Supporting information for "Can civilian attitudes predict insurgent violence? Ideology and insurgent tactical choice in civil war"

Kentaro Hirose, Kosuke Imai, and Jason Lyall

The endorsement experiment

Our endorsement experiment uses four questions regarding domestic policy reform to estimate support levels for the Taliban and ISAF. The exact question wording is reproduced below.

Prison reform

- CONTROL CONDITION: A recent proposal calls for the sweeping reform of the Afghan prison system, including the construction of new prisons in every district to help alleviate overcrowding in existing facilities. Though expensive, new programs for inmates would also be offered, and new judges and prosecutors would be trained. How strongly would you support this policy?

- TREATMENT CONDITION: A recent proposal by foreign forces [or the Taliban] calls for the sweeping reform of the Afghan prison system, including the construction of new prisons in every district to help alleviate overcrowding in existing facilities. Though expensive, new programs for inmates would also be offered, and new judges and prosecutors would be trained. How strongly would you support this policy?

Direct election

- CONTROL CONDITION: It has recently been proposed to allow Afghans to vote in direct elections when selecting leaders for district councils. Provided for under Electoral Law, these direct elections would increase the transparency of local government as well as its responsiveness to the needs and priorities of the Afghan people. It would also permit local people to actively participate in local administration through voting and by advancing their own candidacy for office in these district councils. How strongly would you support this policy?

- TREATMENT CONDITION: It has recently been proposed by foreign forces [or the Taliban] to allow Afghans to vote in direct elections when selecting leaders for district councils. Provided for under Electoral Law, these direct elections would increase the transparency of local government as well as its responsiveness to the needs and priorities of the Afghan people. It would also permit local people to actively participate in local administration through voting and by advancing their own candidacy for office in these district councils. How strongly would you support this policy?

Independent election commission

- CONTROL CONDITION: A recent proposal calls for the strengthening of the Independent Election Commission (IEC). The Commission has a number of important functions, including monitoring presidential and parliamentary elections for fraud and verifying the identity of candidates for political office. Strengthening the IEC will increase the expense of elections and may delay the announcement of official winners but may also prevent corruption and election day problems. How do you feel about this proposal?

- TREATMENT CONDITION: A recent proposal by foreign forces [or the Taliban] calls for the strengthening of the Independent Election Commission (IEC). The Commission has a number of important functions, including monitoring presidential and parliamentary elections for fraud and verifying the identity of candidates for political office. Strengthening the IEC will increase the expense of elections and may delay the announcement of official winners but may also prevent corruption and election day problems. How do you feel about this proposal?

Anti-corruption reform

- CONTROL CONDITION: It has recently been proposed that the new Office of Oversight for Anti- Corruption, which leads investigations into corruption among government and military officials, be strengthened. Specifically, the Offices staff should be increased and its ability to investigate suspected corruption at the highest levels, including among senior officials, should be improved by allowing the Office to collect its own information about suspected wrong-doing. How do you feel about this policy?

- TREATMENT CONDITION: It has recently been proposed by foreign forces [or the Taliban] that the new Office of Oversight for Anti- Corruption, which leads investigations into corruption among government and military officials, be strengthened. Specifically, the Offices staff should be increased and its ability to investigate suspected corruption at the highest levels, including among senior officials, should be improved by allowing the Office to collect its own information about suspected wrong-doing. How do you feel about this policy?

The statistical model for the endorsement experiment

Following Bullock, Imai & Shapiro (2011), we use a statistical model to estimate support levels for ISAF and the Taliban by efficiently combining the responses to multiple endorsement experiment questions. To do so, we model each respondent's answer to a policy question as a function of his or her support for the endorser as well as policy preference. Specifically, we apply the following Bayesian ordered probit factor analytic model:

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

where $Y_{ij} \in \{1, 2, 3, 4, 5\}$ represents respondent i's answer to the jth policy question (1 = Strongly disagree, 2 = Disagree, 3 = Indifferent, 4 = Agree, and 5 = Strongly agree) and respondent i's status regarding the randomized treatment assignment is denoted as $T_i \in \{0, 1, 2\}$ (0 = Control, 1 = ISAF, and 2 = Taliban). The latent variable s_{ijk} measures respondent i's support level for endorser k in policy j with a greater value of s_{ijk} indicating a higher level of support. For identification, s_{ij0} is fixed at zero. Finally, the latent variable x_i represents the degree to which respondent i is in favor of policy reform in general. The 'item difficulty' parameter α_{jl} measures the popularity of the jth policy reform independent of the endorser, while the 'discrimination' parameter β_j expresses the degree to which the reform proposal differentiates pro- and anti-reform respondents. We assume $\alpha \sim \mathcal{TN}_{[0,\infty]}(0, 25)$ and $\beta \sim \mathcal{TN}_{[0,\infty]}(0, 25)$ as the priors.

We model the individual-level support s_{ijk} and ideal point x_i using a hierarchical modeling technique with village-level random effect parameters $\lambda_{village[i]}$ and $\delta_{village[i]}$ as follows,

$$s_{ijk} \quad \sim \quad \mathcal{N}(\lambda_{village[i]} + Z_i^\top \lambda_k^Z, \omega_k^2) \tag{2}$$

$$x_i \quad \sim \quad \mathcal{N}(\delta_{village[i]} + Z_i^\top \delta^Z, 1) \tag{3}$$

where Z_i represents the set of individual-level covariates. As the priors, we assume $\lambda \sim \mathcal{N}(0, \psi^2)$, $\delta \sim \mathcal{N}(0, \sigma^2)$, and $\psi^2, \sigma^2, \omega^2 \sim Inv - \chi^2(5, 2)$,

We use an R package `endorse` developed by Shiraito & Imai (2012) to fit this model. The convergence is monitored by running multiple Markov chains with over-dispersed starting values. Using the posterior simulation draws, we compute each respondent's average support level for each endorser across the four policy areas, and then further aggregate it to village-level support by averaging the individual-level estimates.

Additional descriptive analyses

[Figure 6 about here.]

[Table 1 about here.]

[Figure 7 about here.]

[Figure 8 about here.]

[Figure 9 about here.]

[Figure 10 about here.]

					Violence before survey (March – December 2010)			Violence after survey (March – December 2011)		
Provinces	Surveyed villages	Non-surveyed villages	ANSF/ISAF bases	Aid projects	IED attack	IED found	Non-IED attack	IED attack	IED found	Non-IED attack
Logar	19.6%	14.2%	13.3%	14.3%	5.1%	3.7%	4.9%	6.5%	5.8%	11.3%
Kunar	14.7	18.6	31.3	48.3	3.9	1.5	15.9	3.5	1.8	21.1
Helmand	29.9	35.9	22.1	13.2	79.7	78.2	72.0	72.1	75.7	57.6
Uruzgan	13.7	11.4	12.4	9.6	5.3	10.4	2.9	9.1	9.2	3.3
Khost	22.1	19.7	20.7	14.3	5.8	6.0	4.0	8.6	7.2	6.5
Total counts	204	4,225	217	279	2,765	3,577	12,841	2,674	3,979	9,006

Table I. Distribution of villages, violence, ANSF and ISAF bases, and aid projects. The second and third columns, respectively, indicate the proportion of surveyed and non-surveyed villages in each of the randomly selected five provinces. The fourth and fifth columns present the proportion of counterinsurgent bases that were present at the time of our survey and aid projects started before the survey. The last six columns show the percentage of insurgent attacks in each province during 10 months before and after the survey. The last row gives total counts. In our out-of-sample prediction, we use the 13,381 non-surveyed villages in other Pashtun dominated provinces as well as the 4,225 non-surveyed villages in the randomly selected five provinces shown in the table.

Additional in-sample regression results

[Figure 11 about here.]

[Figure 12 about here.]

[Figure 13 about here.]

[Figure 14 about here.]

[Figure 15 about here.]

[Figure 16 about here.]

[Figure 17 about here.]

[Figure 18 about here.]

[Figure 19 about here.]

[Figure 20 about here.]

[Figure 21 about here.]

[Figure 22 about here.]

Additional out-of-sample prediction results

[Figure 23 about here.]

[Figure 24 about here.]

[Figure 25 about here.]

[Figure 26 about here.]

	In-sample regression			Out-of-sample prediction		
	IED attack	IED found	Non-IED attack	IED attack	IED found	Non-IED attack
Support	2.86	1.33	1.97	16.36%	2.21%	10.95%
Base	0.08	0.04	1.12	0.55	0.56	3.00
Aid	-0.53	-0.15	0.30	0.07	-0.11	-0.42
Violence	38.39	15.94	10.67	318.36	265.86	121.48

Table II. t-statistics of in-sample regression and out-of-sample forecasting improvement rates averaged over a wide range of temporal (1-10 months) and spatial (1 - 60km) window sizes. The former three columns show average t-values of each covariate, which are derived from in-sample regression models explaining the future level of insurgent violence by the relative level of ISAF support, the number of bases, the number of aid projects, and the prior level of insurgent violence. The latter three columns indicate average forecasting improvement rates: the first row corresponds to average improvement rates of adding relative ISAF support to a baseline model with bases, aid projects, and past insurgent violence; the second row corresponds to average improvement rates of adding base to a baseline model with support, aid, and past violence; the third row corresponds to average improvement rates of adding aid to a baseline model with support, base, and past violence; and the fourth row corresponds to average improvement rates of adding past violence to a baseline model with support, base, and aid.

Insurgent violence against civilians

Our theory suggests a second observable implication: insurgents use pro-counterinsurgent attitudes to guide their punishment strategy against civilians. To test this claim, we draw on a new dataset of civilian violence—SIGACTs track only violence against ISAF—and repeat the analysis above, beginning with in-sample testing before moving to out-of-sample predictions. We use data from iMMAP, a non-governmental organization that collates reports of civilian victimization across Afghanistan; data are provided by multiple NGOs and are inputted according to a standardized coding scheme via online data portal.

The dataset represents the best coverage of civilian victimization in Afghanistan to date. It does, however, have two major shortcomings: (1) it is noisier than SIGACT data given its multiple reporting streams and (2) the data suffer from a clear under-reporting problem, yielding far fewer recorded instances of violence against civilians than attacks against ISAF recorded in SIGACT as shown in Table IV in the SI. Together, these limitations suggest that our findings should be interpreted with caution.

We begin by plotting the association between relative ISAF support and two categories of insurgent violence: attacks with IEDs and those without.[15] As an initial exploration of the data, Figure 27 plots this relationship using a negative binomial regression model for attacks within 15km of each village in the five months after our survey. As with SIGACT data, we find a positive association between relative ISAF support and both types of insurgent attacks even after adjusting for prior insurgent attacks, base locations, and aid programs. Substantively, given the low level of civilian victimization reported in the iMMAP data, the effects are fairly modest; villages at the high end of relative ISAF support will observe about three additional reported attacks within these temporal and spatial boundaries. These findings are consistent, however, with the claim that the Taliban are willing to punish civilians for their pro-ISAF views.

Our in-sample estimates of predicted insurgent violence against civilians largely confirm the findings above. In Figure 28, we plot the t values of the association between relative ISAF support and insurgent attacks up to 60km and 10 months after the survey. For IED attacks, the association is strongest at close proximity to the village; at that distance, the positive association extends almost evenly across the 10 month post-survey period. We do note, however, a weakening of our predictive improvement at the spatial mid-range (about 30–40km). We conjecture that this weakening can be attributed to two factors: iMMAP privileges data collection on IEDs that occur within a village and rarely collects data outside of populated locations; and the average Taliban group's operational radius is about 30–40km (see below). Non-IED attacks are strongly associated with relative ISAF support, especially relatively close to the village (at the 20-30 km mark). Taken together (the right panel), relative ISAF support is associated with attacks against civilians, particularly close to the surveyed village, a finding consistent with a punishment strategy.

Our out-of-sample forecasting based on a negative binomial model also demonstrates that including civilian attitudes significantly improves predictive performance, though the results are weaker for IEDs than non-IED attacks (see Figure 29 in the SI). For IED attacks, we find a similar pattern to our in-sample prediction: improvement is highest when violence is closest to the village at all intervals up to 10 months post-survey, though we do observe a weakening of this improvement at the 30–40km range. Once again, our predictive improvement is higher for non-IED attacks against civilians; improvement is significantly higher close to the village and extending up to 40km away. Similar to our SIGACT-based estimates, predictive performance for the pooled insurgent attacks is highest within the 10km range.

[Table 2 about here.]

[Figure 27 about here.]

[Figure 28 about here.]

[Figure 29 about here.]

External generalizability

Our out-of-sample testing helps assuage concerns about generalizability. But there are natural limits to single-country studies. Do our claims about civilian attitudes and insurgent violence travel to other contexts? There are at least three ways to assess generalizability here: as a function of Afghanistan's specific properties; the war's characteristics; and the nature of Taliban organization.

Drawing on Fearon & Laitin (2003), we first plot Afghanistan's location in the distribution of all civil wars (1945–1999) across six characteristics: per capita income, population, mountainous terrain, the regime's polity score, and ethnic and religious fractionalization. As Figure 30 illustrates, Afghanistan is not an outlier in any distribution.

Nor is the war itself an outlier; it shares many properties of long-running insurgencies since 1945. For example, at least one-quarter of all insurgencies since 1945 have witnessed armed intervention by a third-party counterinsurgent like ISAF (Lyall & Wilson, 2009). Most of these counterinsurgency efforts—including prominent examples in Pakistan, Iraq, Colombia, Mexico, Yemen, and the Philippines—have included extensive 'hearts and minds' campaigns to win over public support. And while crossnational data on civilian victimization is poor, the current war in Afghanistan is not an outlier in terms of the magnitude of civilian deaths. The best public estimates suggest that 2,000—3,000 civilians are killed each year, most by the Taliban (United Nations Assistance Mission Afghanistan, 2011). If our results hinge on expectations of violence among civilians, then the relatively low level of civilian fatalities in Afghanistan suggest that our predictive improvement would be even higher in conflicts with (even) higher casualties, as in contemporary Syria.[16]

Finally, there is little unique about the Taliban's intelligence-gathering institutions or its governance project. Other rebel organizations, including Islamic State, FARC, LTTE, Hezbollah, and RCD, have constructed extensive intelligence networks and have sought to engage in systematic hearts and minds or governance campaigns. In fact, most insurgent organizations provide at least some basic services; the Taliban's provision of local dispute adjudication is quite typical. But if civilian attitudes are irrelevant for combatants—i.e., the insurgent organization simply preys upon locals—then we anticipate that other considerations, notably the strategic nature of territory or the presence of lootable commodities, would likely trump attitudes in guiding targeting.

[Figure 30 about here.]

Qualitative evidence

Our in-sample and out-of-sample tests converge on the same finding: including civilian attitudes in our models markedly improves our ability to predict insurgent violence. Yet we require evidence of the mechanics of Taliban intelligence gathering to be confident of the link between civilian attitudes and insurgent targeting. Can the Taliban actually track civilian attitudes with reasonable precision across potentially thousands of villages? And, if so, do its commanders act upon this intelligence when choosing targets and tactics?

While information about the Taliban is necessarily incomplete, there is near universal agreement among researchers and ISAF itself that it possesses a remarkably pervasive intelligence network. ISAF's own Deputy Chief of Staff (Intelligence) publicly declared the need to drastically overhaul ISAF's intelligence collection to compete with Taliban efforts (Flynn, Pottinger & Batchelor, 2010). Similarly, a leaked classified NATO report based on nearly 27,000 interviews with 4,000 detainees in 2011 painted a stark picture of an omnipresent Taliban that had spies on ISAF bases, subverted local ANSF partners, and moved freely among locals in nominally ISAF-controlled villages (Task Force 3-10, 2012).

The Taliban has constructed an extensive surveillance system to support its village-level governance and war-fighting efforts. Each province has a shadow provincial governor and a military commission (plural, *nizami*); these institutions are designed to coordinate, albeit loosely, the implementation of governance programs such as Taliban-run courts while orchestrating attacks against ISAF and ANSF by small, locally-recruited, units (*delgai*) of 15-25 fighters. In theory, these units 'scale-up' to form *mahaz* networks that control parts of a given district. In practice, coordination across these units is often haphazard, owing as much to ISAF efforts at disruption as local ethnic, tribal, strategic, and other disagreements that frustrate broader cooperation (Farrell & Giustozzi, 2013; Giustozzi, 2013; Johnson, 2013; Johnson & DuPee, 2012).

The Taliban collect information about both ISAF movements and civilian attitudes via four principal mechanisms. First, the Taliban have cultivated a network of local supporters who not only provide material assistance but also information about popular opinion and troop movements. In the latter case, everything from smoke signals to information passed via radio repeater stations into Pakistan have been used to collect and disseminate timely information about ISAF actions. In one of our out-of-sample districts, Andar, located in Ghazni province, ISAF estimated that 4,000 locals (out of 150,000 individuals) actively provided information to the Taliban. Indeed, the Taliban were openly known to be running 28 schools in the district despite ISAF's heavy presence (C.J. Chivers, 'In Eastern Afghanistan, at War with the Taliban's Shadowy Rule,' *New York Times*, 6 February 2011: A1.).

Second, Taliban spies and sympathizers have infiltrated ISAF and ANSF bases as well as large-scale development projects. 'The Taliban,' one Afghan National Police officer noted in 2011, 'have spies everywhere.'[17] Even ISAF has conceded that its facilities and, in particular, Afghan security forces, have been compromised by spies (Task Force 3-10, 2012). ANSF units have been caught colluding openly with Taliban forces via local (unauthorized) ceasefire arrangements and through the provision of information about ISAF's movements and security. 'They are on top of every move we make,' one ISAF official bemoaned in 2010 (Brandt, 2011: 1). Extensive back-to-work projects, designed to dampen insurgency by providing steady employment for would-be fighters, are thought especially valuable to penetrate. These programs allow the Taliban to pose as wage laborers and unobtrusively gather information about village reactions toward ISAF and its 'hearts and minds' efforts.

Third, the Taliban's own efforts at providing limited governance represent a 'dragnet' that gathers intelligence about local attitudes from multiple sources. These efforts including: (1) collecting taxes (*zakat*) from villagers; (2) regular meetings with village elders and religious officials; (3) roadside checkpoints that provide opportunities to monitor the population; (4) the creation of complaints commissions that villagers can access for local dispute adjudication, including reporting corrupt or unruly Taliban commanders; and (5) surveillance at Friday prayers, which are typically attended by all males in a village.

As ISAF itself concluded: 'Villagers commonly relay that the Taliban are continually present in their areas solving disputes, purchasing supplies at local bazaars, meeting with tribal leaders or staying overnight in guesthouses or the local mosque' (Task Force 3-10, 2012: 6). These initiatives collectively facilitate the collection of timely information about local attitudes towards the combatants, among other topics.

Finally, by 2010 the Taliban had come to rely heavily on local recruits to staff its units. It has been estimated that between 80-90% of Taliban fighters operate in or close to their communities. Moreover, launching attacks across *mahaz* lines is difficult, requiring permission from both neighboring commanders and from the relevant centralized military commission (Ruttig, 2010: 13). Despite recent efforts to improve the centralization of decision-making, each group retains the authority to conduct attacks within its operating area without prior permission. As a result, the Taliban not only able to obtain local information but are likely to launch attacks in fairly constrained geographic areas.

The Taliban clearly have an impressive, if imperfect, ability to monitor civilian attitudes, one that is likely more sophisticated and extensive than ISAF's own efforts. Does this information inform their choice of targets and tactics? Undoubtedly, yes. It would be odd if the Taliban devoted these resources to surveillance, and to rebuilding them after ISAF counterintelligence operations, only to disregard this information when selecting targets.

The Taliban has drawn on this information to launch an extensive intimidation campaign using 'night letters' affixed to individual's doors warning against continued collaboration with ISAF, for example (United Nations Assistance Mission Afghanistan, 2011; Gopal, 2014). Since 2009, the Taliban has engaged in a systematic assassination campaign that has killed hundreds, including government officials, religious leaders, and individuals who have assisted ISAF in some capacity (notably, as translators or informants). As one US company commander stationed in Ghazni, an out-of-sample province, recalled: 'The guy we had who was willing to give us information about the Taliban is the guy we found dead last week' after he was pulled from his vehicle and executed.[18]

This discussion helps contextualize our argument that insurgents use pro-counterinsurgent attitudes as targeting cues in four ways. First, the Taliban clearly has the means to track ISAF force movements in near-real time, suggesting that striking at ISAF is feasible and a central war-fighting aim. Second, this intelligence system has the capacity to track civilian attitudes; it is realistic to assume that the Taliban is well-informed about local attitudes, not least because the bulk of its fighters are drawn from the same populations they seek to monitor. Third, the Taliban has adapted over time to become more lethal, focusing specifically on IEDs and suicide bombings that maximize the lethality of attacks against ISAF. Fourth, they have shown little hesitation in killing (suspected) informants, suggesting a willingness to consider targeting civilians. These attacks demonstrate that the Taliban can harm the counterinsurgent while also sending a message to would-be ISAF supporters: continue to support the counterinsurgent, and face punishment in the form of increased (indiscriminate) violence (Lyall, 2015). Finally, our in- and out-of-sample findings are strongest at the 25-40 km distance from villages; these distances fit with the local nature of Taliban recruitment and the typical operating radius of its units. In short, pro-ISAF views are an invitation for, rather than a shield against, future localized insurgent violence.

References

Brandt, Ben (June 2011) The Taliban's conduct of intelligence and counterintelligence. *CTC-Sentinel* 4(6): 19–23.

Farrell, Theo & Antonio Giustozzi (2013) The Taliban at war: Inside the Helmand insurgency, 2004-2012. *International Affairs* 89(4): 845–871.

Flynn, Michael, Matt Pottinger & Paul Batchelor (2010) *Fixing intel: A blueprint for making intelligence relevant in Afghanistan*. Washington, D.C.: Center for a New American Security.

Giustozzi, Antonio (2013) Military adapation by the Taliban, 2001-2011. In *Military Adaptation in Afghanistan*, ed. Theo Farrell, Frans Osinga & James Russell. Stanford University Press pp.242–62.

Gopal, Anand (2014) *No Good Men Among the Living: America, the Taliban, and the War Through Afghan Eyes*. New York: Macmillan.

Johnson, Thomas (2013) Taliban adaptations and innovations. *Small Wars and Insurgencies* 24(1): 3–27.

Johnson, Thomas & Matthew DuPee (2012) Analyzing the new Taliban code of conduct (Layeha): An assessment of changing perspectives and strategies of the Afghan Taliban. *Central Asian Survey* 31(1): 77–91.

Meyerle, Jerry & Carter Malkasian (August 2009) Insurgent tactics in southern Afghanistan: 2005-2008. Technical Report CRM D0020729.A2/Final (http://nsarchive.gwu.edu/NSAEBB/NSAEBB370/docs/Document%205.pdf).

Task Force 3-10 (2012) *State of the Taliban (January 6, 2012): Detainee Perspectives*. Bagram, Afghanistan: NATO (http://s3.documentcloud.org/documents/296489/taliban-report.pdf).

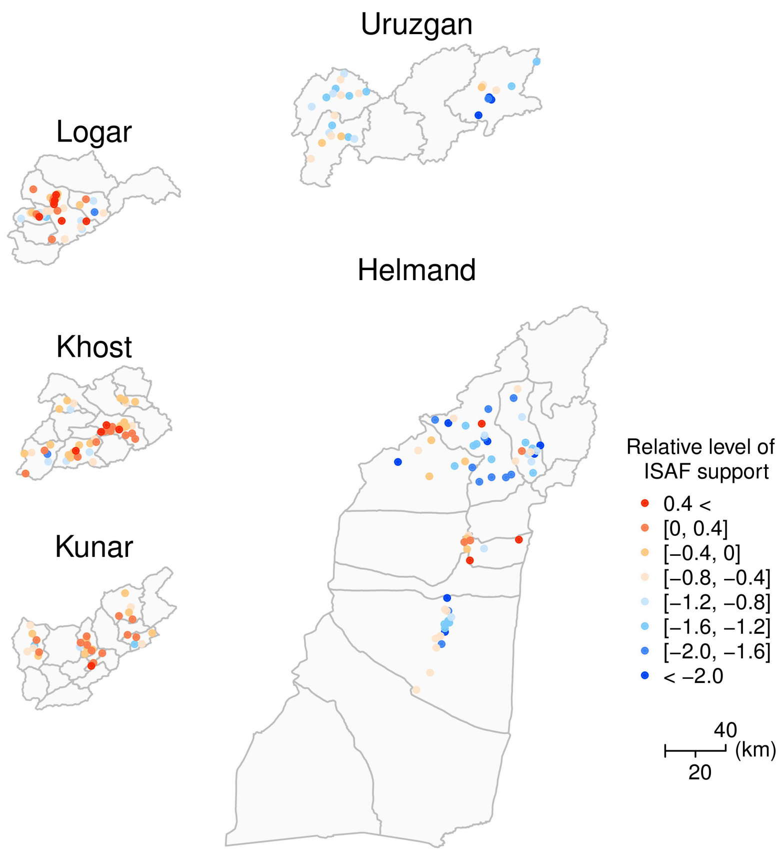

Figure 1. Spatial distribution of relative support for International Security Assistance Force (ISAF) in the 204 surveyed villages from five randomly sampled Pashtun-dominated provinces. The warmer color (red) represents villages which are more supportive of ISAF than the Taliban; the colder color (blue) indicates less supportive villages.

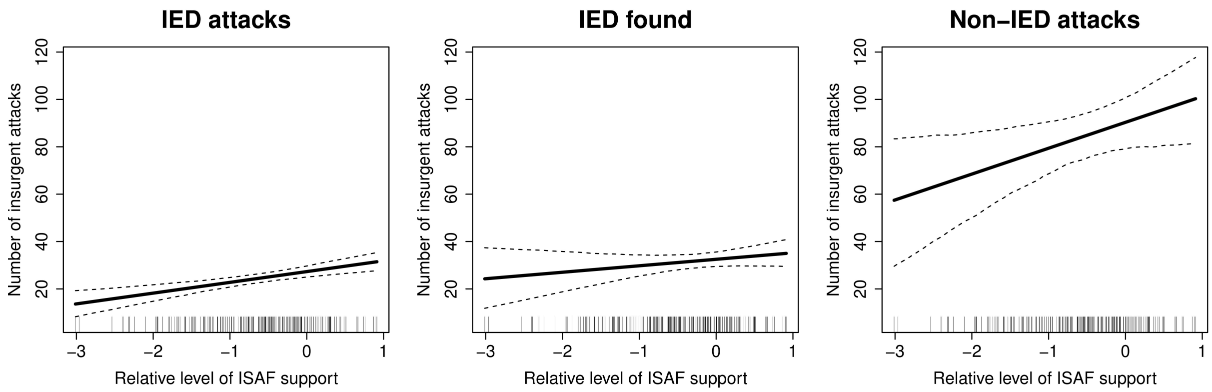

Figure 2. Positive association between the number of future insurgent attacks and relative ISAF support. The plots illustrate the statistically significant association between the number of insurgent attacks that have occurred within 15km of each village during the five months *after* the survey in each village (vertical axis) and its relative level of ISAF support (horizontal axis) while adjusting for the number of insurgent attacks that have occurred (again within a 15km around each village) five months *prior to* our survey, the number of counterinsurgent bases within a 3km radius of each village, and the number of aid programs within a 1km radius of each village. The results are based on the linear regression model estimated separately for each of the three violence categories where the number of future insurgent attacks is regressed on the relative level of ISAF support and the other covariates. The dashed lines represent 95% confidence intervals based on robust standard errors.

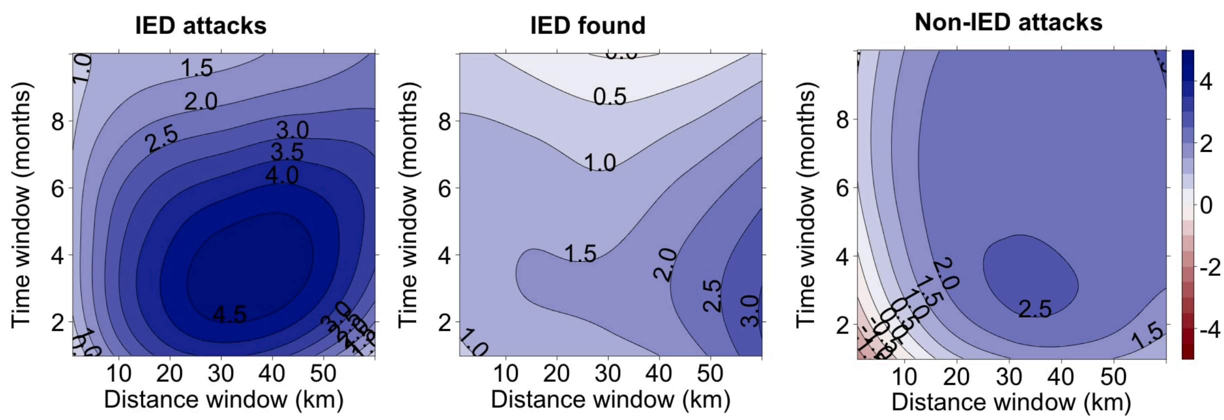

Figure 3. t-statistic of the estimated coefficient for the relative level of ISAF support across a wide range of temporal and spatial window sizes. The dark blue areas represent large values of t-statistics. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks while adjusting for prior insurgent violence and the number of ANSF/ISAF bases and USAID aid projects. The linear regression models which produced the results displayed in Figure 2 are repeatedly estimated using broad time and distance windows (from 1 to 10 months before/after the survey and from 1 to 60km of each village's perimeter). The results illustrate the robustness of the positive association between the number of future insurgent attacks (especially with IEDs) and relative ISAF support.

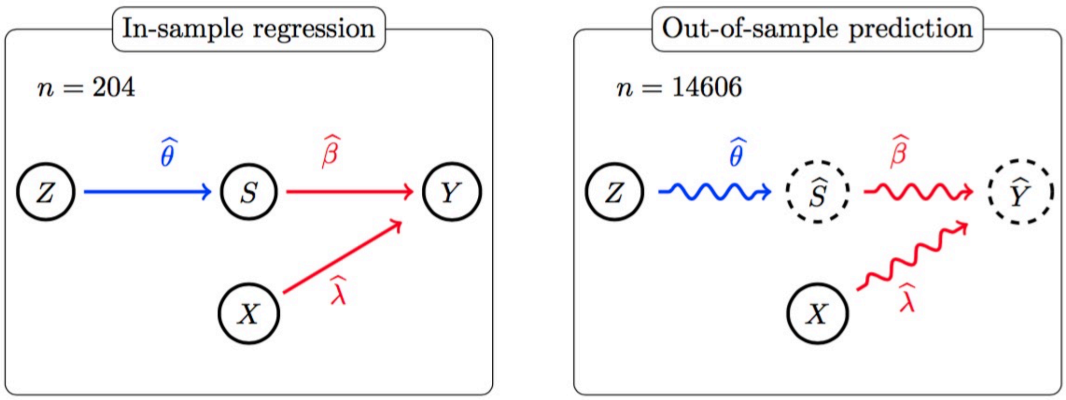

Figure 4. Out-of-sample prediction procedure. In Step 1 (left panel), using in-sample villages, we estimate the *ISAF support model* (blue arrow) and the *Insurgent violence model* (red arrow). For the former model, we regress ISAF relative support S on village- and district-level covariates Z such as log population, log elevation, ISAF control, Pakistan border, and Taliban Sharia. For the latter model, we regress future violence Y on S as well as village-level control variables (X) such as past violence, ANSF/ISAF bases, and aid projects. In Step 2 (right panel), we predict the ISAF support level for out-of-sample villages and then future violence using the models fitted in Step 1. For each out-of-sample village, we first estimate the ISAF relative support level S (blue squiggly arrow) using the covariates Z and then predict Y with this estimated support level $\widehat{S}$ and the other covariates X.

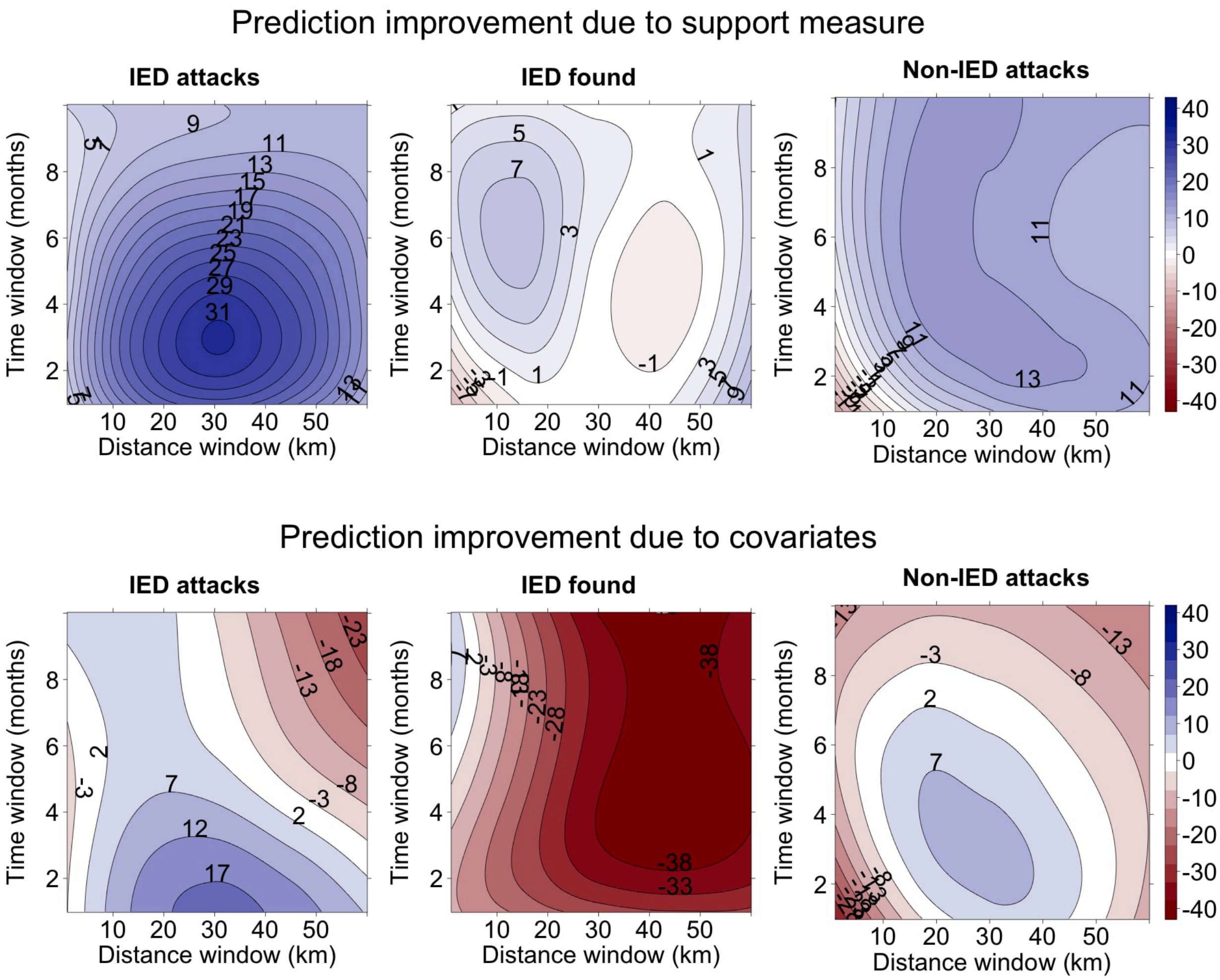

Figure 5. Out-of-sample forecasting performance. The upper panels depict forecasting improvement rates from adding estimated ISAF relative support level to the baseline model with past violence, counterinsurgent bases, and aid projects. Prediction improvement is measured by mean absolute forecasting errors derived from the baseline model (MAFE_2) and the model with the estimated support level (MAFE_1) — i.e., $(\text{MAFE}_2 - \text{MAFE}_1)/\text{MAFE}_1 \times 100\%$. The lower panels depict forecasting improvement rates from adding village- and district-level covariates to the baseline model.

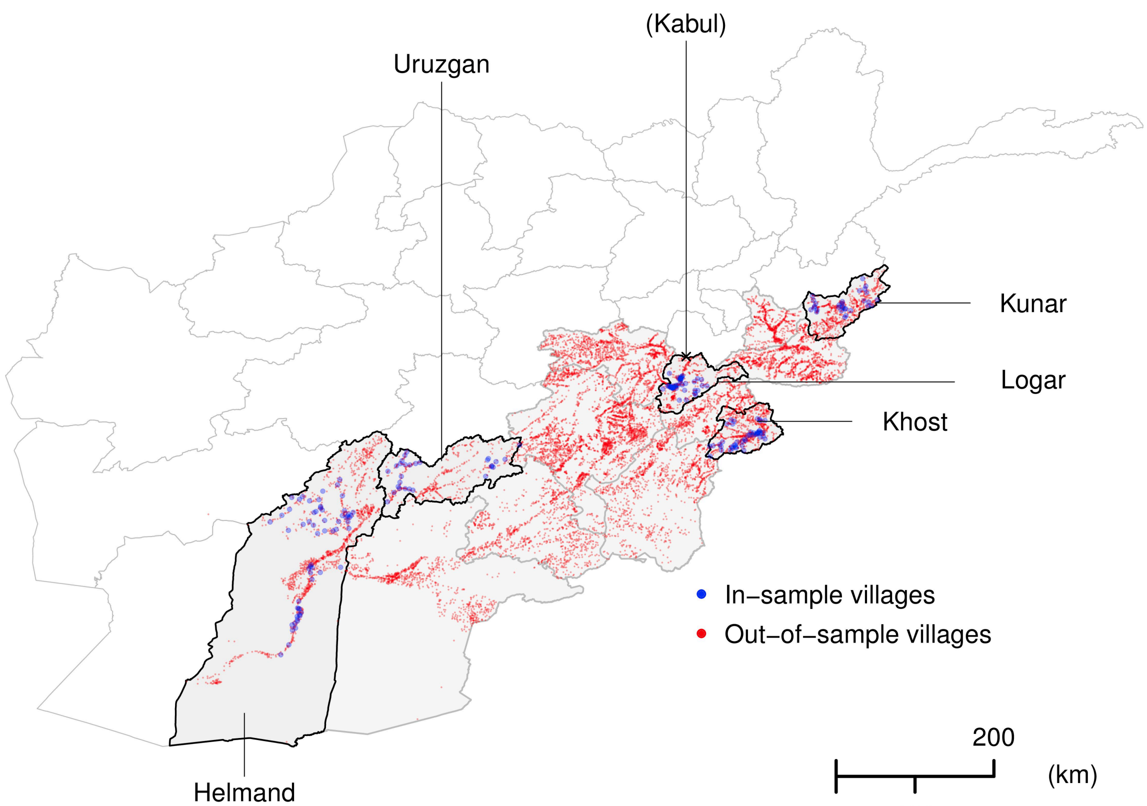

Figure 6. Location of 204 surveyed (in-sample) and 14,606 non-surveyed (out-of-sample) villages in five randomly-chosen Pashtun-dominated provinces (gray areas) of Afghanistan.

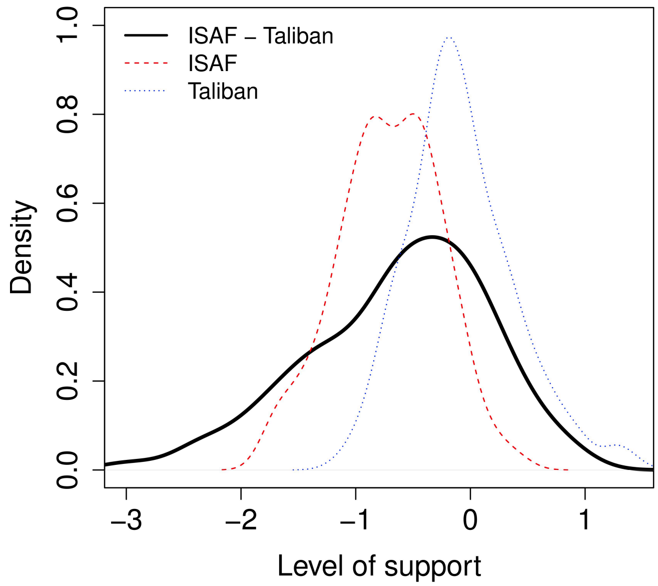

Figure 7. Distribution of the levels of support for Taliban and ISAF as well as the distribution of the difference between the two (ISAF–Taliban). The figure clearly shows that a majority of Afghan population supports Taliban over ISAF.

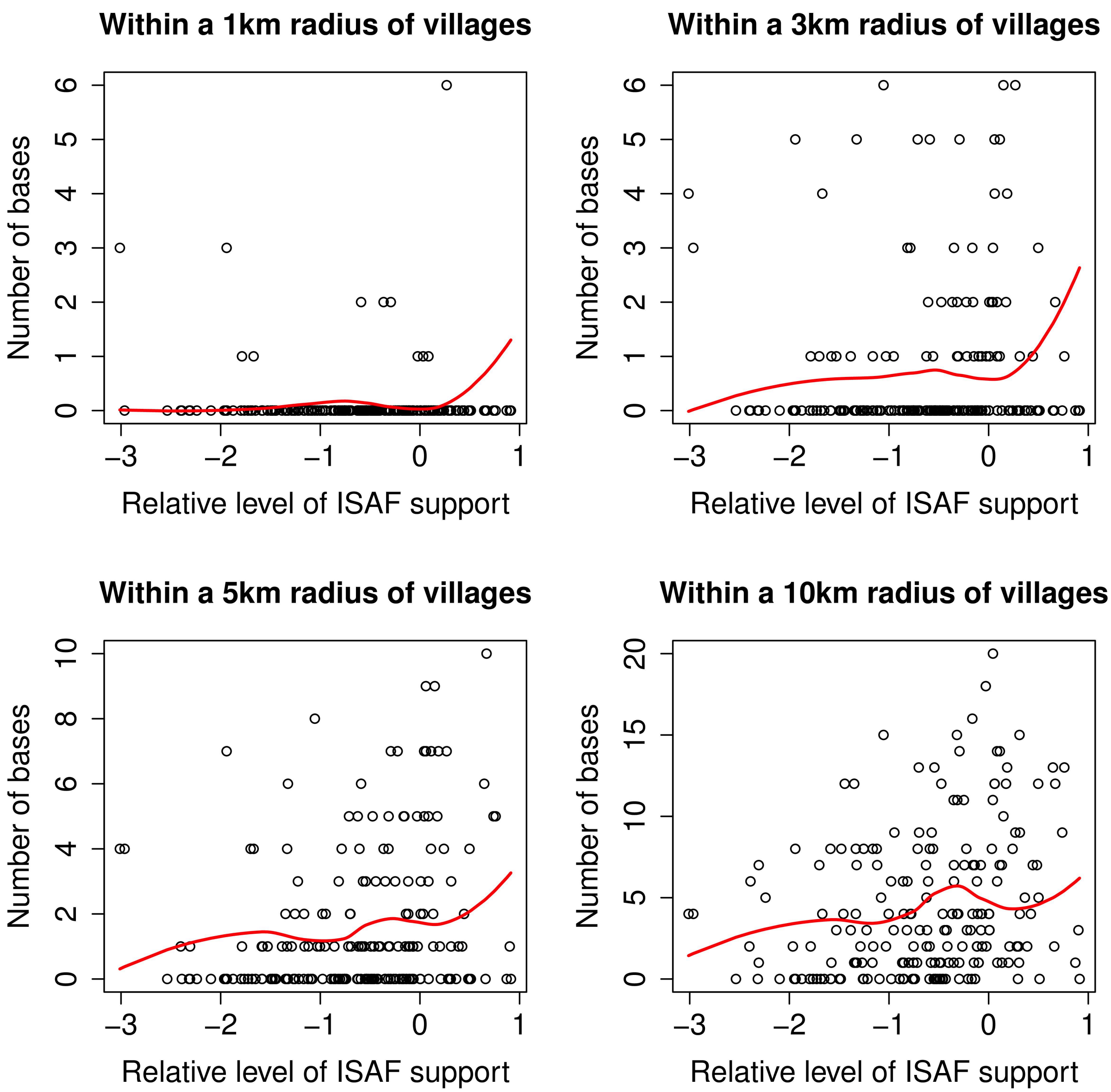

Figure 8. Bivariate relationship between relative ISAF support and the number of ANSF and ISAF bases within a 1km, 3km, 5km, or 10km radius of each surveyed village.

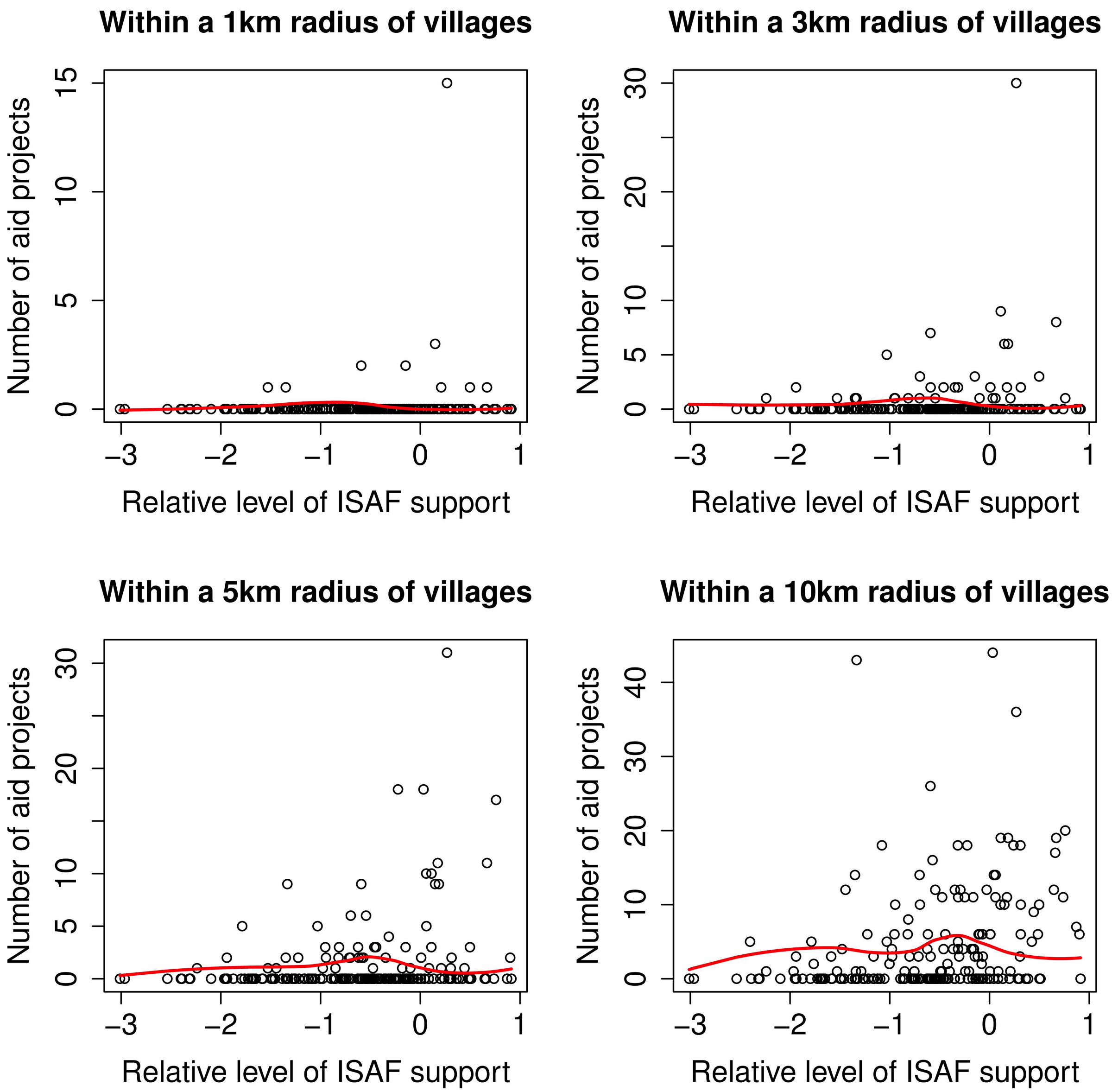

Figure 9. Bivariate association between relative ISAF support and the number of aid projects within a 1km, 3km, 5km, or 10km radius of each surveyed village.

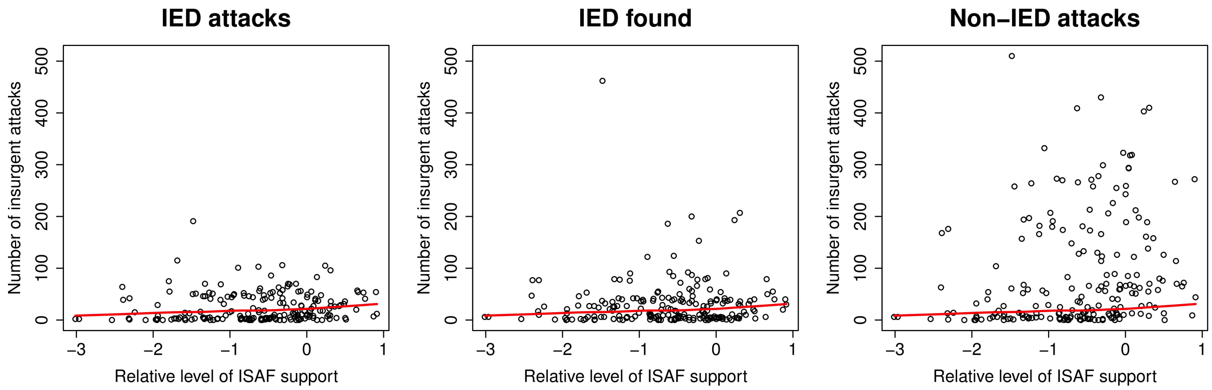

Figure 10. The plots illustrate positive correlations between the number of insurgent attacks that have occurred within 15km of each village during the five months *after* the survey in each village (vertical axis) and its relative level of ISAF support (horizontal axis). Pearson's correlation coefficients are, respectively, 0.08 (IED attacks), 0.04 (IED found), and 0.21 (Non-IED attacks).

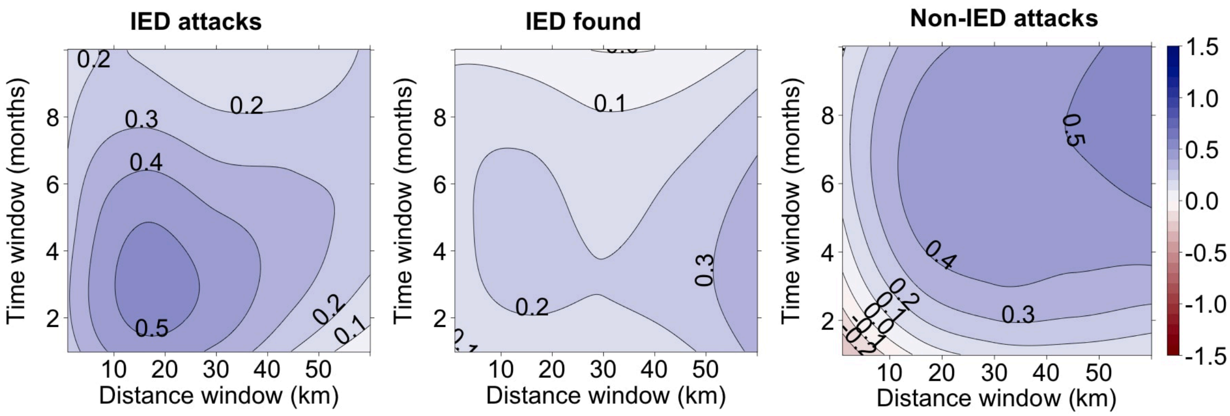

Figure 11. Standardized effects of the *relative level of ISAF support* across a wide range of temporal and spatial window sizes. The effect sizes are [1] computed by taking the differences between the predicted numbers of future attacks when the relative level of ISAF support takes the maximum and minimum values, respectively, and then [2] standardized by the standard deviations of the dependent variables. A linear regression model is used by adjusting for the prior level of insurgent violence, the number of bases, and the number of aid projects.

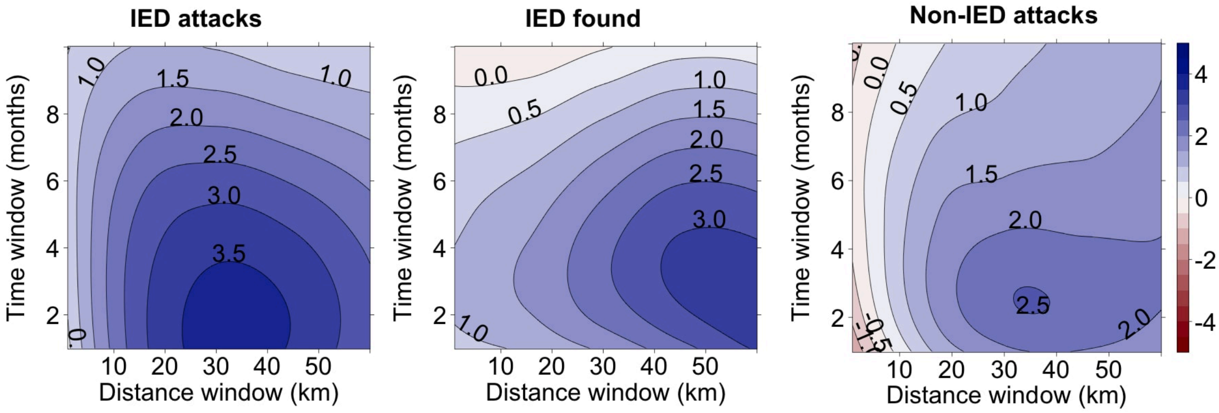

Figure 12. t values of the estimated coefficient for the *absolute* level of ISAF support. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks, while adjusting for the prior level of insurgent violence, the number of ANSF and ISAF bases, and the number of USAID aid projects. Robust standard errors are used.

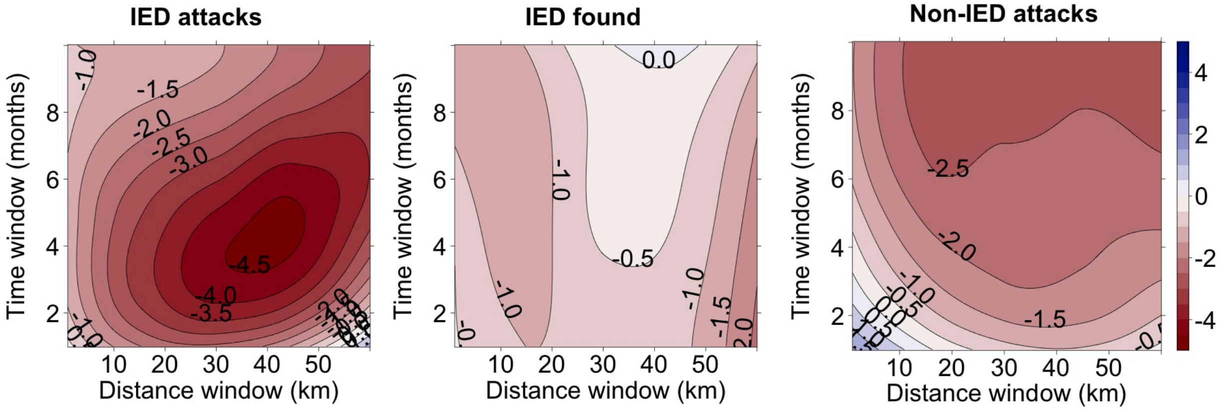

Figure 13. t values of the estimated coefficient for the *absolute* level of Taliban support. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks, while adjusting for the prior level of insurgent violence, the number of ANSF and ISAF bases, and the number of USAID aid projects. Robust standard errors are used.

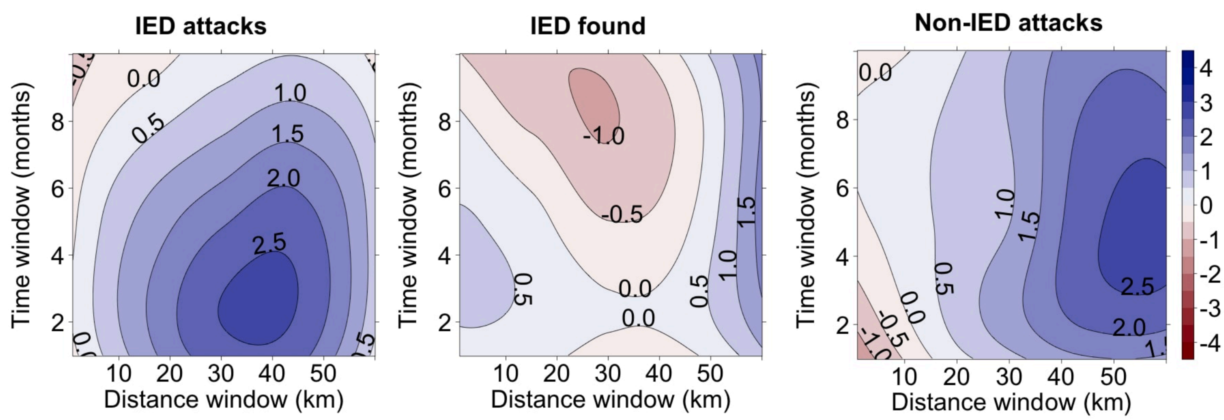

Figure 14. t-statistics of the estimated coefficient of relative ISAF support derived from the Mahalanobis matching analysis. Villages are first paired according to prior violence, bases, and aid projects, and then the pairwise difference in future violence is regressed on the pairwise differences in relative ISAF support, prior violence, bases, and aid projects. Robust standard errors are used.

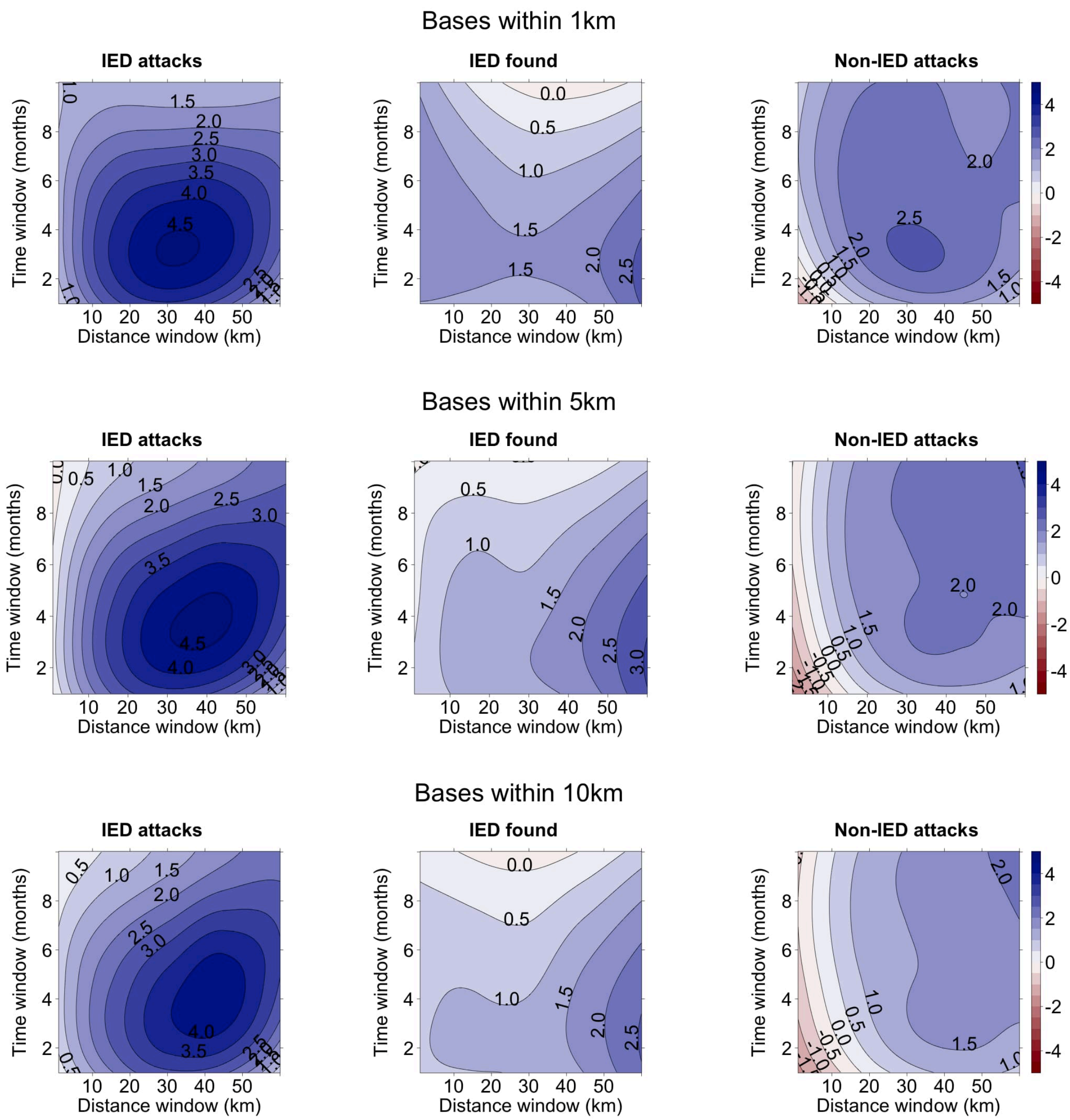

Figure 15. t values of the estimated coefficient of relative ISAF support, corresponding to its marginal effect on the number of future insurgent attacks, while adjusting for the prior level of insurgent violence, the number of bases, and the number of aid projects. The upper, middle, and lower panels use bases within 1km, 5km, and 10km radii of each sampled village, respectively, showing robustness of the positive association between the number of future insurgent violence and relative ISAF support. All the models use aid projects within a 1km radius of each village. Robust standard errors are used.

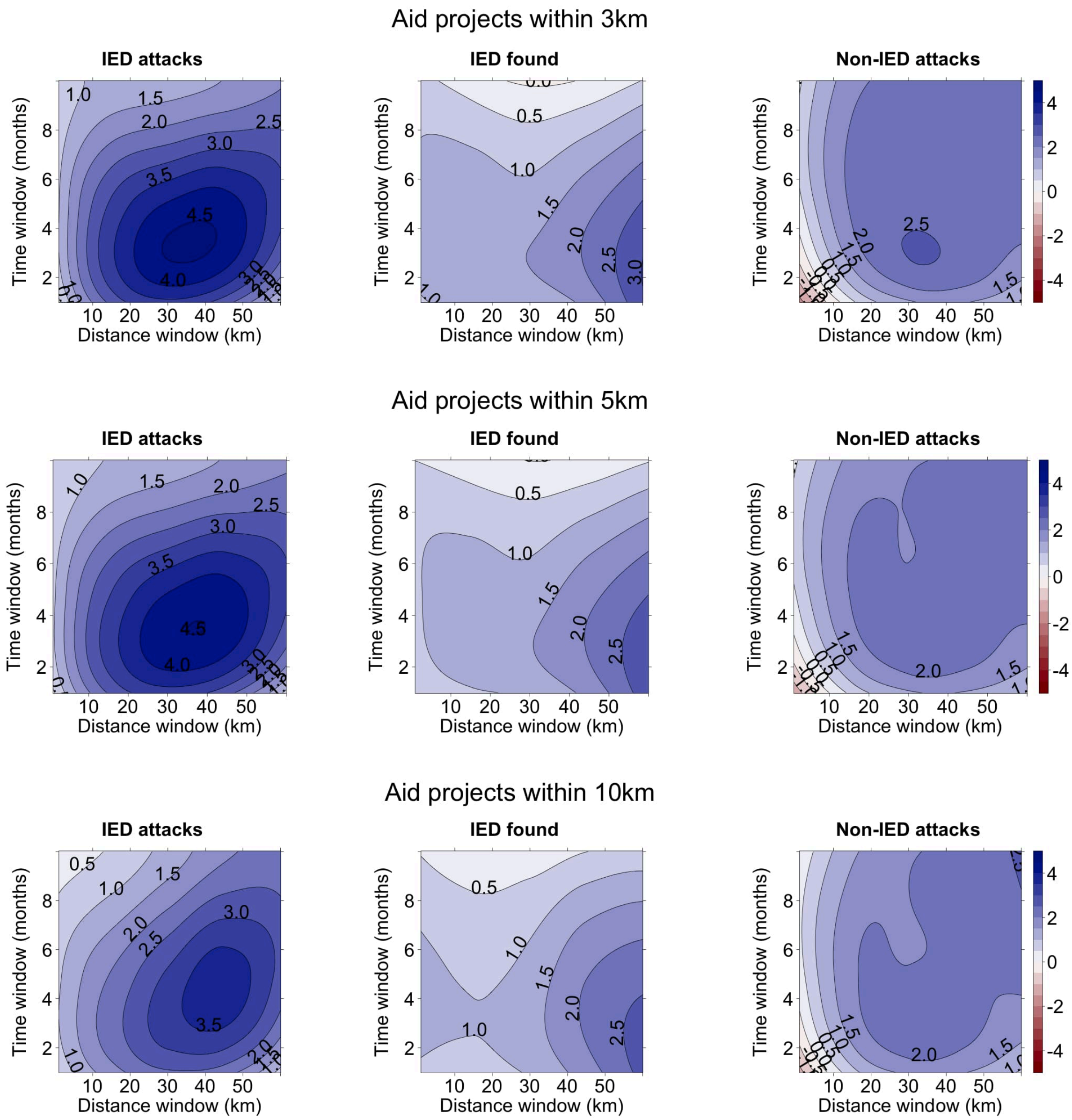

Figure 16. t values of the estimated coefficient of relative ISAF support, corresponding to its marginal effect on the number of future insurgent attacks, while adjusting for the prior level of insurgent violence, the number of bases, and the number of aid projects. The upper, middle, and lower panels use aid projects within 3km, 5km, and 10km radii of each sampled village, respectively, showing robustness of the positive association between the number of future insurgent violence and relative ISAF support. All the models use bases within a 3km radius of each village. Robust standard errors are used.

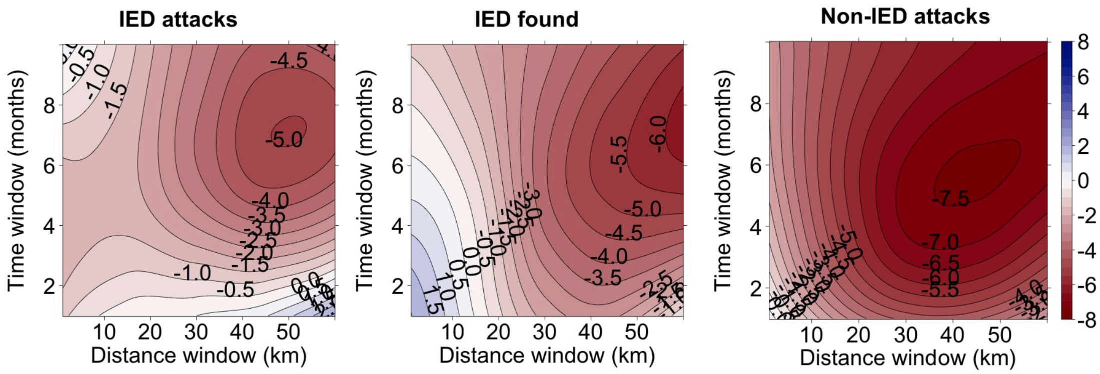

Figure 17. t values of the estimated coefficient for the level of *control by the ISAF* across a wide range of temporal and spatial window sizes. ISAF control is a four-point scale, district-level variable measuring the degree to which the ISAF controls the district. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks while adjusting for the relative level of ISAF support, the prior level of insurgent violence, the number of bases, and the number of aid projects. A linear regression model is used. The results show that the number of attacks by the Taliban is negatively associated with the level of control by the ISAF, suggesting that the Taliban is less likely to attack villages controlled by the ISAF.

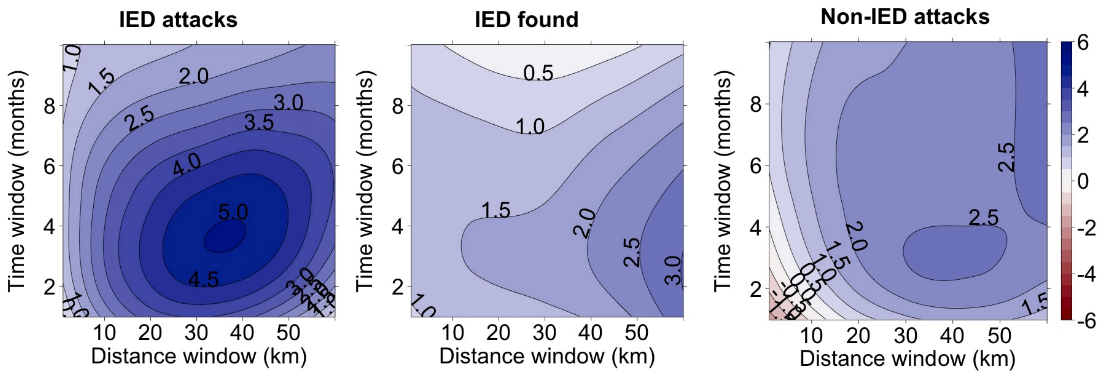

Figure 18. t values of the estimated coefficient for the *relative level of ISAF support* across a wide range of temporal and spatial window sizes. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks while adjusting for the *level of control by the ISAF*, the prior level of insurgent violence, the number of bases, and the number of aid projects. A linear regression model is used. The results show that relative ISAF support is still positively associated with the number of insurgent attacks even when the effect of ISAF control is taken into account.

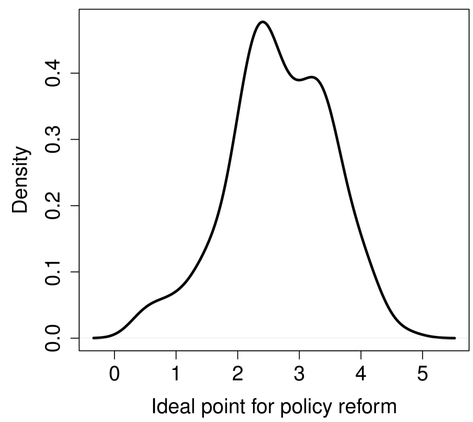

Figure 19. Distribution of the ideal points for policy reform obtained from the IRT endorsement experiment model. These estimated ideal points represent the degree to which the village is in favor of policy reform in general. Four policy reforms are considered: prison reform, direct election, independent election commission, and anti-corruption reform.

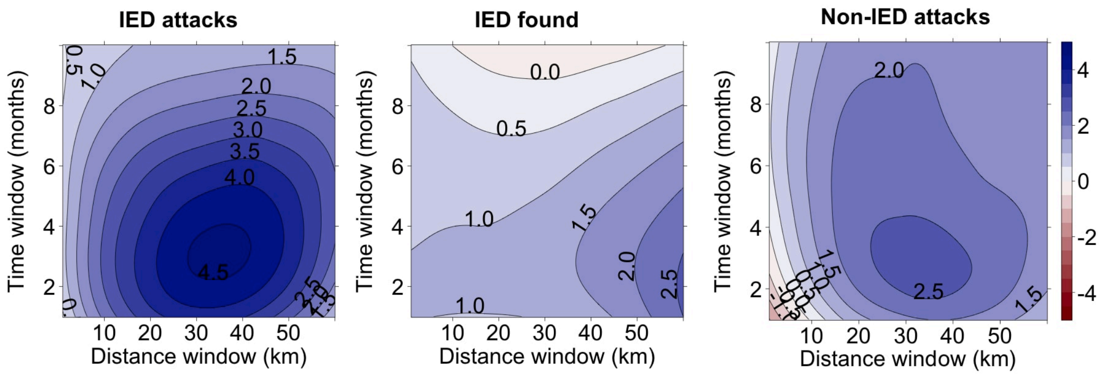

Figure 20. t values of the estimated coefficient for the *relative level of ISAF support* across a wide range of temporal and spatial window sizes. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks while adjusting for the *ideal points for policy reform*, the prior level of insurgent violence, the number of bases, and the number of aid projects. A linear regression model is used. The results show that relative ISAF support is still positively associated with the number of insurgent attacks even when the effect of ideal points for policy reform is taken into account.

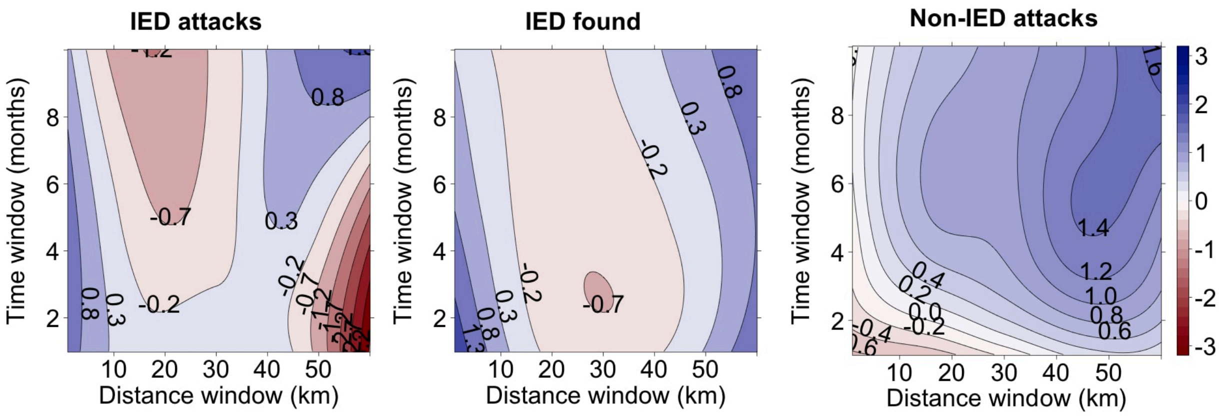

Figure 21. t values of the estimated coefficient for the relative level of ISAF support across a wide range of temporal and spatial window sizes. The estimated coefficient corresponds to its marginal effect on the number of future insurgent attacks while adjusting for the prior level of insurgent violence, the number of bases, the number of aid projects, and *random effects for districts*. A linear regression model is used. 204 surveyed villages were randomly drawn from 21 districts, which were also randomly drawn from 5 provinces. The results show that relative ISAF support is positively, but not strongly, associated with the number of Non-IED attacks when spatial correlations among villages within each district are taken into account.

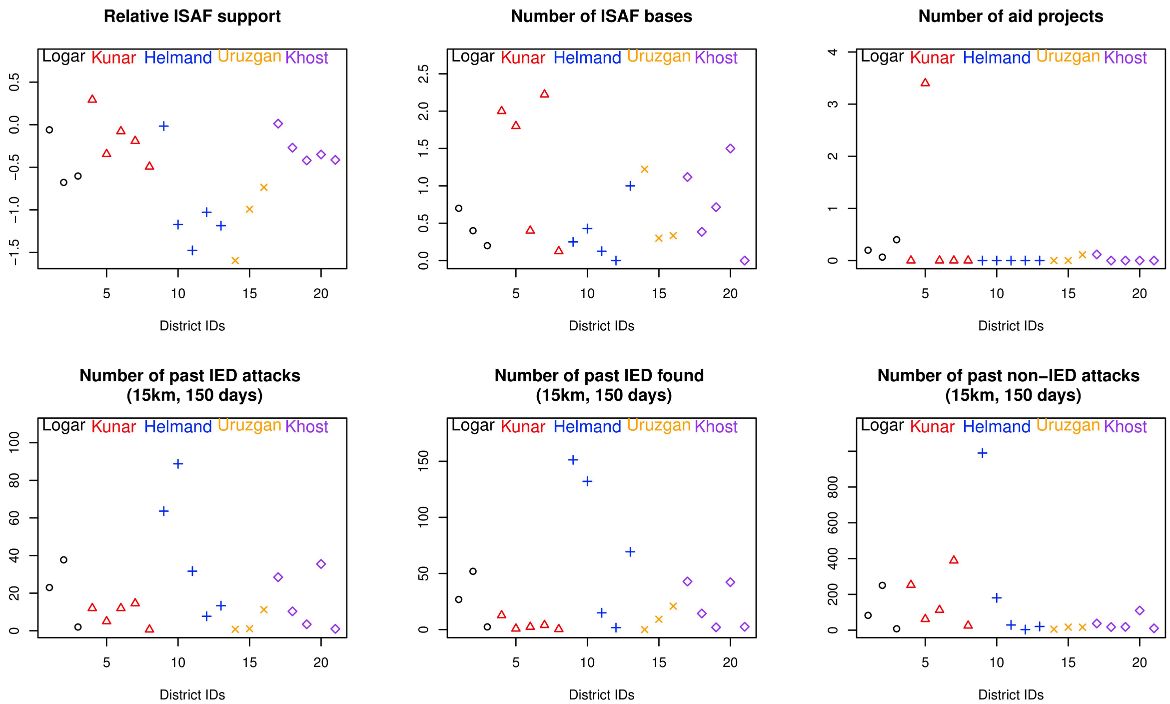

Figure 22. Correlations between districts and covariates. Each mark indicates a district-level average value of a covariate. 204 surveyed villages were randomly drawn from 21 districts, which were also randomly drawn from 5 provinces (Logar, Kunar, Helmand, Uruzgan, and Khost).

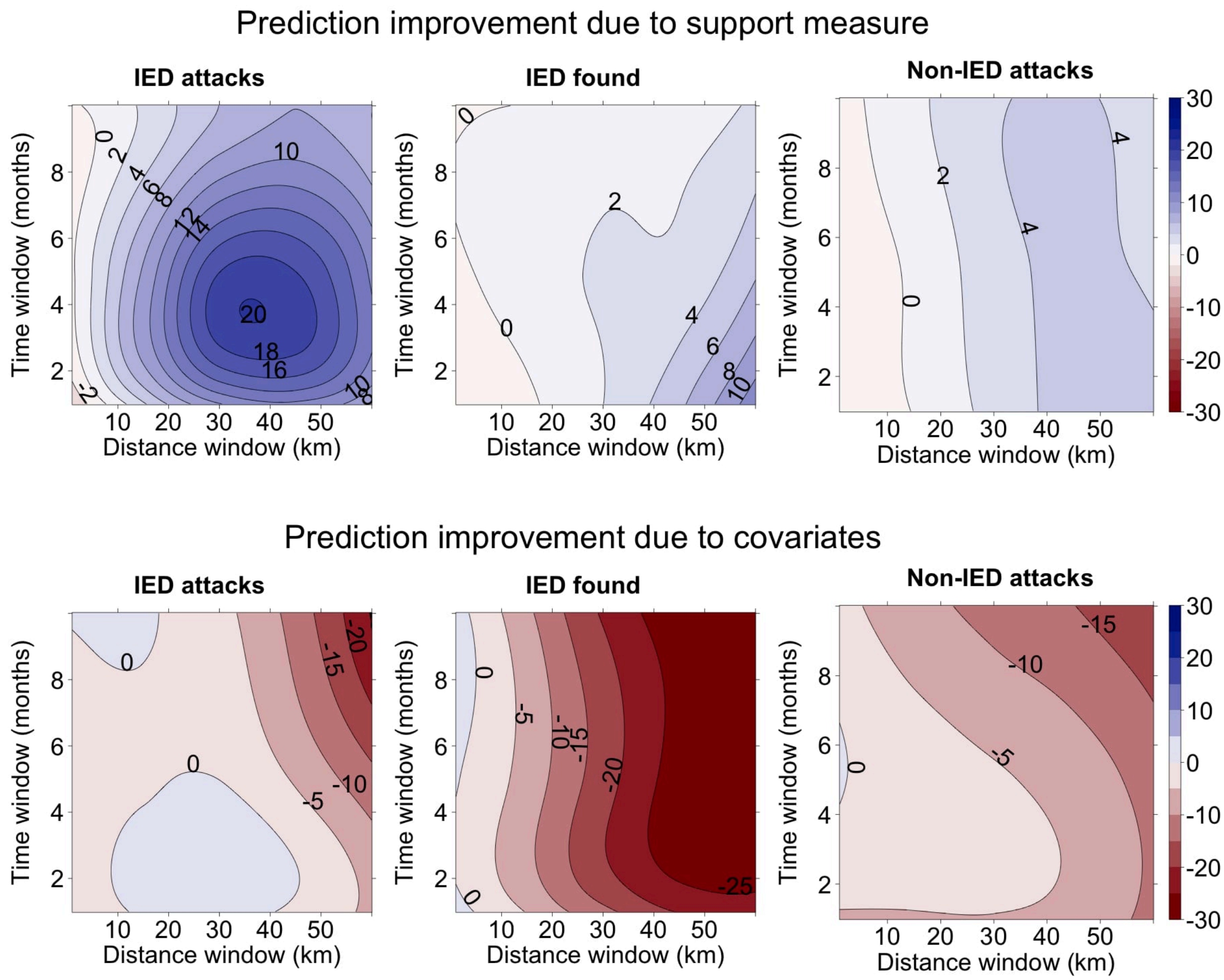

Figure 23. Forecasting improvement rates based on the root mean squared forecasting error (RMSFE). The contour plots show $(\text{RMSFE}_2 - \text{RMSFE}_1)/\text{RMSFE}_1 \times 100\%$, where RMSFE_1 and RMSFE_2 are the root mean squared forecasting errors from the models with and without the predicted ISAF support level. Upper panels represent the forecasting improvement rates of adding the estimated ISAF support level to the baseline model with prior violence, bases, and aid programs. Lower panels show the forecasting improvement rates of adding to the baseline model the village- and district-level covariates, such as log population, log elevation, ISAF control, Pakistan border, and Taliban Sharia, used to estimate the ISAF support level of non-surveyed villages.

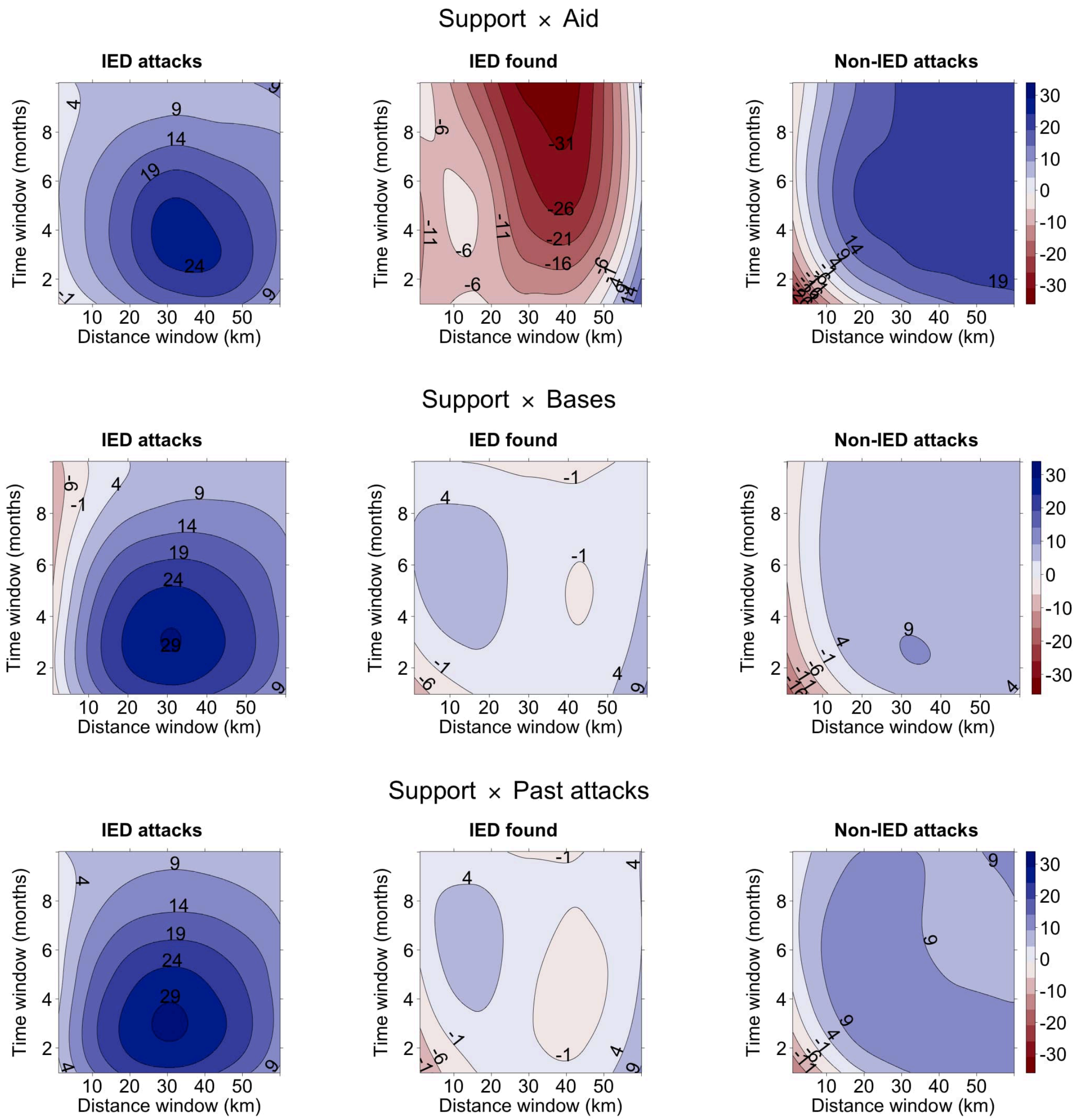

Figure 24. Forecasting improvement rates of adding interaction terms between estimated ISAF support level and prior violence, bases, and aid projects. The baseline model only includes ISAF support, past violence, bases, and aid projects. The contour plots show $(\text{MAFE}_2 - \text{MAFE}_1)/\text{MAFE}_1 \times 100\%$, where MAFE_1 and MAFE_2 are the mean absolute forecasting errors from the models with and without the predicted ISAF support level.

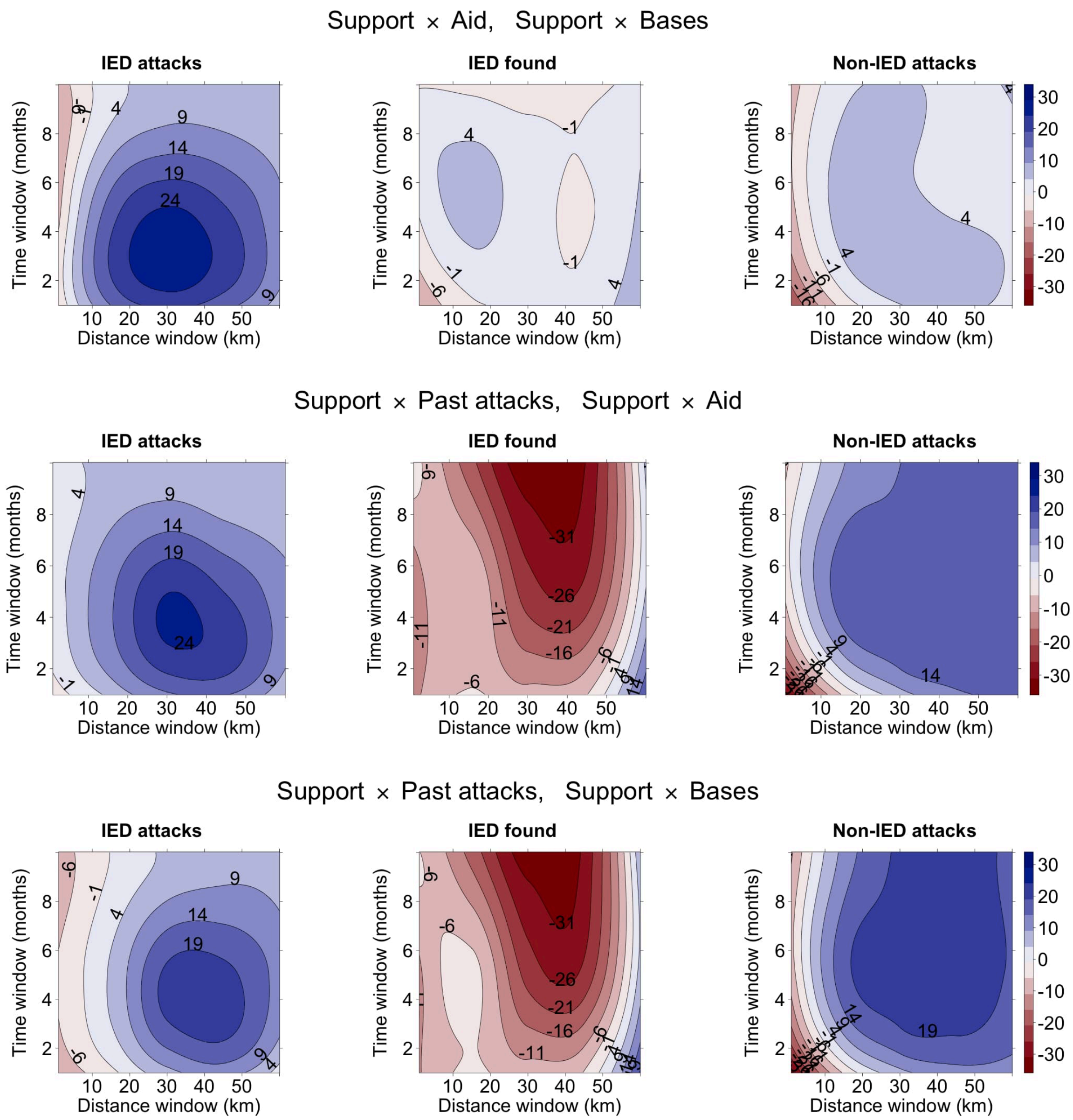

Figure 25. Forecasting improvement rates of adding interaction terms between estimated ISAF support level and prior violence, bases, and aid projects. The baseline model only includes ISAF support, past violence, bases, and aid projects. The contour plots show $(\mathrm{MAFE}_2 - \mathrm{MAFE}_1)/\mathrm{MAFE}_1 \times 100\%$, where MAFE_1 and MAFE_2 are the mean absolute forecasting errors from the models with and without the predicted ISAF support level.

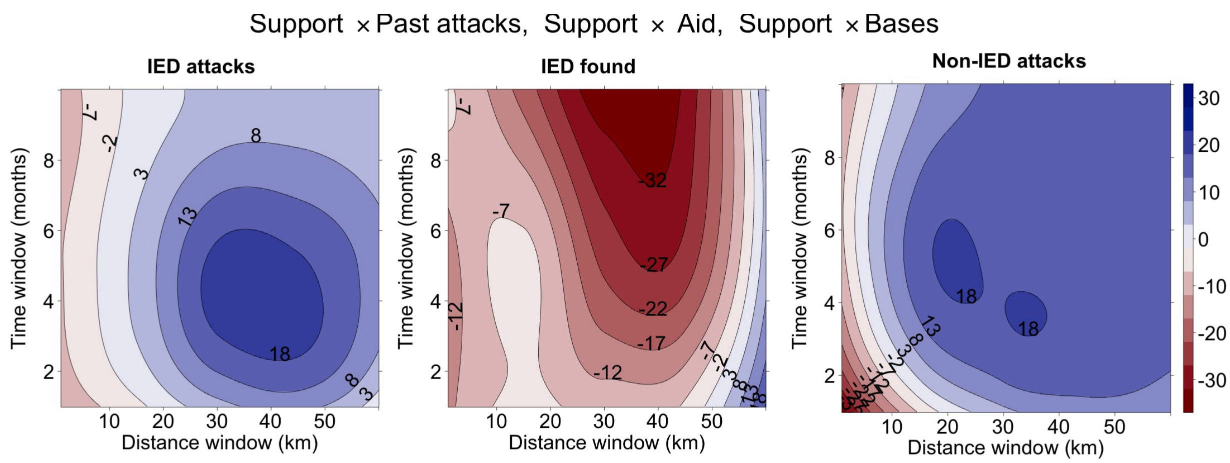

Figure 26. Forecasting improvement rates of adding interaction terms between estimated ISAF support level and prior violence, bases, and aid projects. The baseline model only includes ISAF support, past violence, bases, and aid projects. The contour plots show $(\text{MAFE}_2 - \text{MAFE}_1)/\text{MAFE}_1 \times 100\%$, where MAFE_1 and MAFE_2 are the mean absolute forecasting errors from the models with and without the predicted ISAF support level.