral and undecided about supporting an insurgency. Prewar ties, notably ethnic or ideological allegiances, are viewed as dissolving at the

2. Civilians can also provide information to insurgents to barter for protection or manage conflict. See Kaplan (2013).

war's outset or, at the least, eroding steadily over time as wartime pressures mount. Civilians act on individualistic survival-maximizing imperatives, not group-based identification. Insurgent-counterinsurgent interactions are viewed as a contest in which civilian loyalty is the prize. Both combatants seek to win over fence-sitting civilians using selective incentives such as aid and services while wielding violence (and its threat) to enforce compliance. Indiscriminate violence against civilians, especially by the counterinsurgent, is viewed as especially counterproductive, driving individuals to withhold valuable information from, and possibly to take up arms against, perpetrators. In short, civilians are reactive, individualistic weather vanes, tacking with the prevailing winds and thus capable of shifting support toward either combatant seamlessly as conditions warrant.<sup>3</sup>

Given the presumed fluidity of civilian preferences, leading research emphasizes different battlefield dynamics that drive possible collaboration with the counterinsurgent. We consider three possible arguments: (1) the role played by the distribution of territorial control among combatants; (2) how indirect and personal exposure to violence by the combatants affects attitudes toward informing; and (3) economic incentives, including the receipt of economic assistance from the counterinsurgent and the role of an individual's income. Limited empirical support exists for each position, although tests with individual-level data in settings where informing is a realistic prospect are rare.

First, Kalyvas (2006) has persuasively argued that denunciation tracks closely with the prevailing distribution of territorial control among combatants (see also Baicells 2011; Bhavnani et al. 2011; Kalyvas 2012; Leites and Wolf 1970). This "collaboration-control" model generates two empirical predictions: (1) areas where civilians can safely denounce insurgents—that is, areas totally or mostly controlled by counterinsurgent forces—will record the highest incidence of denunciation since the threat of retaliation is lowest, and (2) no informing should occur in areas of mixed control since individuals cannot judge which side will ultimately prevail, encouraging fence-sitting between combatants (Kalyvas 2006, 111–13, 197–205). In sum, pre-existing (ethnic) preferences are a misleading guide to civilian behavior (103, 113) since "as the war evolves, control is more likely to trump prewar preferences in determining collaboration" (112).

Second, an individual's willingness to consider informing is likely to be shaped by exposure to combatant vio-

lence. We can conceptualize this exposure in two ways: indirect exposure, typically via combat near (or in) one's immediate area, and direct personal exposure, via selective or indiscriminate targeting by at least one combatant. The deterrent effects of violence may become increasingly credible as attacks in one's area increase, underscoring insurgents' ability to punish would-be informants, for example (Mason 1996). In this case, we would expect tips to dry up as insurgent operations become more frequent. Yet, the opposite may be true. As violence increases in a local setting, we may observe a corresponding increase in the likelihood of informing. Civilians may use tip lines to manage insurgent violence or to avoid continued operations by counterinsurgents (Lyal 2009).

Suffering harm is also likely to affect attitudes toward informing. Aggrieved individuals may simply shed their prior loyalties to a particular combatant and use information as a way to punish the combatant that harmed them. Information thus flows symmetrically from harmed individuals to the side engaging in relatively less victimization at that moment in time (Condra and Shapiro 2012; Stoll 1993). More generally, indiscriminate violence is thought to encourage individuals to join the insurgency, thus decreasing tips (Bennett 2008; Findley and Young 2007; Kalyvas 2006; Nagl 2005). But, victimization may deter thoughts of engaging in additional risky practices, leading to the opposite prediction: harmed individuals are less likely to provide information, desires for revenge notwithstanding. After all, rebels typically devote considerable resources to implementing strategies of intimidation ranging from warnings to outright assassination in a bid to forestall denunciation to authorities (Mason 1996).

Finally, counterinsurgents and insurgents alike often provide economic assistance and services in a bid to win over hearts and minds (Beath et al. 2011). This so-called gratitude theory suggests that individuals provide information to the side that delivers the most attractive package of funds, services, and protection. The US Army field manual makes this logic explicit: "People pursue essential needs until they are met, at any cost and from any source. People support the source that meets their needs. If it is an insurgent source, then the population is likely to support the insurgency. If the Host National government provides reliable essential services, the population is more likely to support it" (Department of the Army 2007, 98; see also Berman, Shapiro, and Felner 2011; Findley and Young 2007; Nagl 2005; Trinquier 2006).

We might also imagine that individuals who are unemployed or poor are most likely to be recruited as government informers if monetary rewards are offered (Berman, Callen et al. 2011). In one recent study, negative labor shocks

3. Important representative accounts that rest on these assumptions include Bennett (2008), Berman, Shapiro, and Felner (2011), Epstein (2002), Findley and Young (2007), Kalyvas (2006, 2012), Leites and Wolf (1970), Mason (1996), and Nagl (2005).

in India are linked to increased Naxalite violence against civilians (Vanden Eynde 2013). The mechanism here is cost-benefit calculations of informing: as incomes fall, informing becomes a more attractive prospect, a fact recognized by insurgents who step up their campaign against civilians to forestall collaboration. It stands to reason that individuals at or near subsistence levels will find monetary rewards of informing hard to ignore.

### Coethnic bias and wartime informing

While these explanations for informing are plausible, we believe that existing accounts have downplayed how group identification can shape attitudes during wartime. In particular, we argue that coethnic bias—a persistent preference for cooperation with coethnics over noncoethnics—governs decisions about engaging in risky wartime behavior, including informing. Decisions about informing are therefore conditional on the group identities of the would-be informant and his interlocutor. If we are correct, individuals will be more likely to share sensitive tips with members of their own ethnic group rather than across ethnic boundaries.

Why is coethnicity important? Generally speaking, coethnicity helps respondents solve the problem of assessing the trustworthiness of individuals rapidly in an uncertain environment (Lyall 2010). Coethnicity acts as a visible signpost that allows individuals to gauge the expected behavior of individuals, including the credibility of assurances (such as promises of anonymity when providing tips) and threats of retaliation (Greenwald, McGhee, and Schwartz 1998; Hahn et al. 2014). As a result, shared coethnicity generates incentives to trust fellow coethnics and to believe that punishment will be levied if individuals do not act on behalf of group interests such as collective security.

More specifically, coethnicity influences an individual's decision calculus in three ways. First, considerations of group identification help individuals mitigate the risks of retaliation for informing. Favoring cooperation with coethnics, whose assurances of confidentiality are likely more credible than those of noncoethnics, offers one heuristic that helps would-be informants avoid sanctions for their actions. There may also be costs associated with inaction: coethnics can more readily identify individuals who had useful information and failed to share it than noncoethnics, given the density of ties between coethnics (Fearon and Laitin 1996; Habyarimana et al. 2009, 11–13).

Second, coethnicity moderates the effects of civilian victimization; rather than assuming that individuals respond uniformly to violence, the actual effect of victimization hinges on who inflicted the harm. Suffering harm at the hands of in-group members may not necessarily heighten willingness to

inform. Victimization by members of another ethnic group, however, may increase an individual's willingness to inform, especially if one's interlocutor is a coethnic (Lyall et al. 2013). Finally, as Mason (1996, 73) has argued, prior attachments such as ethnicity can influence the "price" that individuals demand for supportive behavior toward a noncoethnic or external counterinsurgent.

Wartime renders coethnic bias even more salient. The fluid nature of conflict settings increases the importance of accurate, if probabilistic, assessment of others' intentions and behavior. In our view, coethnic bias will persist under wartime conditions rather than fade away, as expected by existing theories. The "sticky" nature of coethnic bias also suggests that counterinsurgents, especially external interveners, will struggle to win civilian hearts and minds. Influencing civilian attitudes is not simply a function of identifying the appropriate mix of selective incentives and (threats of) violence. Instead, counterinsurgents must work to supplant preexisting coethnic biases that civilians rely on as cognitive frameworks for understanding how and when to interact with others. Repeated interaction between counterinsurgents and civilians may only reinforce these biases, undercutting efforts to convince would-be informants to provide tips.

Two additional points bear emphasizing. First, our argument does not assume that coethnics feel a special affinity for one another or that they necessarily share preferences over outcomes (Alesina, Baqir, and Easterly 1999; Alesina and LaFerrara 2005). Ethnic groups can be divided internally, and counterinsurgents and insurgents alike often exploit these cleavages. In some cases, private feuds may trump ethnic identification, leading to some defection. Yet, these internal divisions notwithstanding, we should still observe a persistent aggregate bias toward cooperation with coethnics given the density of preexisting ties.

Second, the magnitude of coethnic bias is not necessarily symmetrical. It is possible, for example, that group A has a profound mistrust of group B, while members of group B exhibit only a weak preference for coethnic collaboration. The magnitude of coethnic bias is likely sensitive to several factors, including demography: the smaller a group is relative to its noncoethnic referent group, the more likely its bias will be large. An individual's self-esteem (and safety) is also tied to a group's relative position within society as well as the individual's own position within her ethnic group (Horowitz 1985; Petersen 2002; Tajfel et al. 1971). Prior experience with violence across ethnic lines, and especially being victimized by a larger group, is likely to reinforce biases within the smaller group. And groups with internal divisions, including strong tribal allegiances, will exhibit

weaker coethnic bias toward noncoethnics as these schisms diminish identification with the broader ethnic group.

### EMPIRICAL APPROACH

There was little warning that the rapid toppling of Afghanistan's Taliban regime in the immediate aftermath of the September 11, 2011, would embroil the United States (and its NATO allies) in the longest war of its history. Yet the fall of Kabul created a political vacuum that gave breathing room for a weakened Taliban to reconstitute its forces. By 2006, the ISAF was confronted by a widening, and grinding, insurgency that forced a "surge" of combat forces in 2009 to bolster Kabul's flagging fortunes. Tied to a mercurial President Karzai, ISAF and various international agencies and institutions sought to marry the selective use of violence with large sums of development funds to convince Afghans of the central government's legitimacy. By early 2014, however, progress at extending Kabul's writ had proven illusive at best as emboldened Taliban forces consolidated their gains throughout eastern and southern Afghanistan. Meanwhile, the United States and its allies planned for an almost complete withdrawal by the end of 2014.

A central plank in these efforts has been the encouragement of civilians to support central authorities through the provision of tips about the nature and identity of insurgent organizations and nonsanctioned militia (*arbaki*). We draw on one such ISAF initiative, the GP program, as a window to explore coethnic bias and informing in wartime. The following sections detail the nature of the GP program as well as our measurement and modeling strategies. We also explain how we measure citizens' willingness to inform, using indirect endorsement experiments. We then present our survey sampling design and provide an initial descriptive analysis. Finally, we introduce our statistical model that allows us to combine survey responses to multiple questions into a single ideal point estimate of GP support.

### Guardians of Peace program

The GP program, first launched in southern Afghanistan in early 2010 but quickly scaled up nationally, was designed to elicit anonymous tips about the identities and activities of insurgents and nonsanctioned militia (*arbaki*) from locals. On the basis of US-style community policing tactics ("see something, say something"), the GP program encouraged villagers to provide information via dedicated hotlines or visits to neighboring ISAF and ANSF bases, where it was hoped that initial walk-up contacts could be cultivated as more regular informants. Financial compensation was awarded for useful tips, although exact details about disbursement remain classified (see, e.g., ISAF 2010).

The GP program was actively promoted through various media, including national television (i.e., Tolo TV) and radio commercials, roadside billboards, village *shuras* (councils), leaflets, business cards, and comic books distributed by patrolling ISAF and ANSF soldiers.<sup>4</sup> Even Afghan celebrities, including singer Habib Qaderi, have been enlisted to promote it.<sup>5</sup>

As a consequence of this media campaign, the GP program was well known among the target population of our survey.<sup>6</sup> The extensive penetration of media technologies into Afghanistan over the past decade also ensured a broad audience for these appeals. A full 94% of our respondents reported owning a radio, for example, while 34% possessed a television. Almost 71% owned a cell phone, making it feasible for would-be informants to provide tips anonymously to state authorities. The near ubiquity of GP advertisements, coupled with high rates of private ownership of radios and (less so) televisions, also reduces possible "information effects" that can arise when new information is presented to poorly informed individuals (Althaus 1998). As a result, we interpret our findings as produced by endorsement effects, not the provision of new information.

### Indirect measurement through endorsement experiments

We rely on an indirect measurement approach—endorsement experiments—to mitigate many of the problems associated with conducting surveys in conflict settings (Blair, Imai, and Lyall 2014; Blair et al. 2013; Bullock et al. 2011; DeMaio 1984; Lyall et al. 2013). We are especially concerned about social desirability bias, where individuals simply provide the answers that they expect will satisfy the interviewer to avoid sanctions or in the hopes of receiving (continued) aid. Such issues are especially likely to rise in wartime, where many seemingly banal issues become highly politicized and where both combatants have been rendering assistance—and threatening or directly harming respondents—in a bid to win hearts and minds.

The mechanics of an endorsement experiment are straightforward. A sample of respondents is first randomly divided into two groups. In the control group, respondents are asked to rate the level of their support for a particular policy. For those in the treatment group, the same question is asked except that the policy is said to be endorsed by an

4. These appeals to report "suspicious activities" typically follow a standardized script. We present an example of these appeals from a nationally televised commercial in appendix sec. 1.1.

5. See, e.g., <http://www.youtube.com/watch?v=8s2VcF0fZmw>.

6. Indeed, our interviewers often discussed their favorite television and radio spots during training sessions.

actor of interest. We therefore take advantage of subtle cues induced by endorsements (or names) and interpret the difference in responses between the treatment and control groups as evidence of support (or lack thereof) for this actor.

We randomly assigned a respondent to one of the three endorsers of the GP program: an ISAF spokesman, to capture the role of the external intervener; a Pashtun Afghan National Police (ANP) official; and a Tajik ANP official.<sup>7</sup> We posed four identical questions about the program; only the endorser's identity was changed across respondents. We selected names in consultation with our field staff that were easily identifiable as Tajik or Pashtun in nature. This setup mirrors reality: both ANSF and ISAF soldiers have participated in disseminating knowledge about the program (including patrols), and so randomizing endorser ethnicity does not raise concerns about unrealistic endorsers. To ensure that individuals could correctly identify the endorser's identity (Habyarimana et al. 2009, 48–57; Harris and Findley 2014), we appended a descriptive phrase (“a Pashtun/Tajik official”) after the endorser's name. We then compare the effects of ISAF, Tajik, and Pashtun endorsements on support for the GP program across Tajik and Pashtun respondents.

Four endorsement questions were posed in our experiment. Specifically, we asked whether the respondent would be willing to call the phone number provided by the GP program (“guardians”), whether the respondent believed that phone calls to the hotline would remain anonymous (“anonymity”), whether respondents would be willing to stop by local ANP stations and outposts to report suspicious activities by armed antigovernment organizations “that may result in harm to innocent Afghans” (“stopping by”), and how likely respondents viewed Taliban retaliation if they participated in the GP program (“retaliation”). We used a five-point response scale for all the questions. For reasons explored below, our analysis indicated that the response distribution to this last question differs significantly from the other questions. We therefore examine the initial three questions together before turning to beliefs about retaliation in the online appendix. These survey questions are reproduced in the appendix (see sec. 1.2).

### Survey sampling and descriptive analysis

Our survey experiment was conducted between November 21 and December 11, 2011, in 100 rural villages located

7. The endorser's identity remained fixed across all four questions rather than being randomly assigned for each question. Covariates at the individual level across our three endorsement conditions are balanced, with no statistically significant differences observed (see online appendix table 2).

in 10 districts of five provinces in Afghanistan. The sampled villages are indicated by circles in figure 1. The survey was conducted by the Opinion Research Center of Afghanistan (ORCA), an Afghan-owned firm that recruits its enumerators from sampled and neighboring villages. A 270-respondent pilot was conducted in 10 randomly sampled villages in these same districts (September 4–10, 2011) to pretest our questions (including alternative endorsement experiment questions), to gauge question-order sensitivity, and to obtain current information about the security environment. These villages were subsequently removed from the sampling frame.

A multistage sampling design was employed to identify our sample of Pashtun and Tajik respondents, the two largest ethnic groups in Afghanistan. Using population estimates from the Central Statistical Office, ORCA, and our purpose-built Afghanistan Population Index database,<sup>8</sup> we first identified 100 ethnically mixed districts from Afghanistan's 400 total districts. All of these districts had GP programming. We also included two districts in Helmand (Lashkar Gah and Nad Ali). Although almost exclusively Pashtun, these districts witnessed the introduction of Tajik-staffed Afghan National Army and Police forces into key towns during a major Taliban-clearing operation in February 2010 (Operation Moshtarak). These districts were GP test bed sites—as part of a larger “government in a box” initiative designed to bring governance and services to newly cleared areas—and so provide an opportunity to explore coethnic bias in strategic areas.

We then randomly sampled 10 districts from these 102 districts (two per province). Next, we stratified villages according to their ethnic composition using categories of “Pashtun only,” “Tajik only,” and “mixed.” We sought a breakdown of 50% mixed villages, 25% Tajik only, and 25% Pashtun only.<sup>9</sup> We obtained 10% or better coverage of villages in each district except for Nad Ali (8%).

Finally, we used a “random walk” procedure to identify households and then selected respondents among adult males (18 years or older) using a Kish grid.<sup>10</sup> Each village had 27 respondents (nine responses for each of the three

8. The Afghanistan Population Index combines multiple data sources, including ORCA, Afghanistan's Central Statistical Office and Ministry of Rural Rehabilitation and Development, and USAID, into a single composite data set of populated locations in Afghanistan.

9. Appendix fig. 1 depicts the distribution. We also collected data on ethnic composition by asking respondents about their perceptions of the village's ethnic breakdown and by having enumerators, who were all locals, classify the village's composition in a postsurvey module.

10. Given our survey locations, we were unable to interview female respondents.

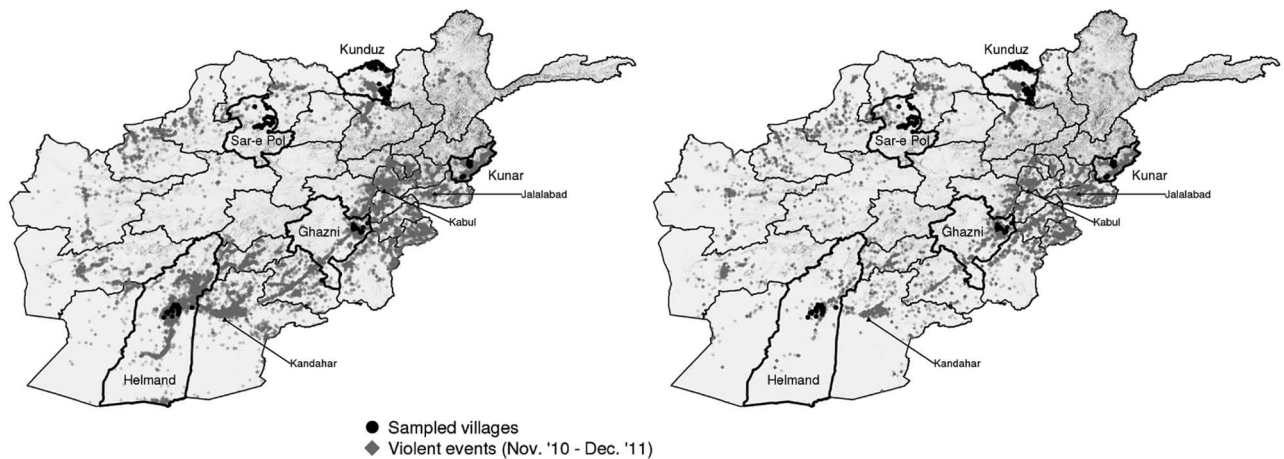


Figure 1. Circles represent 100 randomly sampled villages. Diamonds represent ISAF- (International Security Assistance Force) and insurgent-initiated violent attacks one year before our survey launch (November 2011). *Left*, depicts violence using ISAF's own CIDNE database; *right*, draws on iMMAP, which collates event data from nongovernmental organizations in Afghanistan. Color version available as an online enhancement.

versions). Our refusal rate was less than 15% (2,700 respondents among 3,160 approached). The majority of interviews were completed on the first (62%) or second (24%) attempt. Of the 460 failed contacts, 202 were due to non-response (no adult male was available after three visits), and 258 refused to participate; the most common reason was “in a hurry” ( $N = 70$ ).<sup>11</sup>

Of course, field research in conflict settings poses a special set of challenges. We encountered Taliban and *arbaki* (militia) checkpoints in many of our districts, particularly in Kunduz and Kunar, that restricted movement (although not access). Helmand continued to be extremely violent during our fieldwork, and Ghazni and Kunar both recorded daily (or nearly so) violence in surveyed districts (see table 1). One district in particular, Wati Pur in Kunar Province, proved especially difficult. With Wati Pur long considered a Taliban redoubt, two interviewers were detained by the Taliban as “foreign spies” before conducting their surveys. Village elders interceded on behalf of the interviewers (who were locals), and the Taliban soon released them, albeit with minor injuries. In total, we replaced only four villages of the original 100 selected, a testament to the skill and courage of our survey teams.

Three additional points about the real world nature of this survey experiment are stressed here. First, nearly every village recorded at least some measure of ISAF or ANSF presence, although the strength of government control var-

ied considerably. As a result, the option of providing tips to progovernment forces was a realistic possibility for would-be informants, via either anonymous calls or visiting outposts.

Second, the majority of our Tajik and Pashtun respondents were in a position to provide tips about the actions of both their fellow coethnics and noncoethnics. Pashtuns, for example, could denounce the Pashtun-dominated Taliban (along with some Pashtun-staffed *arbaki*) while also providing information about Tajik *arbaki*, particularly in settings of mixed ethnic settlement and where rebel governance overlapped between multiple armed groups. Tajiks, for their part, could denounce Pashtun members of the Taliban as well as coethnic *arbaki* members. By virtue of the relative distribution of government and rebel control, informing within and across ethnic lines was possible for respondents in our sample.

Third, although insurgent violence tends to be lower in Tajik-majority districts, we should not conclude that Taliban and various *arbaki* forces are absent. In fact, in Tajik-majority Gosfandi, one of the least violent districts in our sample, Taliban members were collecting harvest taxes (*ushr*) and tithing (*zakat*) in April 2011, only months before our survey (Pajhwok Afghan News 2011).

Taken together, these conditions lend credence to our contention that support for the GP program can be interpreted as willingness to inform on insurgent activities to ISAF or ANSF forces. As a first step in this investigation, figure 2 illustrates the distribution of responses for Tajiks and Pashtuns across Tajik, Pashtun, and ISAF endorsers. Considerable heterogeneity is apparent both within and across our respondents' answers. This is an ideal situation,

11. As part of our quality control, district supervisors directly monitored 10% of the interviews and back-checked another 15%, and ORCA's Kabul-based staff randomly audited a further 5% via callbacks.

Table 1. Overview of the Multistage Sampling Design

District	Villages		Individuals		CIDNE: Violent Events Initiated By		iMMAP: Violent Events Initiated By		
	Total	Sample	Total	Sample	Taliban	ISAF	Taliban	ISAF	Unknown
Deh Yak	62	6	72,680	162	157	79	221	191	5
Ghazni Center	74	8	146,403	216	307	134	594	274	26
Lashkar Gah	73	8	126,926	216	546	68	353	144	23
Nad Ali	155	12	290,866	324	6,415	1,417	1,064	488	23
Khas Kunar	43	5	45,166	135	93	21	83	34	2
Wati Pur	82	9	33,023	243	339	26	325	33	2
Imam Sahib	141	14	138,048	378	64	12	100	65	5
Khanabad	148	15	143,243	405	37	12	82	68	3
Gosfandi	109	11	59,570	297	3	1	9	1	0
Sangcharak	122	12	83,412	324	2	0	9	3	1
Total	1,009	100	1,139,337	2,700	7,963	1,770	2,840	1,301	90
92 nonsampled mixed districts	8,205	0	6,102,601	0	5,666	2,303	4,866	3,191	149

Note. Violent events by district (one year before the survey) are based on CIDNE (Combined Information Data Exchange) and iMMAP databases. ISAF = International Security Assistance Force.

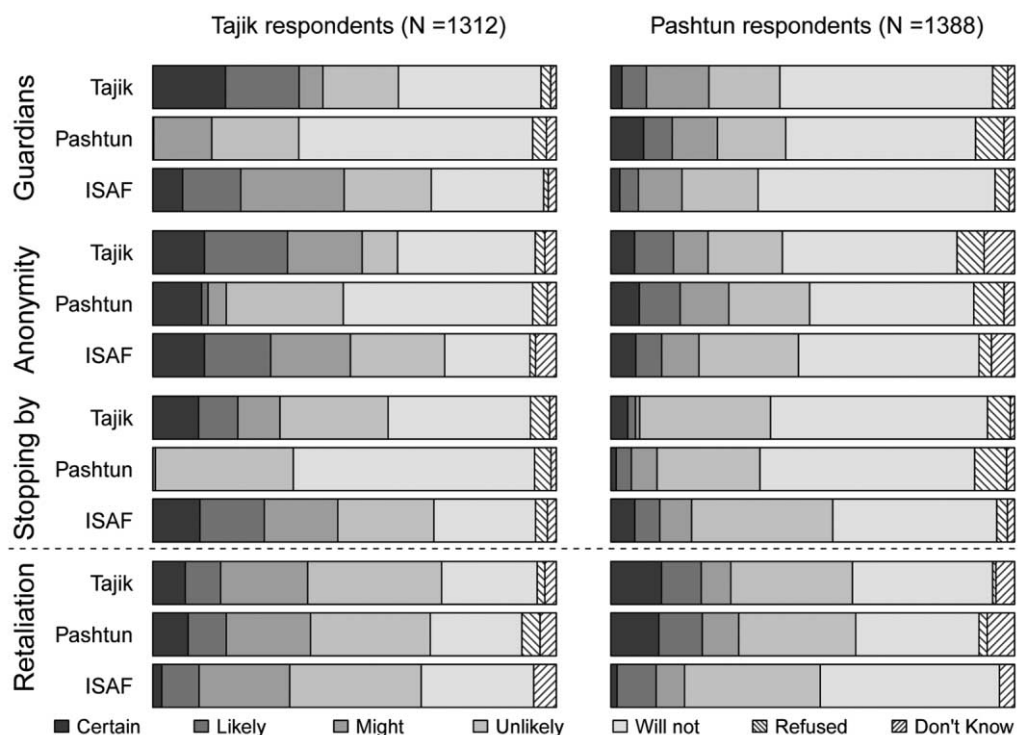


Figure 2. Distribution of responses to four survey questions (guardians, anonymity, stopping by, and retaliation) across three endorsement groups (Tajik/Pashtun/ISAF [International Security Assistance Force]) among Tajik respondents (left) and Pashtun respondents (right). For the retaliation question, labels for the response categories are reversed, so that darker gray indicates higher support for the Guardians of Peace program as in the case of the other questions. Specifically, the darkest gray represents “retaliation is highly unlikely” while the lightest gray represents “retaliation is highly likely.” All the survey questions are reproduced in appendix section 1.2. Sample sizes are shown in parentheses.



enabling us to detect support for endorsers more readily than if attitudes have coalesced around a single position. In addition, the low rate of “refuse to answer” and “don’t know” is encouraging given the sensitive nature of these questions. Finally, we already observe some initial evidence on the possible existence of coethnic bias: compare, for example, the differences between Tajik and Pashtun endorsers among Tajik respondents for questions about the likelihood of calling the dedicated hot line (“guardians”) and stopping by ANP posts (“stopping by”).

**Statistical modeling**

While descriptive analysis is informative, we use a statistical model to combine the responses to three different endorsement questions across three different treatment conditions. The use of statistical modeling also enables us to examine the relationship between theoretically important covariates and the latent willingness to inform. To do this formally, we follow the factor analytic model proposed by Bullock et al. (2011). The model is designed to efficiently extract systematic patterns common across multiple questions. In addition, our multistage sampling design naturally leads to the use of a hierarchical model where the village-specific effects are modeled with village-level covariates. This model also accounts for within-village clustering that may be present.

Formally, let  $Y_{ij}$  denote the observed ordered response variable, which takes one of the following values:  $\{1, 2, \dots, L\}$ . We use  $T_i$  to indicate the randomized “treatment” variable, which represents the endorser assigned to respondent  $i$ . Then, the individual-level model is given by the following ordered probit factor analytic model:

$$\Pr(Y_{ij} \leq l | T_i = k) = \Phi[\alpha_{jl} - \beta_j(x_i + s_{ijk})], \quad (1)$$

where  $\alpha_{j1} = 0$ ,  $\alpha_{jL} = \infty$ , and  $\alpha_{jl} < \alpha_{j,l+1}$  for any  $j$  and  $l$ . In this model,  $x_i$  represents the level of respondent  $i$ ’s willingness to support the GP program, and  $s_{ijk}$  denotes the effect of endorsement by group  $k$  on question  $j$  for respondent  $i$ . As in the standard item response theory,  $\alpha_{jl}$ ’s are the item difficulty parameters, and  $\beta_j$  is the item discrimination parameter. In the current context,  $\alpha_{jl}$ ’s reflect the degree to which a certain aspect of the program is supported, whereas  $\beta_j$  represents the amount of information each question reveals about respondents’ willingness to support the program.

We are interested in  $x_i$  (the level of respondent’s willingness to support the GP program) and  $s_{ijk}$  (the level of support for endorsers), and these parameters are modeled hierarchically as follows using the individual-level covariates  $Z_i$  and the village-level covariates  $V_{\text{village}[j]}$ :

$$x_i \stackrel{\text{indep.}}{\sim} \mathcal{N}(\delta_{\text{village}[j]} + Z_i^T \delta^Z, 1) \quad (2)$$

$$s_{ijk} \stackrel{\text{indep.}}{\sim} \mathcal{N}(\lambda_{k,\text{village}[j]} + Z_i^T \lambda_k^Z, \omega_k^2) \quad (3)$$

$$\delta_{\text{village}[j]} \stackrel{\text{indep.}}{\sim} \mathcal{N}(\delta + V_{\text{village}[j]}^T \delta^V, \sigma^2) \quad (4)$$

$$\lambda_{k,\text{village}[j]} \stackrel{\text{indep.}}{\sim} \mathcal{N}(\lambda_k + V_{\text{village}[j]}^T \lambda_k^V, \psi_k^2) \quad (5)$$

The model therefore allows us to investigate how individual- and village-level covariates are associated with respondents’ willingness to support the program, as well as how these variables determine the size of endorsement effects. The model is completed with the conjugate prior, where the diffuse normal and inverse chi-squared prior distributions are placed on the coefficient and variance parameters, respectively. The model is fitted using the R package endorse (Shiraito and Imai 2012), and the standard convergence diagnostic based on three independent Markov chains was performed.

We fit this model with a series of individual- and village-level covariates that are emphasized by existing theories when explaining willingness to inform. At the individual level, we include measures for standard socioeconomic traits such as age, ethnicity, income, and education. For Pashtun respondents, we include a covariate designating whether that individual is a member of a tribe that has publicly declared its support for the Taliban. In addition, we account for the individual’s frequency of contact with combatants (ISAF and ANP) as well as coethnics and noncoethnics (Allport 1954; Cook 1971). Our models also include self-reports of harm (both personal and property) that the individual or his family has experienced in the past year at the hands of the Taliban, ISAF, and ANP. And we include multiple terms that account for interaction between types of harm as well as between ethnicity and harm and contact with the combatants.

In addition, we incorporate six village-level covariates. These include village elevation (in logged meters) as a proxy for difficulty of state control (Fearon and Laitin 2003), the village’s logged population size and ethnic composition (as a percentage of Pashtun inhabitants), the relative control exercised by combatants over the village (Kalyvas 2006, 421–22),<sup>12</sup> a binary measure of whether the village had received an NSP (National Solidarity Program) grant in the past, the number of ISAF and ANSF military bases within a 2-kilometer radius of the village to account for troop presence and density, and the number of insurgent and ISAF-initiated attacks within a

12. As described below, we use two measures: a dummy indicator for whether the Taliban exercised a high degree of control in the village and a dummy indicator for whether the village was contested between two or more combatants.

2-kilometer radius of the village in the year before the survey (November 2010–December 2011).<sup>13</sup> These event data stem from two sources: ISAF’s own Combined Information Data Exchange (CIDNE) and iMMAP, a web-based database that multiple Western and Afghan nongovernmental organizations use to record and share security and other incidents in near-real time through crowd sourcing. Descriptive summary tables of these covariates, along with the coefficient estimates from our models, are found in appendix sections 1.4 and 1.5.

### EMPIRICAL FINDINGS

We begin by testing for the presence of a persistent coethnic bias among Pashtun and Tajik respondents. We then investigate three additional accounts for individuals’ willingness to consider informing: the control-collaboration model, the role of (in)direct victimization, and economic incentives. Finally, we explore how victimization and the actions of the external intervener (ISAF) affect the magnitude of coethnic bias. For space reasons, we discuss coethnic bias and fear of retaliation in the appendix. We use the ISAF endorser condition as our baseline category for comparison throughout the analyses that follow.

#### Coethnic bias and support for the GP program

Does coethnic bias play a role in explaining support for the GP program? As a first cut, we plot the predicted probability of each response category averaged across three endorsement questions using the ISAF endorser as a our baseline category (fig. 3). The predicted probability is calculated by substituting the Markov Chain Monte Carlo draws of the parameters into equation (1). All covariates other than the endorser variable are set at the observed values.<sup>14</sup> “Certain” represents the predicted probability across our “certain to call this telephone number,” “certain that callers will remain anonymous,” and “certain to stop by a local ANP post” categories, while “not” represents the opposite end of the five-point scale. We calculate these predicted probabilities averaged across respondents separately for each ethnic group using their observed characteristics.

We find that Tajiks are far more likely to consider supporting the GP program than their Pashtun counterparts. In fact, the predicted probability of opposing the program is nearly twice as high among Pashtuns as Tajiks, while

Tajiks are at least three times more likely to report that they are “certain” to support the program than Pashtuns. These differences are both substantively large and statistically significant for all five response categories, suggesting that support varies systematically across ethnic groups.

These interethnic differences may stem from different group-specific wartime experiences over the past decades in Afghanistan. Having experienced the harshness of Pashtun Taliban rule, Tajiks, for example, may be especially loathe to return to Taliban rule and so view the GP program as one means of avoiding this fate. Our predominantly Tajik districts, Sangcharak and Gosfandi in Sar-e Pul province, both witnessed Taliban repression as late as 2000. More generally, Tajiks do typically record higher (although not necessarily high) levels of support for the Karzai government than Pashtuns, which may translate into higher relative support for government programs.<sup>15</sup>

We stress, however, that the median reply for both Pashtun and Tajik respondents was “not participate” and “unlikely to participate,” respectively. These findings suggest that considerable skepticism exists toward the program, leaving a large portion of the population potentially outside of the state’s information-gathering apparatus. These results are a useful diagnostics for our purposes, however. Since we cannot track informing directly due to the absence of unclassified data, these attitudes toward participation in the GP program suggest that willingness tracks with our intuitive understanding of informing as a risky prospect that many individuals are unlikely to consider.

Next, we directly examine how the ethnicity of a program’s endorser influences respondents’ support for the GP program. Figure 4 plots the estimated endorsement effects, which are represented by the estimated probability that having a Pashtun (Tajik) endorser increases support for the program among Pashtun (Tajik) respondents.<sup>16</sup> Our key quantity of interest (i.e., coethnic bias) is estimated as the difference in endorsement effects between Pashtun and Tajik endorsers for Pashtun respondents (or between Tajik and Pashtun

13. We use a 5-kilometer spatial buffer for counting military bases and violent incidents as a robustness check in subsequent analyses. Figures and tables for the robustness check are in appendix sec. 1.7.

14. Quantities presented in appendix figs. 2, 3, and 7 are also computed in the same manner.

15. We also posed a question seeking to tap into group-specific notions of altruism. Tajiks were more likely to agree with the statement “when I do a favor for someone, I expect it to be returned” than Pashtuns; the difference is statistically significant. This may suggest that Tajiks on average are more likely to be drawn to financial incentives for informing than our Pashtun respondents, although the difference is substantively small and should not be exaggerated.

16. We compute this quantity on the basis of the modeling assumption in eq. (3), where the posterior draws of its parameters are used to evaluate the Gaussian distribution function. All covariates other than the endorser variable are set at the observed values. Quantities shown in appendix fig. 9 are obtained in the same manner.

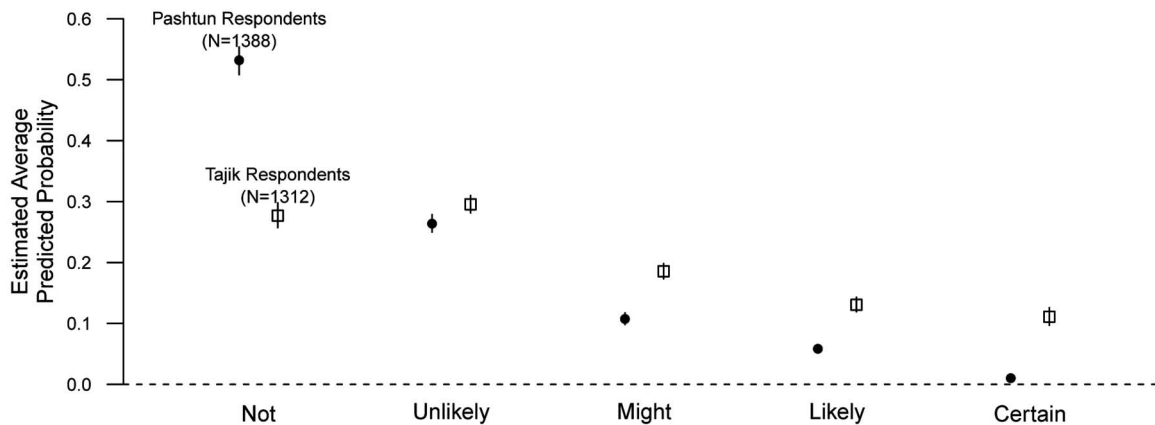


Figure 3. Estimated support for the Guardians of Peace program. Predicted probability of each response category averaged across three endorsement experiment questions under the International Security Assistance Force endorser condition for each ethnic group, with 95% confidence intervals, shown.

tun endorsers for Tajik respondents), with 95% confidence intervals as vertical lines. Here, we observe coethnic bias on full display. Pashtun respondents record higher levels of support for the GP program when endorsed by a coethnic, although the difference is fairly modest. Tajik respondents provide a much sharper example of coethnic bias: support for the program moves only slightly (relative to the ISAF endorser condition) when endorsed by a Pashtun representative, while a Tajik endorsement is associated with a substantial positive increase in estimated support for the program. Notable, too, is the fact that the ISAF endorsement (the baseline endorser category) is not viewed favorably relative to either Pashtun or Tajik endorsers for either ethnic group.

Underscoring the importance of coethnic bias is the fact that these models include standard socioeconomic and demographic variables, and yet they appear to explain little, if any, variation in support for the GP program. Age, wealth, and degree of interaction with noncoethnics, ISAF, and ANSF all appear to have little association with changes in the probability of support. Moreover, village-level covariates, notably elevation, population size, and ethnic composition, also have no consistent relationship with the likelihood of support, and none of these relationships approach conventional levels of statistical significance (see table 5 in the appendix). This indicates that our respondents' choice of strategies hinges on the coethnic status of the proposed interlocutor, particularly among Tajiks, who report little willingness to consider Pashtun-endorsed proposals.

### Additional explanations

We consider four additional explanations for an individual's willingness to inform: the distribution of territorial control between combatants, indirect exposure to violence within and near one's village, direct exposure to victimi-

zation, and the role of economic factors such as receipt of aid or per capita income. Given the observational nature of these data, we emphasize that we are exploring associations between covariates here rather than making causal statements.

**Territorial control.** Following Kalyvas (2006), we expect that the probability of denunciation should increase as the level of control exercised by the counterinsurgent increases since these forces can shield would-be informants from insurgent retaliation. Individuals within villages controlled wholly or mostly by ISAF should therefore record the greatest willingness to consider informing. To explore these claims, our interviewers assessed the combatants' control of each of the 100 surveyed villages using a standardized seven-point scale.<sup>17</sup> We use an indicator variable that assigns villages a 1 if the Taliban had complete or some control (representing a 3, 4, or 5 in our index).

As figure 2 of the appendix demonstrates, there is no statistically significant difference between the average predicted probability of participating in the GP program in

17. (1) ISAF or Afghan security forces were permanently based in this village or nearby; no Taliban activity or presence had been reported; (2) ISAF or Afghan security forces were permanently based in this village or nearby; some Taliban activity or presence had been reported, especially at night; (3) ISAF or Afghan security forces were permanently based in this village or nearby but did not move freely at night; village administrators usually do not sleep in their homes, and Taliban activity took place regularly; (4) Taliban forces were permanently based in this village or nearby and operated freely; ISAF or Afghan security forces might visit the village on occasion but did not stay; (5) Taliban forces were permanently based in this village or nearby and operated freely; no ISAF or Afghan security force presence or activity at all; (6) local *arbaki* (militia) control this village; minimal Taliban, ISAF, or ANSF presence; and (7) there were no ISAF, Taliban, ANSF, or *arbaki* controlling this village.

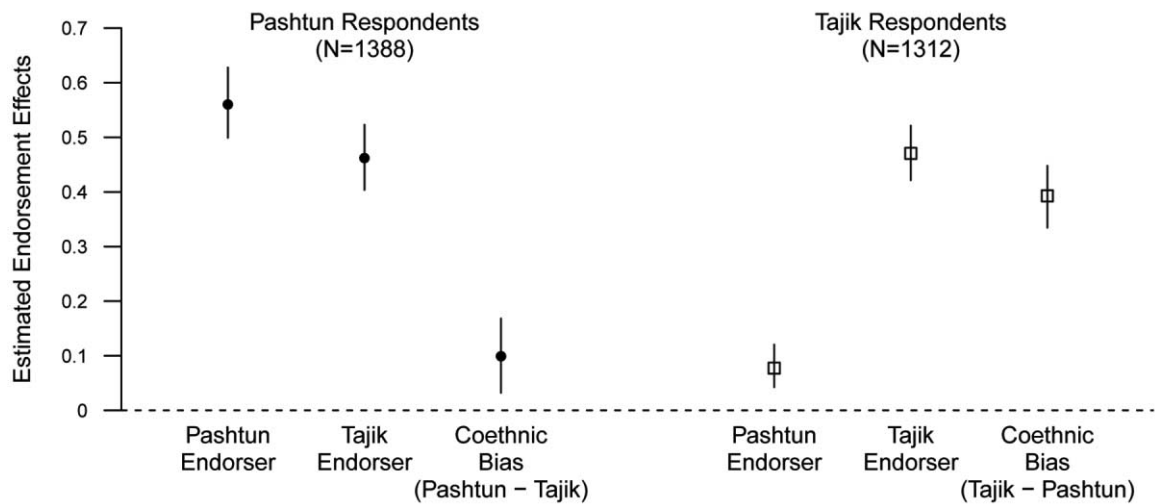


Figure 4. Estimated endorsement effects on respondents' support for the Guardians of Peace program. Effects are estimated for each ethnic group separately under the Pashtun and Tajik endorser conditions. Differences between these endorsement effects under these two endorser conditions represent coethnic bias. Lines represent 95% confidence intervals.

Taliban- and non-Taliban-controlled areas.<sup>18</sup> Indeed, the predicted probability of participation appears insensitive to the broader distribution of combatant control, with no statistically significant differences at any of our levels of control. This includes the density of military installations surrounding a given village, which is negatively associated with GP support but does not reach conventional levels of significance. A robustness check reported in figure 4 of the appendix tests the related claim that contested areas should witness no denunciation as embattled civilians opt to fence-sit rather than commit to a side. Once again we find that there is little difference between contested villages and those controlled by ISAF, local militia, or the Taliban. The same empirical pattern is repeated if we reestimate the models using an expanded 5-kilometer radius for measuring ISAF and insurgent violence (see appendix figs. 13 and 16).

Perhaps the most surprising null result is simply that Taliban control is not associated with a lower predicted probability of supporting the GP program. While any answer must remain tentative given our sample size, the non-responsiveness of support to Taliban control suggests that individual decision making may not be guided solely by security motives. It is possible, for example, that the individuals within Taliban-controlled villages are concerned about retaliation (see below), but a combination of war weariness, unease at the prospect of extended Taliban rule, and membership in non-Taliban tribal networks (if Pashtun) may conspire to mitigate fears of supporting the GP pro-

gram. Intriguingly, the share of respondents “likely” or “certain” to support the GP program is actually higher in Taliban-controlled villages, although the difference does not reach conventional levels of statistical significance. In short, exposure to Taliban rule might actually induce, rather than dissuade, individuals to consider informing.<sup>19</sup>

**Violent settings.** We also examine the more general claim that attitudes are driven by indirect exposure to violence in one’s immediate area. We draw on two different data sources—ISAF’s CIDNE and iMMAP’s Security Incidents Tracker—to measure attacks within a 2-kilometer radius of each village. CIDNE records the date and spatial location of 17 different types of violent events initiated by either ISAF or insurgent organizations.<sup>20</sup> While CIDNE is focused almost exclusively on ISAF-insurgent interactions, iMMAP records nearly 50 types of violence against civilians and Afghan government institutions (including ANSF). There are 42,248 events in CIDNE and 30,640 events in iMMAP in the year preceding our December 2011 survey.

Echoing earlier findings by Lyall et al. (2013), we find no consistent association between indirect exposure to vi-

18. We derive these estimates from the model detailed in appendix table 5; ISAF is the baseline endorser.

19. More generally, we might expect higher rates of informing in rebel-controlled areas since individuals are likely to observe far more insurgent actions (and identities) than in government-controlled areas, where such actions are necessarily clandestine.

20. The specific event categories are as follows: ISAF (Cache Found, Direct Fire, Escalation of Force, and Search and Attack) and insurgents (Assassination, Attack, Direct Fire, IED Explosion, IED False, IED Founded/Cleared, IED Hoax, Indirect Fire, Mine Found, Mine Strike, SAFIRE, Security Breach, and Unexploded Ordinance).

olence and individual attitudes (see appendix table 5). This pattern holds even if we reset the spatial buffer around the village at 5 kilometers to capture a wider geographic area (see appendix table 7).

Neither Taliban nor ISAF violence appears to influence support for the GP program, at least within our specified temporal and spatial parameters. This nonfinding represents something of a challenge for existing theories of civil war violence that test their claims solely through event data. In our case, we cannot recover the individual-level linkage between indirect exposure to violence and civilian attitudes that is posited by several leading theories of civil war dynamics. The problem likely worsens if even more aggregate spatial (e.g., district-level) and temporal (e.g., annual) categories are used to record the amount of violence that individuals are indirectly exposed to in a given location.

**Individual victimization.** Perhaps, however, direct exposure to harm, as captured by our survey instrument, provides a more precise means of estimating the association between combatant violence and GP program support. To test this relationship, we gathered data on whether a respondent or members of his family had been physically injured or killed or experienced property damage within the past year. These self-reports, elicited before the endorsement experiments, were preceded by a script defining “harm” as first death or physical injury and then property damage. Respondents were asked to identify which combatant was responsible and whether restitution had been offered.<sup>21</sup>

Yet our findings are again inconsistent with the expectation that exposure to violence by either ISAF or the Taliban is associated with an increased likelihood of supporting the GP program. This is true even of “most likely” cases for support, notably Tajik respondents who have experienced Taliban violence and who are interacting with co-ethnic interviewers. In fact, for several combinations, notably ISAF victimization of both Pashtuns and Tajiks under the ISAF endorsement (appendix fig. 3), we find that individuals victimized by third parties are much less likely to express support when compared with nonvictimized individuals.

We plot the remaining combinations of respondent, endorser, and victimizer in figures 5 and 6 of the appendix. Remarkably, with the key exception of ISAF violence, our respondents’ willingness to inform does not depend on victimization. Revenge motives that lead civilians to punish combatants for indiscriminate violence do not appear to be

operating in our sample. Instead, while ISAF is clearly punished for its violence against civilians (under the ISAF endorsement), Taliban victimization itself does not lead to an increase in support for the GP program. These patterns continue to hold even if we reset the spatial boundaries at 5 kilometers around the sample village (see appendix figs. 17 and 18).

Our findings fit instead with a commonsense view of civilian behavior in wartime: victimized individuals, absent special circumstances, are unlikely to engage in high-risk behavior such as informing to punish combatants. Yet co-ethnic bias persists even in the face of victimization. Compare, for example, the estimated average predicted probability of victimized Tajiks informing with a Pashtun endorser (about 5%) and a Tajik endorser (about 20%). While individuals are clearly willing to withhold information from ISAF if victimized, it also appears that civilians are guided by the logic of coethnic identification, not revenge, even when they have suffered at Taliban hands.

**Economic factors.** Does the receipt of economic assistance—in this case, from the NSP, the largest aid program in Afghanistan—increase support for the GP program? Over one-third (38 out of 100) of our sampled villages had received an NSP-funded block grant by the time of our survey. Administered through gender-balanced Community Development Councils, these \$60,000 block grants were typically devoted to either infrastructure or human capital development (e.g., training). All but one of our districts (Deh Yak in Ghazni province) had at least one NSP-funded village Community Development Council.

In figure 7 of the appendix, we plot the predicted probability of participating in the GP program in NSP and non-NSP villages. No statistically significant difference emerges between these two village types across any of the response categories. While the NSP program may contribute to the increased legitimacy of the Afghan government, at least in safe districts (Beath et al. 2011), this improved standing in the eyes of citizens does not appear to translate into increased tips.

Finally, we also find little evidence to support the claim that relatively poorer individuals, as measured by monthly household income, are more receptive to promises of payment for information than wealthier individuals. As figure 8 of the appendix shows, there are no differences between Tajiks’ and Pashtuns’ support for the GP program when shifting from the lowest income category ( $\leq 2,000$  afghanis per month) to the median (10,001–20,000 afghanis per month). In fact, individuals with the lowest monthly incomes are actually more likely to answer that they will

21. See Lyall et al. (2013) for a discussion of the advantages and limitations of self-reports of victimization.

not participate in the GP program under both Tajik and Pashtun endorsement conditions. Fully 70% of Pashtuns with a monthly income of  $\leq 2,000$  afghanis state that they will not participate in the program under the Tajik endorsement. On average, there is a 5%–10% decrease in the likelihood of supporting the GP program when shifting from the lowest income category to the median across both Tajiks and Pashtuns.

### **VIOLENCE AND THE ROLE OF THE EXTERNAL INTERVENER**

We also explore whether interaction with an external intervener affects the magnitude of coethnic bias. We expect that experiencing violence at the intervener's hands will have differential effects on coethnic bias (and, ultimately, on support for the GP program, as detailed below) depending on whether an individual shares ethnicity with the insurgents. If the individual and insurgents are coethnics, then we are unlikely to observe any change in coethnic bias since these differences are already accounted for by the fact that these individuals view the intervener as an out-group. If, however, the individuals and insurgents do not share ethnicity, then we argue that violent interaction with the intervener is likely to increase coethnic bias, as these aggrieved individuals bandwagon with coethnics while casting the intervener as unreliable and untrustworthy. Note that in the absence of a strong identification strategy, the estimates reported here are largely associational and thus only suggestive of causal relationships.

Figure 9 of the appendix plots the estimated differences of coethnic bias across Pashtun and Tajik respondents who have (not) experienced ISAF victimization. Victimized Pashtuns reported little coethnic bias, representing a slight decrease in coethnic bias relative to nonvictimized Pashtuns (col. 3). Within the Tajik sample, however, there is only suggestive evidence for our argument. Although ISAF victimization is associated with an increase in estimated coethnic bias, the posterior probability of the effect being positive is .885, and hence the 95% confidence interval does overlap with zero. With relatively fewer Tajiks harmed by ISAF than Pashtuns (ISAF caused 14.6% of Tajiks and 45.6% of Pashtuns in the sample both physical and property damage), we must be cautious in assigning too much weight to this finding.

In short, the magnitude of coethnic bias among our respondents appears largely unaffected by ISAF victimization. Despite some suggestive evidence for Tajik coethnic bias, neither Pashtun nor Tajik respondents who were victimized by ISAF exhibit statistically significant differences in observed coethnic bias. These findings suggest the co-

ethnic bias is not merely the by-product of civil war violence—or, at least, the actions of the external intervener—and that it has its own independent role to play in shaping individual actions in wartime.

### **CONCLUSION**

Drawing on the Guardians of Peace program in Afghanistan, we find evidence that our respondents' attitudes toward informing are influenced by persistent coethnic bias. Support for the program itself is divided along ethnic lines, while an individual's willingness to consider informing hinges on the endorser's ethnicity. Tajiks in particular have a strong preference for cooperation with fellow Tajiks compared with Pashtun endorsers. Beliefs about the odds of retaliation for participating in the GP program also track closely with the ethnic makeup of respondent-endorser pairings.

We found only mixed evidence for explanations that emphasized the territorial distribution of control, economic factors, and (in)direct exposure to combatant violence. Victimization by ISAF was associated with a reduced willingness to consider participation in the GP program, for example, while beliefs about retaliation did follow the distribution of control and the respondents' proximity to military installations. In other cases, however, our findings run counter to expectations. Victimized individuals were not more likely to express willingness to participate in the GP program, and per capita income, event counts of ISAF and Taliban violence, and economic assistance programs all appear unrelated to attitudes about informing.

It is natural to wonder, however, about the generalizability of these findings. In a narrow sense, there are no statistically significant differences for any covariates, including violence, between our sampled villages and the remaining 1,009 villages in our sampled districts (see appendix table 3). Compared to the average Afghan village, our surveyed villages have experienced more ISAF- and Taliban-initiated violence, are closer to military bases, and are somewhat larger in population (see appendix table 4). As a result, while we hesitate to extend our findings up to all of Afghanistan, our villages are representative of violent areas where the act of informing is realistic. Moreover, these villages pose a difficult test for our claims since it precisely in this environment where battlefield dynamics are thought to trump group identification.

These findings raise several implications for current theories of civil war violence. Most important, the influence that coethnic bias wields on high-stakes wartime decision making suggests that existing accounts of civilians as individualistic security-seeking actors may be too simplistic. Considerations of group identification influenced our respondent's willingness to support the GP program, the likelihood of in-

forming, and their estimates of the odds of suffering retaliation. The persistence of coethnic bias in this setting suggests that the assumption of widespread civilian neutrality—a staple of hearts-and-minds theorizing—may not have empirical support. Instead, these findings suggest that existing research designs and data may not be “micro” enough since they cannot capture the role of cognitive heuristics and other individual-level processes that shape civilian behavior, including the flow of information, in wartime.

The interplay of coethnic bias and informing also highlights the occupier’s dilemma. There is almost no support among our respondents for ISAF endorsements, underscoring how difficult it is for external actors to gain traction in the face of persistent skepticism toward noncoethnics. Indeed, ISAF endorsements fare even worse among Tajiks—its most likely set of supporters—than Pashtun ones. Hearts and minds may be far less malleable than assumed. Instead, counterinsurgent-led initiatives, if not prefaced with the right local attributes, may be self-defeating. Moreover, state-building efforts by the external occupier, including the construction of a new military apparatus and the delivery of goods and services, may be jeopardized if these coethnic biases are not addressed. And it is unclear how these efforts would gain legitimacy in the eyes of the local population. Without excellent knowledge of the direction and magnitude of coethnic biases, these initiatives can actually reinforce grievances within the population.

Our results also suggest future avenues of research. The fact that the magnitude of coethnic bias can vary substantially across ethnic groups is intriguing, for example. Tajiks report an overwhelming preference to collaborate with one another; Pashtuns, by contrast, exhibit markedly weaker coethnic biases. This variation may stem from Tajiks’ minority status (about 20% of the population) in Afghanistan, where Pashtuns (at about 40%) predominate. It may also reflect Tajiks’ prior exposure to violence by the Pashtun-dominated Taliban during its 1996–2001 rule. It is also likely that strong tribal allegiances among Pashtuns serve as cross-cutting cleavages that weaken, although do not entirely supplant, ethnic group identification.

Exploring how contextual factors such as local demography or prior interactions affect the magnitude of coethnic bias is an important next step for understanding the links between group identification, attitudes, and behavior. So too is testing both the direction and magnitude of coethnic bias in other issue areas such as collective goods provision, police-citizen relations, and perceptions of government legitimacy in fragile states. Finally, examining whether (and when) coethnic bias can be reduced is an important task for academics and policy makers alike. Minimizing the nega-

tive externalities that flow from persistent coethnic biases could result in improved services, less interethnic violence, and greater legitimacy for new governments as they seek to put the ravages of war behind them in multiethnic settings.

## ACKNOWLEDGMENTS

We thank the Opinion Research Center of Afghanistan and especially Rafiq Kakar, Mr. Abdul Nabi Barezai, Mr. Zabih Osmani, Mr. Asadi, and Mr. Tahiri, along with the 110 district managers and interviewers who conducted the survey, for helpful feedback and excellent work under trying conditions. Our program manager, Prakhar Sharma, deserves special thanks. We also thank Joshua Gubler for helpful comments as well as our editors and anonymous reviewers. An earlier version of this article (“Dropping a Dime: Coethnic Bias and Wartime Informing”) was presented at the 2013 American Political Science Association meeting.

## REFERENCES

- Akerlof, George, and Janet Yellen. 1994. “Gang Behavior, Law Enforcement, and Community Values.” In Henry Aaron, Thomas Mann, and Taylor Timothy, eds., *Values and Public Policy*. Washington, DC: Brookings, 173–204.
- Alesina, Alberto, Reza Baqir, and William Easterly. 1999. “Public Goods and Ethnic Divisions.” *Quarterly Journal of Economics* 114:1243–84.
- Alesina, Alberto, and Elina LaFerrara. 2005. “Ethnic Diversity and Economic Performance.” *Journal of Economic Literature* 63:762–800.
- Allport, Gordon. 1954. *The Nature of Prejudice*. Reading, MA: Addison-Wesley.
- Althaus, Scott. 1998. “Information Effects in Collective Preferences.” *American Political Science Review* 92:545–58.
- Baiccels, Laia. 2011. “Continuation of Politics by Two Means: Direct and Indirect Violence in Civil War.” *Journal of Conflict Resolution* 55:397–422.
- Beath, Andrew, Fotini Christia, and Ruben Enikolopov. 2011. “Winning Hearts and Minds? Evidence from a Field Experiment in Afghanistan.” Political Science Working Paper no. 2011-14, MIT.
- Bennett, Scott. 2008. “Governments, Civilians, and the Evolution of Insurgency: Modeling the Early Dynamics of Insurgencies.” *Journal of Artificial Societies and Social Simulation* 11 (4).
- Berman, Eli, Michael Callen, Joseph Felter, and Jacob Shapiro. 2011. “Do Working Men Rebel? Insurgency and Unemployment in Afghanistan, Iraq, and the Philippines.” *Journal of Conflict Resolution* 55:496–528.
- Berman, Eli, Jacob Shapiro, and Joseph Felter. 2011. “Can Hearts and Minds Be Bought? The Economics of Counterinsurgency in Iraq.” *Journal of Political Economy* 119:766–819.
- Bhavnani, Ravi, Dan Miodownik, and Hyun Jin Choi. 2011. “Three Two Tango: Territorial Control and Selective Violence in Israel, the West Bank, and Gaza.” *Journal of Conflict Resolution* 55 (1): 133–58.
- Blair, Graeme, Christine Fair, Neil Malhotra, and Jacob Shapiro. 2013. “Poverty and Support for Militant Politics: Evidence from Pakistan.” *American Journal of Political Science* 57 (1): 30–48.
- Blair, Graeme, Kosuke Imai, and Jason Lyall. 2014. “Comparing and Combining List and Endorsement Experiments: Evidence from Afghanistan.” *American Journal of Political Science* 58 (4): 1043–63.
- Bullock, Will, Kosuke Imai, and Jacob Shapiro. 2011. “Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan.” *Political Analysis* 19 (4): 363–84.

- Cederman, Lars-Erik, Andreas Wimmer, and Brian Min. 2010. "Why Do Ethnic Groups Rebel? New Data and Analysis." *World Politics* 62 (1): 87–119.
- Chandra, Kanchan, and Steven Wilkinson. 2008. "Measuring the Effect of 'Ethnicity.'" *Comparative Political Studies* 41:515–63.
- Condra, Luke, and Jacob Shapiro. 2012. "Who Takes the Blame? The Strategic Effects of Collateral Damage." *American Journal of Political Science* 56:167–87.
- Cook, S. W. 1971. *The Effect of Unintended Interracial Contact upon Racial Interaction and Attitude Change*. Washington, DC: US Department of Health, Education, and Welfare.
- DeMaio, T. J. 1984. "Social Desirability and Survey Measurement: A Review." In Charles F. Turner and Elizabeth Martian, eds., *Surveying Subjective Phenomena*. Vol. 2. New York: Russell Sage, 257–81.
- Department of the Army. 2007. *U.S. Army Field Manual No. 3-24*. Chicago: University of Chicago Press.
- Epstein, Joshua. 2002. "Modeling Civil Violence: An Agent-Based Computational Approach." *Proceedings of the National Academy of Sciences of the USA* 99:7243–50.
- Fearon, James, and David Laitin. 1996. "Explaining Interethnic Cooperation." *American Political Science Review* 90 (4): 715–35.
- Fearon, James, and David Laitin. 2003. "Ethnicity, Insurgency, and Civil War." *American Political Science Review* 97 (1): 75–90.
- Findley, Michael, and Joseph Young. 2007. "Fighting Fire with Fire? How (Not) to Neutralize an Insurgency." *Civil Wars* 9 (4): 378–401.
- Galula, David. 2006. *Counterinsurgency Warfare: Theory and Practice*. New York: Praeger.
- Greenwald, Anthony G., Debbie E. McGhee, and Jordan L. K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74:1464–80.
- Habyarimana, James, Macartan Humphreys, Daniel Posner, and Jeremy Weinstein. 2009. *Coethnicity: Diversity and the Dilemmas of Collective Action*. New York: Russell Sage.
- Hahn, Adam, Charles M. Judd, Holen K. Hirsh, and Irene V. Blair. 2014. "Awareness of Implicit Attitudes." *Journal of Experimental Psychology* 143 (3): 1369–92.
- Harris, Adam, and Michael Findley. 2014. "Is Ethnicity Identifiable? Lessons from an Experiment in South Africa." *Journal of Conflict Resolution* 58:4–33.
- Horowitz, Donald. 1985. *Ethnic Groups in Conflict*. Berkeley: University of California Press.
- ISAF (International Security Assistance Force). 2010. "Afghans Report, Turn in Explosives, Weapons." Press release, October 12. <http://www.isaf.nato.int/article/isaf-releases/afghans-report-turn-in-explosives-weapons.html>.
- Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.
- Kalyvas, Stathis. 2012. "Micro-Level Studies of Violence in Civil War: Refining and Extending the Control-Collaboration Model." *Terrorism and Political Violence* 24:658–68.
- Kalyvas, Stathis, and Matthew Kocher. 2009. "The Dynamics of Violence in Vietnam: An Analysis of the Hamlet Evaluation System (HES)." *Journal of Peace Research* 46 (3): 335–55.
- Kaplan, Oliver. 2013. "Protecting Civilians in Civil War: The Institution of the ATCC in Colombia." *Journal of Peace Research* 50 (3): 351–67.
- Kocher, Matthew, Thomas Pepinsky, and Stathis Kalyvas. 2011. "Aerial Bombing and Counterinsurgency in the Vietnam War." *American Journal of Political Science* 55:201–18.
- Leites, Nathan, and Charles Wolf. 1970. *Rebellion and Authority: An Analytic Essay on Insurgent Conflicts*. Chicago: Markham.
- Lyall, Jason. 2009. "Does Indiscriminate Repression Incite Insurgent Attacks? Evidence from Chechnya." *Journal of Conflict Resolution* 53:331–62.
- Lyall, Jason. 2010. "Are Co-ethnics More Effective Counterinsurgents? Evidence from the Second Chechen War." *American Political Science Review* 104:1–20.
- Lyall, Jason, Graeme Blair, and Kosuke Imai. 2013. "Explaining Support for Combatants in Wartime: A Survey Experiment in Afghanistan." *American Political Science Review* 107 (4): 679–705.
- Lyall, Jason, and Isaiah Wilson. 2009. "Rage against the Machines: Explaining Outcomes in Counterinsurgency War." *International Organization* 63:67–106.
- Mason, T. David. 1996. "Insurgency, Counterinsurgency, and the Rational Peasant." *Public Choice* 86 (1–2): 63–83.
- Nagl, John. 2005. *Learning to Eat Soup with a Knife: Counterinsurgency Lessons from Malaysia and Vietnam*. Chicago: University of Chicago Press.
- Pajhwok Afghan News. 2011. "One Dead as Sar-i-Pul Residents Clash with Taliban." Pajhwok Afghan News, April 7.
- Petersen, Roger. 2002. *Understanding Ethnic Violence: Fear, Hatred, and Resentment in Twentieth-Century Eastern Europe*. New York: Cambridge University Press.
- Posen, Barry. 1993. "The Security Dilemma and Ethnic Conflict." *Survival* 35:27–47.
- Sambanis, Nicholas. 2001. "Do Ethnic and Nonethnic Civil Wars Have the Same Causes?" *Journal of Conflict Resolution* 45:259–82.
- Shiraito, Yuki, and Kosuke Imai. 2012. "endorse: R Package for Analyzing Endorsement Experiments." Princeton University. <http://imai.princeton.edu/software/endorse.html>.
- Stoll, David. 1993. *Between Two Armies in the Ixil Towns of Guatemala*. New York: Columbia University Press.
- Tajfel, Henri, Michael Billig, R. P. Bundy, and Claude Flament. 1971. "Social Categorization and Intergroup Behavior." *European Journal of Social Psychology* 1:149–78.
- Thompson, Robert. 1966. *Defeating Communist Insurgency*. London: Chatto & Windus.
- Trinquier, Roger. 2006. *Modern Warfare: A French View of Counterinsurgency*. New York: Praeger.
- Vanden Eynde, Oliver. 2013. "Targets of Violence: Evidence from India's Naxalite Conflict." Working paper, Paris School of Economics.
- Weidmann, Nils. 2011. "Violence 'From Above' or 'From Below'? The Role of Ethnicity in Bosnia's Civil War." *Journal of Politics* 73:1178–90.