Mobile Phone Access and Insurgent Violence: Evidence from a Radio Wave Propagation Model in Afghanistan

Ryan Ellis*

Robert Gonzalez†

Lin Yang‡

September 13, 2024

Abstract

We provide evidence that mobile phone coverage lowers insurgent violence using two separate empirical designs in Afghanistan. First, we propose a novel regression discontinuity (RD) design for multiple, unknown cutoffs that uses signal strength as the forcing variable. We estimate signal strength for each cell of a $1X1$ kilometer grid using a high-resolution radio wave propagation model. Second, we employ a differences-in-differences design leveraging the staggered rollout of cell towers for the largest operator in Afghanistan. We explore insurgent coordination and information sharing as possible mechanisms. We find weak evidence of spatial and temporal changes in the clustering of attacks following the installation of a tower in a district. However, we find evidence of coverage having larger effects near populated areas, primary roads, during morning hours, and a sharp increase in call traffic preceding certain attacks, all suggesting that information sharing is a key mechanism.

Keywords: cell phone access, spatial regression discontinuity, insurgent violence

JEL Classification: P16, P35, P37, D72, D73

*School of Economics, Georgia Institute of Technology, Atlanta GA (e-mail: ryan.ellis@gatech.edu)

†School of Economics, Georgia Institute of Technology, Atlanta GA (e-mail: robert.gonzalez@gatech.edu)

‡School of Economics, Georgia Institute of Technology, Atlanta GA (e-mail: lyang371@gatech.edu)

1 Introduction

This paper examines the impact of mobile phone coverage on insurgent violence. In theory, access to coverage can lower violence by increasing the flow of information from civilians to the government engaging in counterinsurgency and by better shielding informants from retaliation. This is particularly the case if norms favoring insurgents are not too high among the general population. On the other hand, cell phone access can increase violence by reducing the cost of producing violence. For example, facilitating coordination among insurgents, and the remote detonation of cell phone-triggered improvised explosive devices (IEDs).

Researchers studying the impact of mobile phone coverage on any outcomes of interest face two main hurdles: how to accurately measure coverage and how to address the endogenous selection of locations and people into coverage. This paper proposes a novel method that addresses these issues. Specifically, we estimate a high spatial resolution radio-wave propagation model that uses variations in terrain topography and the spatial distribution of mobile phone towers to accurately predict signal strength on the ground for each cell of a $1X1$ kilometer grid of Afghanistan. The predicted signal strength is then used in a regression discontinuity design that compares grid cells within a small bandwidth around the signal strength threshold required for coverage. At this margin, access to coverage is mostly determined by minor exogenous changes in terrain features that lead to arbitrary diffraction and signal blocking.

We find considerable evidence that the net effect of access to mobile phone technology is to *lower* insurgent violence. Specifically, grid cells with just enough coverage experience a 2 percentage point drop in the likelihood of any attack and a 0.8 percentage point drop in the likelihood of an IED.

Results from a novel interval regression discontinuity design support this. We propose this design for cases where there are multiple unknown cutoffs. In this project, we do not observe the true received signal strength of each location. A village or grid cell has many potential signal cutoff values due to differences in network technology, call traffic, or even weather. However, empirical findings from the signals propagation literature provide us with a distribution of known signal reception, from which we construct a probability-weighted interval of signal cutoffs for Afghanistan during our sample period. We present identifying assumptions for the interval regression discontinuity (IRD) design, and describe the estimation of an interval-wide treatment effect that is generalizable to other research settings with multiple unknown cutoffs. Estimates from this specification are consistent with our traditional "sharp" RD design, though moderately larger in magnitude, indicating a 2.8-3.4

percentage point drop in the likelihood of any attack for grid cells with just enough coverage.

We estimate the dynamic effects of mobile coverage on violence leveraging the staggered rollout of cell tower installations from Afghanistan's largest mobile network operator. Arranged as a district-by-year panel, these data allow us to employ a staggered difference-in-differences strategy to identify treatment effects across time for groups of districts as they receive coverage. Our results are consistent with those we obtain from the regression discontinuity designs. Conditional on a set of fixed effects, violent events trend downward following the introduction of coverage, and decrease further as additional towers are installed within a district, on average.

We explore two key mechanisms by which mobile phone coverage can affect insurgent violence: rebel coordination and information sharing by civilians. To investigate the mechanism of coordination, we conducted both spatial and temporal analyses of violent incidents after cell towers installation. Specifically, we explore how the spatiotemporal clustering of violence changed. Our findings show that the presence of cell towers did not significantly alter the spatial clustering of violent incidents. Temporally, we measured changes in various moments of the temporal distribution of violence (i.e., standard deviation, skewness, and kurtosis). The results indicate that while additional towers reduce the volatility (lower standard deviation), asymmetry (lower skewness), and the concentration of extreme violence (lower kurtosis), the installation of the first tower results in an increase in kurtosis, suggesting short-term spikes in violence. In all, we find that coverage may lead to changes in the spatiotemporal coordination of attacks but we cannot directly say whether this can be interpreted as an improvement in coordination of insurgents. We note that, even if insurgent coordination improves, in equilibrium we find an overall drop in violence in coverage areas. Therefore, any gains in coordination are clearly offset by larger drops in violence due to the information sharing channel.

Further analysis suggests that mobile phone access facilitating information gathering from civilians is a key mechanism: The deterring effect of coverage is significantly larger in cells where detection of insurgent activities by civilians is more likely: near populated areas, near primary roads, and during morning hours. Similarly, the effect of coverage on the failure rate of attacks–measured as the share of unsuccessful IEDs–significantly increases in these cells. Notably, phone usage spikes around days where an IED is found by counterinsurgency forces.

The paper is structured as follows: Section 2 provides a conceptual framework for the effect of cell phone coverage on insurgent violence. Section 3 describes and project's primary datasets and presents the radio propagation model. Section 4 contains our primary and supporting empirical

strategies, and introduces the Interval Regression Discontinuity (IRD), a novel contribution of this paper. Sections 5 and 6 discuss results and mechanisms, respectively. Section 7 concludes.

2 Conceptual Framework

This section presents a theoretical model that illustrates the relationship between cell phone coverage and insurgency. It builds on the "Hearts and Minds" model presented in Berman et al. (2011) and extended in Shapiro and Siegel (2015). Coverage affects the equilibrium level of violence through three channels: (1) the cost that individuals pay to transmit information to the government, (2) by reducing the likelihood that individuals face retaliation from insurgents when sharing information, and (3) as a shifter in the cost of producing violence for insurgents. The model is a sequential game with the rebels R moving first by choosing violence level v, then the community C moves by choosing how much information i to share with the government.

2.1 Players and Payoffs

Community: C chooses how much information $i \in [0,1]$ to share with the government by maximizing their expected utility:

$$EU_C(i,v) = (c+g-n)p(i,\epsilon,\gamma) + (c-v)(1-p(i,\epsilon,\gamma)) - (1-\tau)i(R+T) \tag{1}$$

where $p(i,\epsilon,\gamma) = \gamma(i+\epsilon)$ gives the probability that the government succeeds in controlling the territory. $\epsilon \in [0, 1-i]$ is the information obtained by the government through other means. This could be, for instance, intelligence obtained through the interception of communications between insurgents. $\gamma \in [0,1]$ is a parameter that captures the ability of the government to convert information into a non-trivial increase in the probability $p(.)$ of defeating the insurgents. Under government rule, the community receives utility $c+g-n$ where c denotes consumption, g denotes government public goods, and n denotes norms favoring rebel rule and hence disutility from government control. If rebels succeed, the community receives the same level of consumption c net of disutility v resulting from insurgent violence. Sharing information with the government is costly. The community faces communication costs T (e.g., transportation costs, equipment costs, etc.), and retaliation costs R (e.g., insurgents targeting cooperating communities). The community faces these costs only if they share information (i.e., $i > 0$). Furthermore, the degree of cell phone coverage availability $\tau \in [0,1]$ reduces the costs faced by the community. For instance, coverage significantly reduces the costs of

communication T and shields the community from retaliation costs.

Rebels: The objective of rebels R is to maximize the violence-related costs incurred by the government subject to the cost of producing such violence. Specifically, rebels maximize:

$$EU_R(i, v) = A(v)(1 - p(i, \epsilon, \gamma)) - B(v; \tau) \tag{2}$$

where $A(v)$ gives the costs imposed on government, with $A^{'}(v) > 0; A^{''}(v) < 0$. $B(v; \tau)$ is the cost of producing violence that the rebels face ($B^{'}(v) > 0; B^{''}(v) > 0$). We assume that $B(v; \tau)$ decreases with coverage τ. This captures the fact that coverage can improve coordination among rebels or reduce the cost of violence-producing technology such as remotely-detonated explosive devices.

2.2 Best Responses and Equilibrium

The game is solved by backwards induction starting with the community. Given that i enters linearly in Equation (1), the community's best response is simply given by:

$$i^* = \begin{cases} 0 & \text{if } n \geq g + v - \dfrac{1-\tau}{\gamma}(R+T) \\ 1 & \text{if } n < g + v - \dfrac{1-\tau}{\gamma}(R+T) \end{cases} \tag{3}$$

where C alternates between zero information sharing or full cooperation depending on whether rebel affinity n is sufficiently large to offset public goods provision, disutility from violence, and the effective costs of communication and retaliation.

Rebels choose the level of violence v by maximizing Equation (2) expecting optimal behavior by C. The equilibrium level of violence solves the following first order condition:

$$A'(v^*)(1 - \gamma(E(i^*) + \epsilon)) - A(v^*)\frac{\gamma}{n_U - n_L} - B'(v^*; \tau) = 0 \tag{4}$$

where $E(i^*) = P(i^* = 1) = (n_U - n_L)^{-1}\left[g + v - (1-\tau)(R+T)\gamma^{-1} - n_L\right]$ if we assume that community norms are distributed uniformly in $[n_L, n_U]$.

The main takeaway from the model is that the relationship between coverage and insurgent violence is theoretically ambiguous: (i) violence can increase with coverage as coverage can lower the costs of producing violence (e.g., reducing the costs of coordination among rebels, or the detonating of IEDs), (ii) coverage can lower violence if it also reduces the cost of information sharing between

civilians and the government engaging in counterinsurgency. This cost reduction can either be the result of lower communication costs or by more effectively shielding civilian informants from retaliation from insurgents. However, this mechanism is more nuanced as it is only effective if norms favoring rebels within the population are not too high. (iii) Coverage can reduce violence if the government engaging in counterinsurgency can capture information being transmitted among rebels using phones.

3 Data

This section describes the primary datasets used in the analysis: insurgent violence data and the measure of cell phone coverage.

3.1 Insurgent Violence Data

Insurgent violence data are obtained from time-stamped and georeferenced records collected by International Security Assistance Forces (ISAF) and Afghan forces in Afghanistan. The data include, among other things, time (by the hour), geolocation (within meters), and the type of incident (direct fire, IED, etc.). The analysis in this paper uses four categories of violence: improvised explosive device (IED) explosions, direct fire attacks, and indirect fire attacks. Direct fire refers to an attack on a target that is visible to the attacker. Examples include small arms fire, rocket propelled grenades, or a thrown hand grenade. Indirect fire refers to an attack characterized by a relatively high trajectory and where the attacker fired from a distance beyond line-of-sight. Examples include artillery, mortar and rocket (NATO, 2016).[1] In order to define a spatial unit of aggregation, we create a 1-kilometer by 1-kilometer grid covering Afghanistan. The main measure of violence at the cell level is given by an indicator for whether there has been at least one incident within that cell of the grid. The primary motivation for creating the grid is that a significant number of violent incidents take place away from populated centers, however, additional results using whether an incident occurred within a 2-kilometer radius around villages are also presented.[2]

Figure 1 presents a time-series of IED explosions, direct fire, indirect fire, and all attacks combined in 2012, the year of analysis used in this paper. Note that direct fire incidents is the largest category of violence while all categories exhibit a similar within-year pattern: attacks start

[1] Refer to Condra et al. (2018) for a more detailed description of this dataset.

[2] Village location data are obtained from the Measuring Impacts of Stabilization Initiatives (MISTI) project sponsored by US Agency for International Development (USAID) (MISTI, 2013). Refer to Gonzalez (2019) for more information on this dataset.

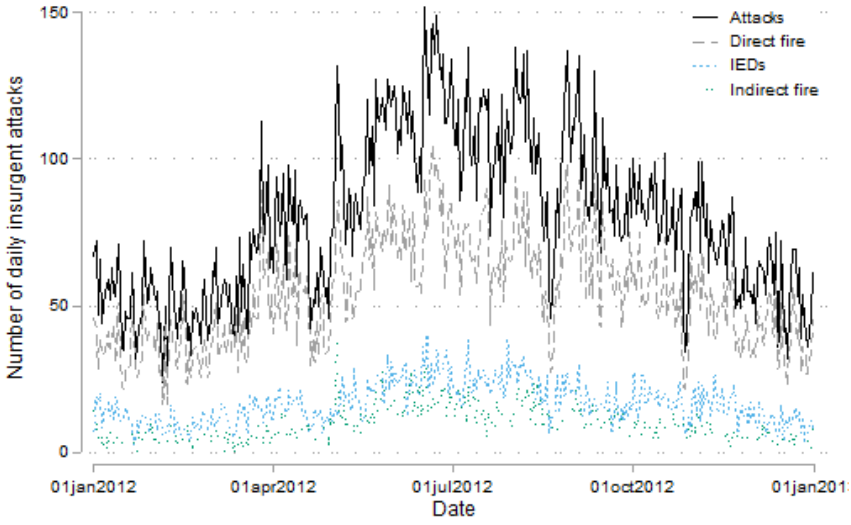


Figure 1: Insurgent Violence, 2012

Notes: Daily violence.

gradually increasing during the spring and gradually decrease during winter months. Table 1 presents summary statistics for the four categories. On average, about 24%, 17%, and 9% of villages experienced a direct fire, IED, or indirect fire attack within 2 kilometers, respectively.

3.2 Measure of Coverage

Cell phone technology uses high frequency radio waves transmitted by cell phone tower antennas (transmitters) to enable communication between mobile, hand-held devices (receivers). Given the location of an antenna, one can therefore assess the strength of coverage at a given point on the

Table 1: Insurgent Violence Summary Statistics, 2012

	Villages				Grid [1km×1km]			
	Whole sample		Within 10 dBm		Whole sample		Within 10 dBm	
	Mean (1)	S.D. (2)	Mean (3)	S.D. (4)	Mean (5)	S.D. (6)	Mean (7)	S.D. (8)
Any attack	0.289	0.45	0.384	0.49	0.016	0.13	0.036	0.19
Direct fire	0.236	0.42	0.328	0.47	0.012	0.11	0.027	0.16
IED	0.172	0.38	0.224	0.42	0.006	0.08	0.014	0.12
Indirect fire	0.094	0.29	0.131	0.34	0.003	0.05	0.005	0.07
Observations	45017		7853		647463		92584	

Notes:

ground using a radio wave or signal propagation model.

This paper uses the Irregular Terrain Model (ITM) to determine coverage strength across Afghanistan. The ITM is the workhorse model used by the United States government, the Federal Communications Commission, and businesses to model coverage and signal propagation. This is primarily due to its high accuracy and ability to capture terrain topography.[3]

The model provides a measure of received power or signal strength at a given point in space taking as inputs three primary sets of information: (1) the characteristics of the transmitter or cell tower antenna (e.g., latitude and longitude, antenna height, frequency of radio wave), (2) the characteristics of the receiver or mobile device (e.g., antenna height and gain, receiver sensitivity), and (3) geographic characteristics of the terrain (e.g., topography, climate, ground conductivity). Figure 2a maps the location of cell towers for all mobile network providers in Afghanistan in the year 2012. Location data and other characteristics of cell towers are obtained from the Afghan Telecommunication Regulatory Authority (ATRA).[4] We use the 30-meter resolution ALOS Global Digital Surface Model to accurately capture the effect of topography on signal propagation (JAXA, 2016). This is currently, the most precise global-scale elevation data model available (Open Topography, 2017). Appendix **??** provides a detailed discussion of the variables and parameters used in the estimation of the ITM model in this paper.

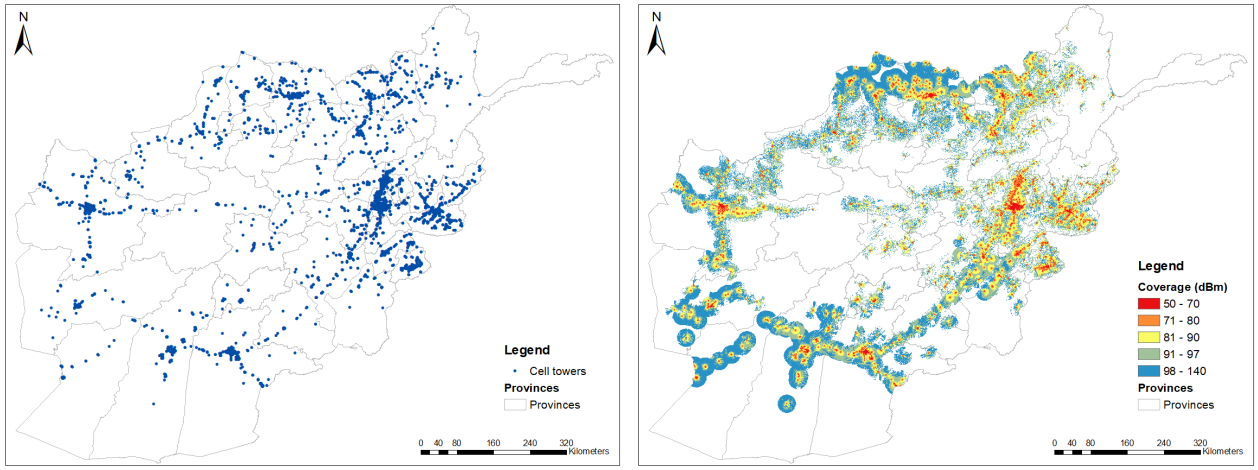
The model output is presented in map form in Figure 2b. Coverage strength–measured as received power on the ground–is measured in decibel-milliwatts (dBm). Received power typically ranges between 0 and -140dBm. However, for ease of interpretation, our measure of coverage uses the absolute value of received power. Therefore, lower dBm levels should be interpreted as stronger coverage. Note how received power is significantly affected by topography: signal propagation is much more scattered in the northern and central mountainous parts of the country.

4 Empirical Strategy

Our empirical strategy consists of two sets of analysis. First, we implement a set of regression discontinuity designs: a sharp RD design, a donut RD design, and an interval RD design–this is a

[3]Predictions from the ITM have been repeatedly validated via on-the-ground measurements (Longley and Rice, 1968; Eppink and Kuebler, 1994; Seybold, 2005; Lazaridis et al., 2013). Refer to Crabtree and Kern (2018) for a detailed discussion of the ITM. Other propagation models specifically designed to model cell phone coverage exist. However, these were mainly designed for urban and suburban environments where obstacles to propagation come from building footprints rather than topography.

[4]Tower location data obtained from ATRA are for the year 2016 while insurgent violence data end in 2014. To obtain an overlapping sample of tower and violence data, I use a 2012 tower location map from ATRA to extract the towers from the original 2016 data.



(a) Cell Towers, Afghanistan (2012)

(b) ITM Model Output

Figure 2: Cell Phone Towers and ITM Output

Notes: Cell tower location obtained from the Afghan Telecommunications Regulatory Authority (ATRA). Estimates of the ITM model described in Section 3.2 and in Appendix **??**. Lower dBm values mean higher signal strength (i.e., more coverage).

novel design that accommodates the case of multiple unknown cutoffs. Second, we use a difference-in-differences design exploiting the staggered rollout of cell towers for the largest mobile network operator in Afghanistan.

4.1 Sharp Regression Discontinuity Design (SRD)

For the sharp RD design, we estimate the effect of cell phone coverage on violence following the regression discontinuity (RD) design described in Gonzalez and Maffioli (2024). The design uses signal strength as the forcing variable and the receiver sensitivity threshold as the treatment cutoff. This is a network-specific threshold that determines whether the mobile device is receiving enough signal strength to be able to make voice calls, send SMS, etc. In the case of GSM networks such as that of Afghanistan, receiver sensitivity typically ranges between 90 to 105 dBm, in absolute value. Specifically, we estimate the following baseline specification:

$$v_i = \gamma + \beta D_i + f(\tilde{R}_i) + h(\mathbf{G}_i) + \epsilon_i \tag{5}$$

where v_i is a measure of insurgent violence in cell i of the grid. $\tilde{R}_i = R_i - c$ is received power (measured in dBm) at cell i net of the receiver sensitivity cutoff c. $D_i = \mathbb{1}\{R_i \geq c\} = \mathbb{1}\{\tilde{R}_i \geq 0\}$ is an indicator for whether cell i has coverage (i.e., received power is higher than cutoff c). $f(\tilde{R}_i)$ is the

RD polynomial. The main results use a local linear specification $f(\tilde{R}_i) = \alpha_1 \cdot \tilde{R}_i + \alpha_2 \cdot D_i \times \tilde{R}_i$ with a bandwidth h around the cutoff c optimally determined (Calonico et al., 2014).[5] All specifications include polynomial in elevation and terrain slope $h(\mathbf{G}_i)$ to capture any differences in topography near the cutoff that may be correlated with violence and coverage.[6]

We choose a cutoff c of 97 for two reasons. First, mobile network operators in Afghanistan used a combination of GSM900 and GSM1800 networks during our sample period. Repeated testings from the engineering literature and the GSM Association (GSMA), as well as industry standards proposed by the GSMA and the Cellular Telephone Industries Association (CTIA) suggest an average signal cutoff of 95dBm for GSM900 networks and 99dBm for GSM1800 networks (Razally, 2015; GSMA, 2019). We pick an average of the two (i.e., 97 dBm) since we do not observe the specific network used by each tower. Second, employing the Difference-in-Kernels estimator described in Qiu (2011) and Porter and Yu (2015) to empirically detect unknown discontinuities points yields a clear cutoff at 97dBm. This exercise is presented in Figure 3.[7]

Although there is clear evidence of a response at 97dBm. The value of the cutoff depends on the network used by the antenna (GSM900 or GSM1800) and even then there could be variation in the network-specific cutoff due to factors like daily variation in weather and call traffic. We do not observe these cutoffs. Therefore, one should think of our chosen cutoff as the expected cutoff and our results from the sharp RD design as capturing an intent-to-treat similar to the intent-to-treat captured by a sharp RD in a fuzzy RD setting. β estimates the causal effect of coverage access on insurgent violence under certain assumptions. These assumptions are discussed next.

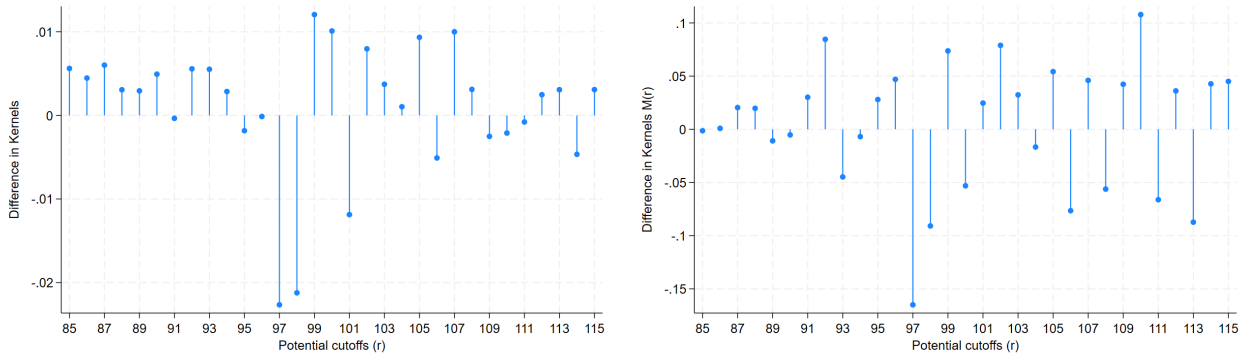
4.2 Validity of the RD Identifying Assumptions

Intuitively, the RD design presented above compares violence levels between cells of the grid that receive just enough power with cells where received power is just below the sensitivity threshold. Using Appendix Figure A2 as a reference, this is tantamount to comparing violence levels between green-shaded and blue-shaded areas. This exercise identifies the causal effect of coverage under one key assumption: potential outcome functions $E\left[v(1)|\tilde{R}\right]$ and $E\left[v\left(0\right)|\tilde{R}\right]$ are continuous at the

[5] This is essentially a local linear regression design that uses a rectangular kernel with bandwidth h. We choose a rectangular kernel since kernel choice has little impact in practice (Lee and Lemieux, 2010).

[6] In the discussion of the identifying assumptions below, we note the importance of topography as a determinant of coverage near the cutoff.

[7] Refer to Gonzalez and Maffioli (2024) for more details on the implementation of this method in a similar RD design. This method compares the kernel-weighted average of the outcome on the right and left sides of potential cutoff r. The difference between the weighted averages is large if r is a cutoff point and small if r is not a cutoff or if there is no treatment effect.



Potential Cutoffs (grid) Potential Cutoffs (Village)

Figure 3: Detecting Treatment Cutoff

Notes: Each dot represents the estimate of $M(r) = \frac{1}{nh}\sum_{i=1}^{n} Y_i K_1\left(\frac{R_i - r}{h}\right) - \frac{1}{nh}\sum_{i=1}^{n} Y_i K_0\left(\frac{R_i - r}{h}\right)$ for each potential signal strength cutoff r. R_i denotes the signal strength in each location i. Y_i is our outcome of interest. $K_1(.)$ and $K_0(.)$ are one-sided kernel estimators on the right and left side of potential cutoff r. We use a triangular kernel and a bandwidth h of 2 dBms. We restrict values of r to be within 85 to 115 dBm. There is a clear jump in the difference in the kernels estimator at $r = 97$. The smaller jump at 98 is due to the 2dbm bandwidth used in the estimation.

coverage cutoff c, where one and zero denote assignment and non-assignment into treatment, respectively. Simply put, observables and unobservables must transition smoothly across the coverage cutoff so that grid cells with received power just below the cutoff can serve as a valid counterfactual for treated cells. In the context of this study, this is a plausible assumption. We compare areas that are at the margin of minimal coverage (and hence away from the endogenously selected tower locations). As shown in Appendix Figure A2, we compare locations that are relatively close to each other, and whether some receive enough signal strength or not is purely determined by exogenous variations in topography, which we control for.

Table 2 provides summary statistics for several socioeconomic and geographic characteristics of the sample grid. Columns (1) and (2) report the mean of these variables for areas within 20 dBm on each side of the sensitivity cutoff. Columns (4) and (5) repeat the exercise using a 10 dBm bandwidth. Columns (3) and (6) report the clustered standard errors of the difference in means between covered and non-covered cells of the grid. Comparing columns (1) and (2) suggests that areas with coverage tend to be closer to health facilities, roads, rivers, and have higher populations. Most importantly, notice that as we narrow the bandwidth of analysis (columns (4) and (5)), many of the differences fade away. Of all socioeconomic characteristics tested, only two remain statistically significant as the bandwidth is narrowed down. This provides support for the continuity assumption discussed above.

Table 2: Summary Statistics

	Within 20 dBm			Within 10 dBm		
	Coverage (1)	No coverage (2)	S.E. (3)	Coverage (4)	No coverage (5)	S.E. (6)
Socioeconomic characteristics:						
Distance (km) to:						
Health facility	19.13	21.83	(0.64)***	20.49	21.38	(0.74)
Primary road	30.00	34.72	(3.27)	31.71	30.71	(3.32)
Secondary road	2.93	3.72	(0.18)***	3.35	3.48	(0.16)
River	5.94	8.99	(0.49)***	6.61	8.83	(0.60)***
District center	0.01	0.01	(0.00)	0.00	0.00	(0.00)
Population	545.60	449.15	(27.33)***	470.78	472.72	(35.38)
Language spoken:						
Dari	0.32	0.29	(0.03)	0.32	0.24	(0.03)**
Pashto	0.55	0.56	(0.03)	0.55	0.60	(0.04)
Other	0.13	0.16	(0.03)	0.13	0.16	(0.04)
Topographic characteristics:						
Elevation (m)	1,454.07	910.25	(46.36)***	1,487.03	813.51	(53.19)***
Slope (%)	3.57	1.14	(0.14)***	3.71	0.76	(0.17)***
Observations	104,470	35,142		74,469	18,115	

Notes: Columns (1), (2), (4), and (5) give the means of the corresponding variable. Columns (3) and (6) give the clustered standard errors for the difference in means in parenthesis. Clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance respectively. "District center", "Population", "Language spoken" uses the village sample instead of grid sample.

Topographic characteristics, on the other hand, differ significantly even as we restrict the bandwidth. This is somehow expected: a key determinant of signal propagation is terrain ruggedness.[8] For this reason, all specifications of Equation (5) include controls for elevation and slope. Moreover, we present results including all socioeconomic and topographic characteristics as covariates and show that the estimated impact of coverage is not sensitive to this inclusion.

4.3 Interval Regression Discontinuity Design (IRD)

Locations have different signal strength cutoffs due to the specific network of the antennas servicing them (GSM 900 versus GSM 1800) or due to daily fluctuations in weather and call traffic. We do not observe the specific cutoffs but we know, based on how the technology works that these cutoffs can only be between a given interval. This scenario can be framed as one with multiple, unknown treatment cutoffs.

[8]For a graphical depiction of the continuity of baseline covariates across the coverage cutoff, refer to Appendix Figure A6 which presents RD plots for all covariates in Table 2.

In this section, we propose a novel RD design to accommodate this scenario. This design builds on the donut RD design and the multiple cutoffs RD design (Barreca et al., 2011; Cattaneo et al., 2016). The key distinction is that, unlike typical donut RD and multiple cutoffs RD designs where the cutoff(s) are known, our framework allows for estimating treatment effects when there are multiple, unknown cutoffs. We refer to this as an Interval RD (IRD) design. We proceed by describing this framework.

Let $R \in [\underline{R}, \overline{R}]$ be the forcing variable (i.e., signal strength in our setting). Suppose c is one of infinitely many possible cutoffs in $C = [c_-, c_+]$, where c is unknown but c_- and c_+ are known from the engineering literature on signal propagation. In other words, the exact cutoffs are unknown, but their minimum and maximum possible values are known.

Assume that any location with signal strength $R = r$ such that $r < c_-$ is untreated with certainty, and any location with $r > c_+$ is treated with certainty. Let $E[v(1)|R]$ and $E[v(0)|R]$ be the potential outcome functions in the presence and absence of coverage, respectively.

To estimate the treatment effect at cutoff c, we proceed as follows: 1. Estimate the potential outcome function $E[v(1)|R]$ in the interval $(c_+, \overline{R}]$ and $E[v(0)|R]$ in the interval $[\underline{R}, c_-)$ by excluding all observations within C.[9] 2. Extrapolate $E[v(1)|R]$ and $E[v(0)|R]$ within the interval $C = [c_-, c_+]$. Let $m_1(c)$ and $m_0(c)$ be the predicted values of the potential outcome functions at the specified cutoff $c \in C$. 3. The estimated treatment effect at a given cutoff $c \in C$ is given by:

$$\tau(c) = m_1(c) - m_0(c)$$

4. To estimate an interval-wide treatment effect, select I random cutoffs within C and compute the following weighted average:

$$\tau = \frac{\sum_{i=1}^{I} \tau(c_i) f(c_i)}{\sum_{i=1}^{I} f(c_i)} \tag{6}$$

where $f(c_i)$ is the density of the cutoffs, i.e., the probability of having cutoff c_i among all possible cutoffs within C.

Figure 4 gives a graphical representation of the IRD design. Panel (a) shows interval C. Panel (b) depicts the estimation of the potential outcome functions outside of interval C. Panel (c) depicts the extrapolation exercise while panel (d) shows a sample treatment effect at a given point $c \in C$. Calculating the interval-wide treatment effect in step 4, entails estimating a number of treatment effects like the one depicted in panel (d) and then averaging them following equation (6). Note

[9] We can also estimate the potential outcome functions using smaller intervals, similar to local linear regression.

in panel (a) that there is a clear jump within C at a signal strength value of -97. This suggests that although there might be multiple cutoffs, given the 2G technology used in our setting, there is room to also employ a sharp RD design using this clear cutoff. We will take advantage of this in section 4.1.

In terms of the endpoints of C, we base our choice on empirical findings from the signals propagation literature (Razally, 2015; GSMA, 2019). Specifically, we pick $C = [-103, -91]$ because these where the minimum and maximum values for signal cutoffs found in repeated testings of phones using the same technology (GSM 2G) and frequency (900MHz and 1800MHz) used in Afghanistan during our sample period. We estimate the density of cutoffs $f(c_i)$ in Equation 6 using the observed cutoffs measured in several engineering experiments in Razally (2015); GSMA (2019) for the same network available in Afghanistan during our sample period.

IRD Design Assumptions Identification of treatment effects in the IRD design proposed in this paper requires two key assumptions. First, the potential outcome functions $E[v(1)|R]$ and $E[v(0)|R]$ have to be continuous in signal strength R at any cutoff point $c \in C$. This ensures that control locations near the cutoff can be a proper counterfactual for treated locations near this cutoff. This is, in essence, the continuity assumption of any typical RD design just extended to each possible cutoff. Second, the potential outcome functions $E[v(1)|R]$ and $E[v(0)|R]$ are independent of the cutoff $c \in C$. In other words, for any cutoff c, the potential outcome functions would be the same. This is the cutoff ignorability assumption in the multiple cutoffs RD literature (Cattaneo et al., 2016). This assumption is what allows extrapolating the expected outcomes for each cutoff from a single potential outcome function as depicted in panel (c) of Figure 4.

This assumption means that the violence production function does not change with different signal strength cutoffs. This is a plausible assumption in our setting. A rebel group will take into account the presence of coverage when choosing violence (e.g., can a cellphone-triggered IED be used in this area), however, it is unlikely that this decision will depend on the network-specific cutoff. Similar to the researcher, for the rebel group the cutoff is unknown, therefore, it is unlikely that it plays any part in the group's violence production function. What matters is whether there is coverage or not.

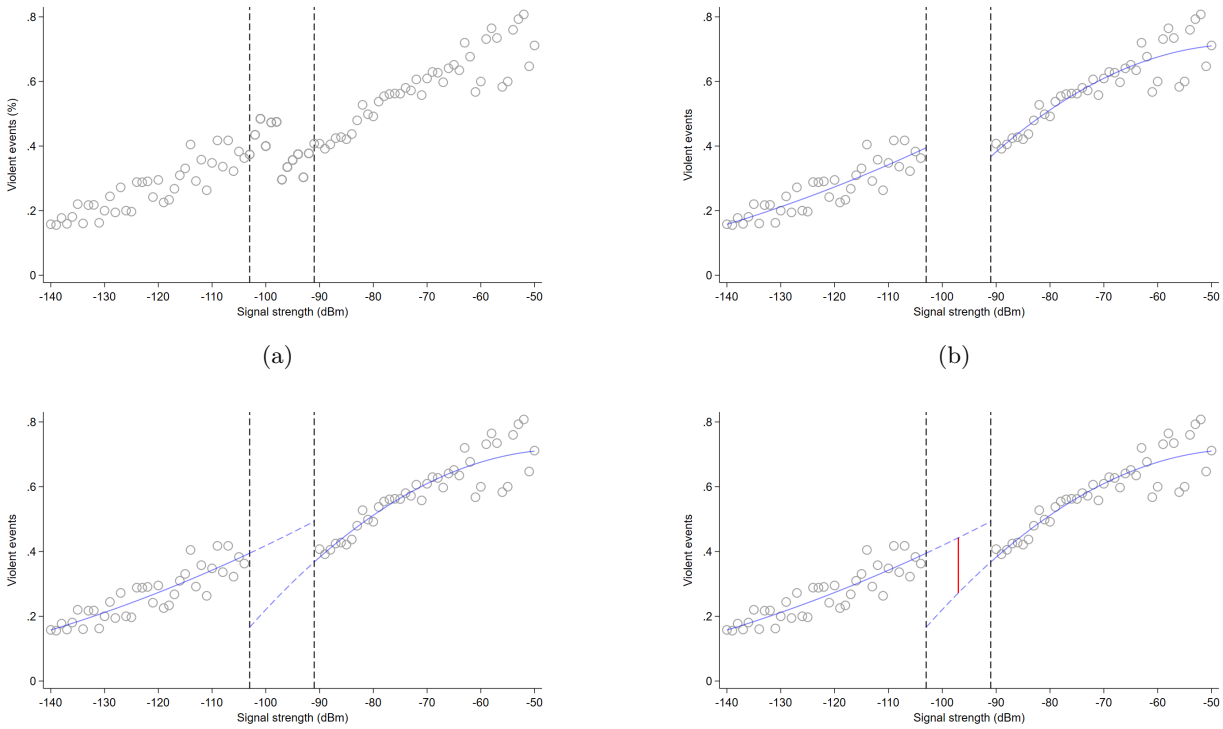


Figure 4: Interval Regression Discontinuity Design Illustration

Notes: Dots plot the average of the outcome variable (all incidents within 2km of a village) for a signal strength bins of 1 dBm.

4.4 Panel Regressions

In addition to RD results, we present further evidence that coverage may decrease violence using a panel of cell tower installations over time. These data come from Roshan, Afghanistan's largest mobile network operator. Roshan was an effective first-mover in many Afghan provinces, rapidly expanding the reach of cellular coverage in the country, beginning in 2003.[10] By 2009, Roshan had at least one installation in each of Afghanistan's 34 provinces. Our data contain the geographic placement of all Roshan towers brought into service between 2003 and 2011, as well as each tower's date of activation, allowing for panel estimation techniques that leverage differences in coverage across time and geography.

Our outcome of interest is the combined likelihood of all direct fire, indirect fire, and IED events from the SIGACTS dataset for a given district in a given year. Even aggregated in this way, a large portion of district-year observations contain zero violent events. Given the unsuitability of OLS for data with abundant zeros, we follow convention and estimate effects via Poisson Pseudo-Maximum Likelihood (Silva and Tenreyro, 2006). To identify the treatment effects of coverage in a staggered rollout, we employ the following Two-Way Mundlak Difference-in-Differences strategy (Wooldridge, 2021):

$$E(y_{it}|\mathbf{d}_i) = exp\left[\eta + \sum_{q=2003}^{T} \gamma_q d_{iq} + \sum_{s=2003}^{T} \theta_s fs_t + \sum_{r=2003}^{T} \sum_{\substack{s=2002 \\ s\neq(r-1)}}^{T} \tau_{rs}(d_{ir} \cdot fs_t)\right] \tag{7}$$

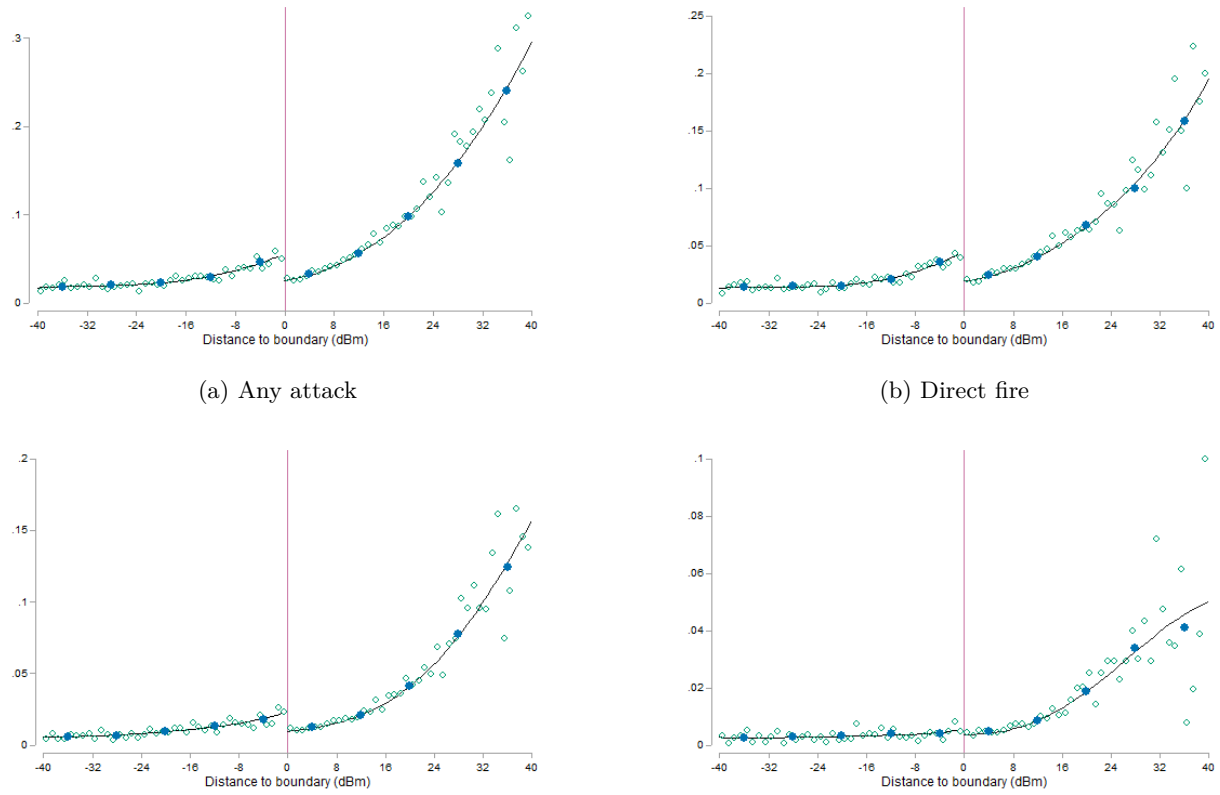
where binary outcome y_{it} is equal to 1 when any violent events occur in district i in year t. The indicators d_i equal 1 in the year a district first receives a cell tower, beginning with the first coverage-year cohort (represented by subscript g). fs_t is a year dummy equal to 1 when $s = t$. Effectively, this approach interacts every treatment cohort with every post-treatment time period. Treatment effects (τ_{rs}) estimated by pooled regressions are unbiased when at least one new unit enters into treatment in every period $(g, ..., T)$.[11]

5 Results

5.1 Graphical Analysis

Figure 5 plots the likelihood of an insurgent attack (Panel (a)), a direct fire attack (Panel (b)), an IED explosion (Panel (c)), and an indirect fire attack (Panel (d)) at a given cell of the grid

[10] The country's first operator, Afghan Wireless, was first to market in Kabul, in 2002.
[11] See Wooldridge for more.



(a) Any attack

(b) Direct fire

(c) IED

(d) Indirect fire

Figure 5: Regression Discontinuity Plots

Notes: "Negative" distances refer to no signal areas. Positive distances refer to areas with coverage. Cutoff at zero represents a signal strength threshold of 97 dBm. Solid dot give the average of the outcome variable (listed below each figure) for a signal bin of 8 dBm. Hollow dots gives the average of the outcome variable (listed below each figure) for a signal bin of 4 dBm.

for different levels of signal strength (measured in dBm). The solid vertical line gives the coverage cutoff. Signal strength is normalized so that negative (positive) dBm values represent no coverage (coverage). The figures provide two levels of smoothing: solid dots represent the averages of the outcome variables for 8 dBm signal strength bins while hollow dots use 4 dBm bins. Recall that in this paper, the sensitivity threshold used for the cutoff is 97 dBm.

Figure 5 shows that the likelihood of an attack exhibits a significant drop as soon as an area receives enough signal strength to allow cell phone use. This is a consistent result across the different insurgent violence categories. Specifically, the likelihood of a direct fire incident drops by about 2.5 percentage points while the likelihood of an IED explosion decreases by about 1.5 percentage points. Indirect fire incidents show only a slight decrease at the coverage boundary.

5.2 Sharp RD Estimates

Table 3 presents the estimated effect of coverage on insurgent violence using Equation (5). The outcome variables are indicators for whether there was at least one incident in cell i of the grid. All results use the local linear regression specification of Equation (5) described in section 4.1. All specifications include controls for elevation and slope and standard errors clustered at the district level.

Results in Panel A indicate a significant drop in all violence categories once signal strength surpasses the cutoff required for cell phone coverage. In particular, comparing violence levels for areas within 10 to 14 dbm of the cutoff, the likelihood of an insurgent attack drops by about 2.4 percentage points on the coverage side. For reference, the average likelihood of an attack in non-coverage areas within the specified bandwidth is about 4%. The drop in the likelihood of an attack is explained by statistically significant drops in all attack types: direct fire, IEDs, and indirect fire attacks (columns (2)-(4)).

Using the same bandwidths of analysis, Panels B and C add to the baseline specification a set of covariates (Panel B) and a more flexible polynomial in elevation and slope (Panel C).[12] In both cases, we document qualitatively similar drops in all attack types. In the case of indirect fire attacks, although the statistical significance goes away, the magnitude remains robust. Overall, Panels B and C provide evidence that, consistent with a valid RD design, our results are robust to the inclusion of covariates (Panel B) and that the observed drops in violence are the result of coverage access and not the terrain characteristics used to model coverage using the ITM (Panel C). Lastly, we note that the results are also robust to different choices of bandwidth.[13]

We proceed by exploring whether our main results are robust to the functional form of Equation (5) and the sample used. To do this, Table 4 replicates the results in Table 3 using a nonlinear specification of Equation (5) (Panel A), and an alternative sample (Panel B). Given that the outcomes of interest are binary variables, Panel A presents results using a Probit model. Note that the estimated marginal effects are comparable in magnitude and precision to the coefficients reported in Table 3. Panel B defines a 2-km radius around a village's location (instead of the 1-km by 1-km grid) as the unit of spatial aggregation. Data on the location of villages comes from the Measuring Impacts of Stabilization Initiatives project (MISTI, 2013) sponsored by US Agency

[12] Panel B includes the elevation and slope controls of Panel A and adds controls for distance to: health facility, primary roads, secondary roads, rivers. Panel C uses a topography polynomial given by: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$

[13] Refer to Appendix Figure A7 for estimated RD coefficients for all violence categories using a host of bandwidths.

Table 3: Effect of Coverage on Insurgent Violence, RD Estimates

	Dep. Variable = $\mathbb{1}\{\text{Number of attacks} > 0\}$			
	Any attack	Direct fire attack	IED explosion	Indirect fire attack
	(1)	(2)	(3)	(4)
Panel A: Baseline RD Estimates				
Coverage	-0.024***	-0.020**	-0.008***	-0.003*
	(0.008)	(0.008)	(0.003)	(0.002)
Panel B: Baseline RD Estimates + Controls				
Coverage	-0.021***	-0.018**	-0.007**	-0.002
	(0.008)	(0.008)	(0.003)	(0.002)
Panel C: Baseline RD Estimates + Topography polynomial				
Coverage	-0.019***	-0.016**	-0.006**	-0.002
	(0.007)	(0.007)	(0.003)	(0.002)
Observations	108,471	113,436	97,279	84,333
Number of districts	371	371	371	371
Mean Outside coverage	0.040	0.030	0.018	0.0042
Bandwidth (dBm)	13.7	14.1	12.0	9.36

Notes: Results present estimates of β using a local linear regression specification of equation (5). Panel A controls for elevation and slope. Panel B adds controls for distance to: health facility, primary roads, secondary roads, rivers. Topography polynomial is given by: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. Optimal bandwidth chosen as in Calonico et al. (2014). All specifications use standard errors clustered at the district level. *, **, and *** indicate 10, 5, and 1 percent significance respectively.

for International Development (USAID). The data include geographic coordinates from various data sources between the years 2012 and 2013 for more than 45,000 villages across Afghanistan. Results in Panel B estimate Equation (5) using an indicator for violence within a 1km radius of a village's location. Consistent with the results reported above, violence drops in villages just receiving sufficient signal strength: the likelihood of an attack decreases by about 11 percentage points while the likelihood of an IED explosion decreases by close to 10 percentage points.

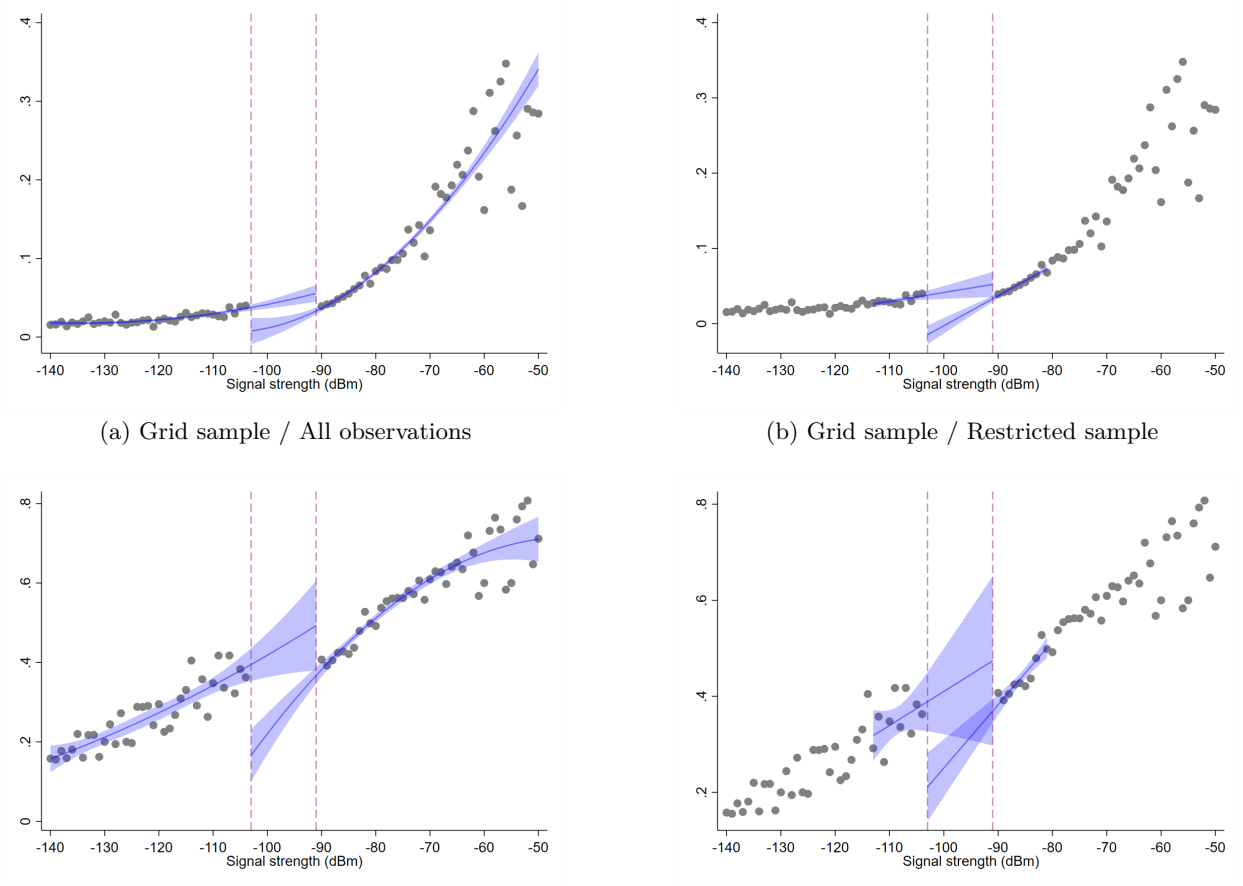
5.3 Interval RD Estimates

Table 5 presents our Interval RD results. Specifically, it presents estimates of τ in Equation (6) using the grid sample (columns 1 and 2) and the village sample (columns 3 and 4). Columns 1 and 3 use all observations outside of interval $C = [c_-, c_+]$ to estimate the potential outcome functions using a second degree polynomial. Columns 2 and 4 restrict the sample to observations within

Table 4: Effect of Coverage on Insurgent Violence, Alternative Sample and Functional Form

	Dep. Variable = $\mathbb{1}\{\text{Number of attacks} > 0\}$			
	Any attack	Direct fire attack	IED explosion	Indirect fire attack
	(1)	(2)	(3)	(4)
Panel A: Functional Form: Probit Estimates				
Coverage	-0.290***	-0.288***	-0.183***	-0.204*
	(0.055)	(0.062)	(0.067)	(0.112)
Marginal effect	-0.024	-0.019	-0.0049	-0.0027
Observations	108,471	113,436	97,279	84,333
Number of districts	371	371	371	371
Bandwidth (dBm)	13.7	14.1	12.0	9.36
Panel B: Alternative Sample: Villages				
Coverage	-0.111**	-0.114**	-0.095**	-0.038
	(0.044)	(0.045)	(0.042)	(0.049)
Observations	8,642	9,488	8,642	8,642
Number of districts	344	346	344	344
Mean Outside coverage	0.41	0.36	0.25	0.13
Bandwidth (dBm)	11.7	12.8	11.3	11.6

Notes: Results present estimates of β using a local linear regression specification (Panel A) and a Probit specification (Panel B) of equation (5). All results include controls for elevation and slope. Grid sample used in Panel B. Optimal bandwidth chosen as in Calonico et al. (2014). All specifications use standard errors clustered at the district level. *, **, and *** indicate 10, 5, and 1 percent significance respectively.



(a) Grid sample / All observations

(b) Grid sample / Restricted sample

(c) Village sample/ All observations

(d) Village sample / Restricted sample

Figure 6: Interval Regression Discontinuity Plots

Notes: Red dashed lines denote interval $C = [c_-, c_+]$ discussed in section 4.3. Solid dots give the average of the outcome for a signal bin of 1 dBm. Panels b and d restrict sample to observations within 10dBm outside of C and use a linear polynomial. Panels a and c include all observations and use a quadratic polynomial.

10dBm outside the interval $C = [c_-, c_+]$. We find that coverage leads to a 2.8-3.4 percentage point drop in the likelihood of any attack occurring a grid cell (columns 1 and 2). Similarly, villages with coverage experience a 13.7-16.7 percentage point drop in the likelihood of any attack near the village. This results are quite similar to our sharp RD results presented in column 1 of Table 3 and column 1 of Table 4.

Figure 6 presents graphical results of the interval RD design. The design in panels a and b in the figure correspond to columns 1 and 2 in Table 5 while panels c and d correspond to columns 3 and 4. The results in column 1 of Table 5, for example, can be interpreted as the estimated weighted average (i.e., τ in Equation (6)) of all the discontinuities within C in panel a of Figure 6.

Table 5: Effect of Coverage on Insurgent Violence, Interval RD Estimates

	Grid data		Village data	
	(1)	(2)	(3)	(4)
Coverage	-0.028	-0.034	-0.167	-0.137
	[-0.044, -0.012]	[-0.054, -0.015]	[-0.282,-0.051]	[-0.307,0.032]
Obs.	126,744	61,385	19,751	8,222
Bwidth (dBm)	.	10	.	10
Degree	2	1	2	1

Notes: Results present estimates of τ in Equation (6). 95% confidence intervals around τ presented in brackets. Confidence intervals calculated via bootstrap with 250 repetitions and resampling clustered at district level. Degree refers to the degree of the polynomial used in estimating the potential outcome functions in 4. Grid sample used in columns 1 and 2. Village sample used in columns 3 and 4.

5.4 Panel Regression Estimates

We analyze SIGACTS events beginning in 2002. The first Roshan cell towers were built in Kabul in 2003, and towers were installed in additional districts each year until 2011, the final year of our panel. This gives $q = 9$ coverage-year cohorts (see Figure 8). Because our setting includes one period before any units are treated, and maintains a subset of "never treated" districts, this method produces $\frac{(T-q+1)(T-q+2)}{2}$ treatment effects (τ_{rs}) which can be aggregated by weighted averaging over calendar time, cohort year, or relative time-to-treatment, known as an event study.

Figure 7 shows such an event study aggregation, representing dynamic treatment effects. The upper left graph shows the downward trend of the likelihood of any violent event following coverage for a given district (point estimates and errors are shown in Table **??**). The other graphs show similar patterns for specific attack types: IEDs, direct fire, and indirect fire. In each case, visual evidence suggests the effect is no different from zero in pre-treatment periods.

We next explore the cumulative effect of coverage, based on our expectation that effects due to additional towers is likely nonlinear. First, we consider the that mobile phone service is a type of network good; the effect on a single node receiving the technology could be multiplicative with the addition of new nodes, up to a saturating point. Additionally, in our model of insurgent violence, it's unclear whether a group like the Taliban would react consistently along the intensive and extensive margins of coverage expansion. In other words, it may be the case that rebels increase violence to halt or interrupt the introduction of coverage to a new area, but reduce it as the network expands and locals make use of the amenity. Of course, the Taliban themselves adopted new methods of violence that utilized mobile phones explicitly (i.e. as IED detonation devices), and eventually

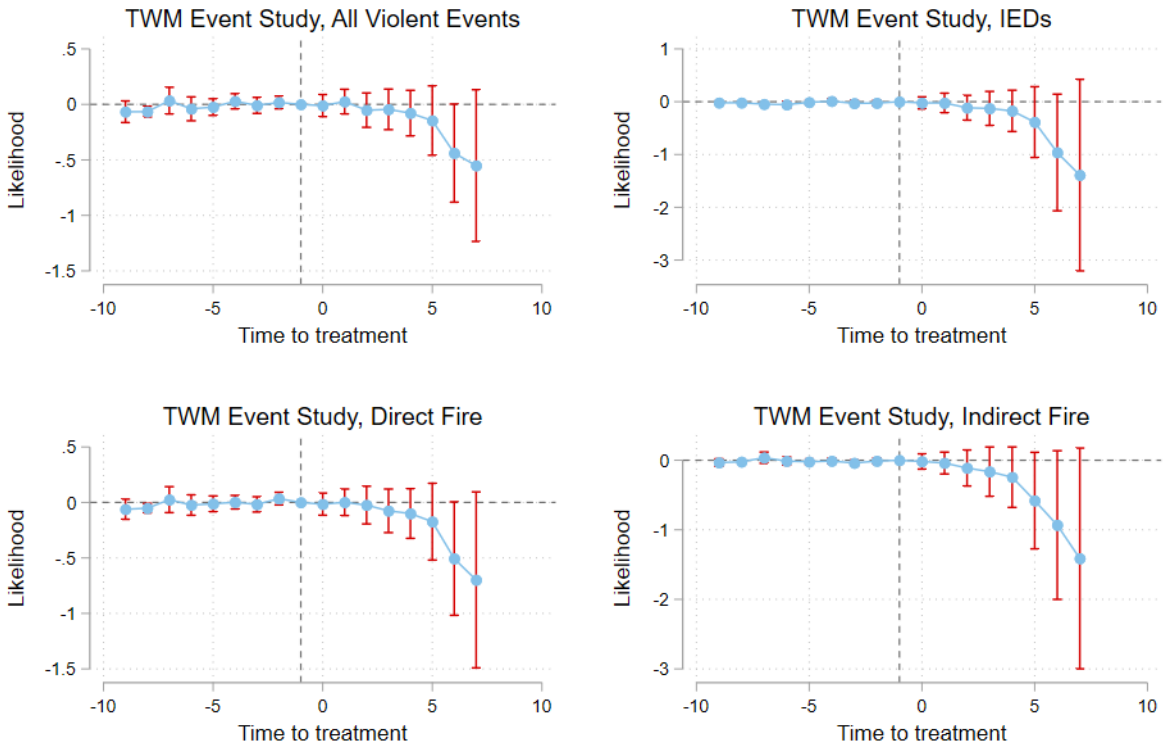


Figure 7: Two-Way Mundlak Event Study graphs. Estimation is by Poisson Pseudo-Maximum Likelihood, with errors clustered at the district level. Treatment effects are normalized relative to 1 year prior to a district receiving its first cellular tower. 95% confidence intervals are represented by spikes around each point estimate.

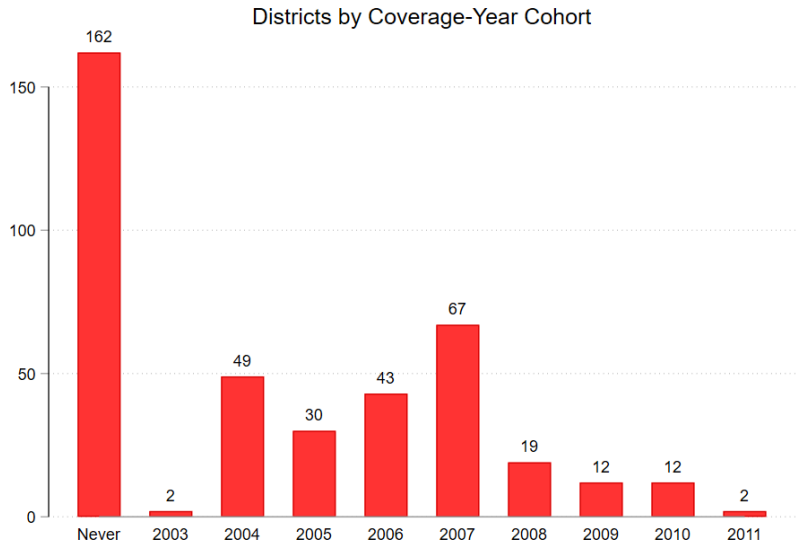


Figure 8: The number of districts in each coverage-year cohort. Every Afghan province had at least one Roshan tower by 2009, but many districts, a smaller geographic division, did not.

threatened towers directly.

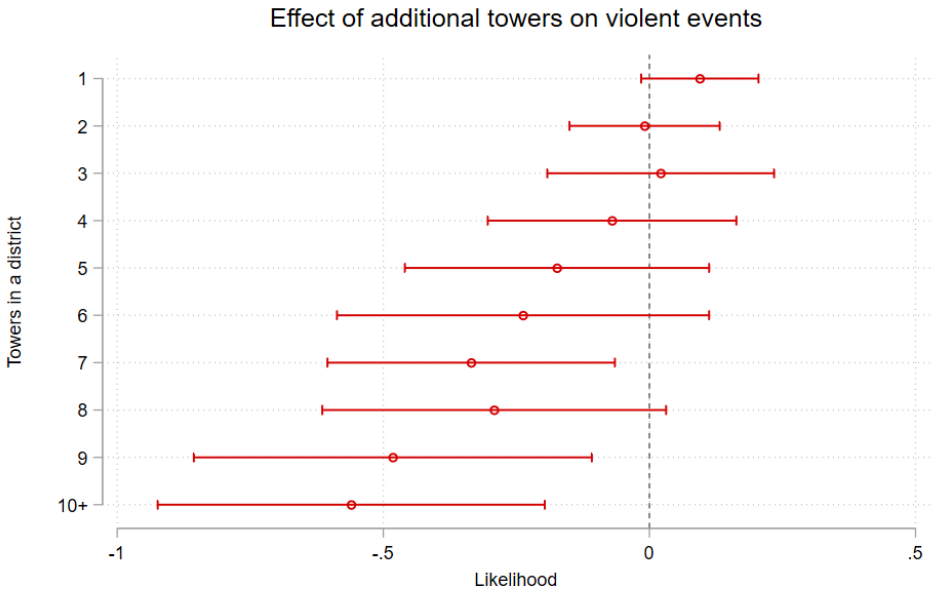


Figure 9: PPML Fixed Effect regression with district-clustered errors.

Figure 9 reveals a downward trend in the likelihood of violent events as towers are added in a district. Here we estimate a panel fixed effects model by Poisson Pseudo-Maximum likelihood, with indicators equal to 1 when a district-year cell has exactly $n \in [1, 9]$ towers, plus a binned indicator for district-years with $n \geq 10$ towers.[14]

6 Coverage-Violence Channels

Recall from our theoretical framework that coverage can affect violence through two broad channels: (1) reductions in the cost of producing violence (e.g., coordination of attacks, remote detonation of IEDs), and (2) information sharing (e.g., the transfer of insurgent-related information from civilians to the government). This section explores these channels empirically.

6.1 Violence Production Costs

We focus on two sub-channels that can lower the costs of producing violence for rebels: (i) remote detonation of IEDs using phones and (ii) improved rebel coordination.

[14] By 2015, the median number of towers per district was 1, the third quartile was 2, and the 95th percentile was 8.

6.1.1 Cellphone triggered IEDs

The results in column 3 of Table 3 and Table 4, and panel c of Figure 5 all suggest that coverage leads to a *decrease* in the likelihood of an IED. Therefore, it is unlikely that coverage is leading to a significant shift towards more effective IED detonation tactics such as cellphone-triggered IEDs. This is plausible in this context since counterinsurgency forces rapidly adapted to this strategy by deploying radio jamming devices designed to address cellphone-triggered IEDs.

6.1.2 Spatiotemporal coordination

We do not observe rebel coordination therefore, we cannot directly explore whether coverage makes rebels more effective when coordinating attacks. However, we provide evidence on how coverage affects different measures of spatial and temporal patterns of violence. Changes in spatiotemporal behavior might suggest changes in rebel coordination.

Changes in the spatial clustering of violence over time may suggest changes in coordination among groups responsible for the violence. Violent incidents occurring in close proximity may imply some level of organization or collaboration among the perpetrators.

We rely on two spatial statistics from the geography literature to assess changes in the spatial clustering of violence across districts and years. Specifically, we compute the Getis-Ord Gi* statistic and Anselin Local Moran's I, for each district and year. Broadly speaking, the statistics compare the number of violent events in a given district with its contiguous districts. [15]. Both measures are expressed as z-scores with higher z-scores indicating significant clustering of events in the district.

Following the construction of these spatial statistics, we use a district-by-year panel and employ a difference-in-differences setup to investigate how these statistics change after the introduction of a cell tower in a district and the number of towers in the district. Our setup includes the conventional district and year fixed effects as well as a control for district population. Tables 7 and 6 present these estimates. We find no discernible changes in the spatial clustering of violence in districts after they obtain a cell tower or when the number of towers increases.

We also examine changes in temporal patterns of district-level violence after the introduction of a cell tower. Specifically, for each district and year, we calculate the standard deviation, absolute skewness, and kurtosis for the distribution of days-of-the-week (0-6) and months-of-the-year (1-12) in which violent incidents take place.[16]

[15]The Getis-Ord Gi* is computed using the Hot Spot Analysis tool in ArcGIS Pro Esri (2024b). Anselin Local Moran's I is computed using the Cluster and Outlier Analysis tool in ArcGIS Pro Esri (2024a)

[16]High absolute values of skewness indicate greater skewness in the temporal distribution of violence. Kurtosis

Table 6: Effect of Cell Towers on Spatial Statistics

	HotSpot Analysis Z-score Contiguity	Cluster and Outlier Analysis Z-score Contiguity
Towers	-0.016	0.006
	(0.010)	(0.010)
Dep-var Mean	0.021	0.53
# Cluster	398	398
# Observations	5174	5174
Year Fixed Effect	YES	YES
District Fixed Effect	YES	YES

Linear Models With Many Levels of Fixed Effects regressions control for district population.
All standard errors clustered at the district level.

Table 7: Effect of First Tower on Spatial Statistics

	HotSpot Analysis Z-score Contiguity	Cluster and Outlier Analysis Z-score Contiguity
First Tower	0.100	0.144
	(0.105)	(0.116)
Dep-var Mean	0.021	0.53
# Cluster	398	398
# Observations	5174	5174
Year Fixed Effect	YES	YES
District Fixed Effect	YES	YES

Linear Models With Many Levels of Fixed Effects regressions control for district population.
All standard errors clustered at the district level.

We replicate our previous difference-in-differences setup to investigate whether the temporal distribution of violence changes after the introduction of a cell tower in a district or when the number of towers in a district increases. Tables 9 and 8 present the results for the effect of first cell tower installation and for the number of cell towers, respectively.

Overall, the results provide some evidence of a temporal response in violence. Table 9 shows that the temporal distribution of attacks increases in kurtosis (i.e., becomes more peaked/heavy tailed) after the introduction of the first cell tower. The standard deviation and absolute skewness do not change significantly. Table 8 shows that the temporal distribution becomes more compact, more symmetric, and less peaked/lighter tailed with each additional tower installed. In all, this suggests that a larger number of attacks becomes more concentrated around the modal month/day as the number of cell towers increase.

measures the "tailedness" or "peakedness" of a distribution. Higher kurtosis indicates a sharper peak and heavier tails. Lower kurtosis suggests a flatter distribution with lighter tails.

Even though we find no major changes in the spatial concentration of attacks following the deployment of cell towers, we do find evidence of a temporal response in the timing of attacks. Based on our data, we cannot say whether this can be interpreted as improved insurgent coordination. We can only interpret this as suggestive evidence that cell phone access is leading to changes in temporal coordination. However, we note that even if this is leading to more violence due to improved coordination, the overall drop in violence documented in section 5 suggests that, in equilibrium, additional channels such as information sharing offset any gains to insurgents resulting from improved coordination. We discuss this channel next.

Table 8: Effect of Cell Towers on Temporal Statistics

	Standard Deviation Month of Year	|Skewness| Month of Year	Kurtosis Month of Year	Standard Deviation Day of Week	|Skewness| Day of Week	Kurtosis Day of Week
Towers	-0.005***	-0.002**	-0.007***	0.000	-0.001**	-0.005**
	(0.002)	(0.001)	(0.003)	(0.001)	(0.001)	(0.002)
# Cluster	303	300	300	303	302	302
# Observations	1820	1765	1765	1820	1785	1785
Year Fixed Effect	YES	YES	YES	YES	YES	YES
District Fixed Effect	YES	YES	YES	YES	YES	YES
Baseline Group Average	2.539	0.401	2.014	1.900	0.260	1.680
# Control Group Observations	886	846	846	886	866	866

Linear Models With Many Levels of Fixed Effects regressions control for district population.
Standard errors are clustered by district.
The difference in observations is due to the exclusion of days and months without violent incidents.
The baseline group consists of districts without tower construction.

Table 9: Effect of First Tower on Temporal Statistics

	Standard Deviation Month of Year	|Skewness| Month of Year	Kurtosis Month of Year	Standard Deviation Day of Week	|Skewness| Day of Week	Kurtosis Day of Week
First Tower	0.038	0.017	0.232**	0.067	0.027	0.102**
	(0.120)	(0.042)	(0.090)	(0.067)	(0.031)	(0.047)
# Cluster	303	300	300	303	302	302
# Observations	1820	1765	1765	1765	1785	1785
Year Fixed Effect	YES	YES	YES	YES	YES	YES
District Fixed Effect	YES	YES	YES	YES	YES	YES
Baseline Group Average	2.539	0.401	2.014	1.900	0.260	1.680
# Control Group Observations	886	846	846	886	866	866

Linear Models With Many Levels of Fixed Effects regressions control for district population.
Standard errors are clustered at the district level.
The difference in observations is due to the exclusion of days and months without violent incidents.
The baseline group consists of districts without tower construction.

6.2 Information Sharing

We do not observe information shared between civilians and counterinsurgency forces. However, we can provide indirect evidence on information flows.

We rely on three strategies for this. First, we explore how the rate of found-and-cleared IEDs changes with coverage. Found-and-cleared IEDs are IEDs that were planted but were unsuccessful

because they were found and safely disposed by counterinsurgency forces. Information sharing plays a key role in whether IEDs are "found". If coverage leads to an increase in information-sharing then we should expect the rate of found-and-cleared IEDs to be significanlty higher once a location has enough coverage for reporting to take place.

Second, we use data on millions of call detail records (CDRs) for the largest mobile network operator in Afghanistan to explore call behavior around the date of a found-and-clear IED. If information sharing plays a role in explaining the reported drop in violence in coverage areas, we should expect increased call frequency just prior to a found-and-clear IED. We do not directly observe whether individuals called to report an IED but we can observe whether there is a spike in all calls/phone usage just prior to a found-and-cleared IED.

Third, we look at how our estimated RD coefficient varies depending on temporal and spatial markers that correlate with civilian presence (how the effect of coverage varies in morning hours, near populated areas, etc.)

Lastly, we also explore whether coverage lowers the risk of retaliation by looking at an outcome that combines assassination, kidnapping, murder threats, and reports of intimidation.

Table 10 Column (1) tests whether coverage increases the "failure rate" of IEDs. Failure rate refers to the ratio of found-and-cleared IEDs over all IEDs. We find that coverage leads to a significant increase in the rate of found-and-cleared IEDs. Columns 2 and 3 explore whether this effect is larger near populated areas (i.e., closer in distance to a village) and therefore information sharing might play a bigger role. We find that coverage increases the failure rate if close to a village, however, coverage is not significant in finding more IEDs as we move away from civilians' "eyes".

We explore call behavior around a found-and-cleared IED using call detail records for the largest mobile network operator in Afghanistan at the time. These data contain the daily centers of gravity (COG) for each individual user. The COG is an estimate of the individual's location based on the latitude and longitude of the nearest towers to the individual. This allows to determine whether a given user was geographically close to a found-and-cleared IED. Refer to Blumenstock et al. (2024) for more details on these data and the construction of COGs. We then create an indicator for whether a given user has used their phone in a given day based on whether the COG was able to be calculated on that day.

Figure 10 plots the coefficients from a regression using leads and lags around the day that an IED is found and cleared for individuals that were withing five kilometers of the found-and-cleared IED. We find that the likelihood of phone usage spikes just the day before an IED is found for

Table 10: Effect of Coverage on Unsuccessful IEDs

	Dependent variable= IED Failure Rate			
	RD	$RD_{\leq 2km}$	$RD_{>2km}$	DD-RD
	(1)	(2)	(3)	(4)
Coverage	0.097*	0.136**	-0.007	-0.010
	(0.051)	(0.060)	(0.083)	(0.049)
Near village				-0.044
				(0.051)
Coverage×Near village				0.075
				(0.050)
Observations	1468	1086	382	1468
Number of districts	141	127	63	141
Mean Outside coverage	0.60	0.60	0.57	0.59
Bandwidth (dBm)	6.20	6.20	6.20	6.20

Notes: "IED Failure Rate" is given by the share of IEDs that are found and cleared out of all exploded and found and cleared IEDs. All specifications control for elevation and slope and standard errors clustered at the district level. "DD-RD" refers to a difference-in-differences design within the bandwidth used in column (1). "Near village" means that grid cell is within 2 kilometers of a village's location. Columns (2) and (3) restrict the analysis to grid cells within and beyond 2 kilometers of a village's location, respectively. ***, **, * indicate 10, 5, and 1 percent significance, respectively.

individuals that are close to the found IED relative to those farther away. Phone usage remains significantly higher the day the IED is found and the following day. Phone usage goes back to baseline levels after that. This provides plausible evidence on information sharing as individual phone usage spikes around the time of IEDs that were precisely found by counterinsurgency forces.

Table 11 presents our main RD estimates separately by grid cells that are within two kilometers of a population center (column 1), within one kilometer of a road (column 4), and within one kilometer of a road and village (column 7). Columns 2, 5, and 8 present the respective counterfactual cases. As reference, Appendix Figure A5 shows a sample of grid cells within two kilometers of a road. If significant information sharing takes place, we should expect a significant effect near populated centers and roads where citizens reside and work.

Table 11 shows that most of the effect documented in our main results is explained by grid cells near populated centers and roads. The effect of coverage away from villages and roads is null in most specifications. Column (10) provides further evidence: with each additional 100 population increase, the presence of coverage decreases the likelihood of an insurgent attack by two percentage points.

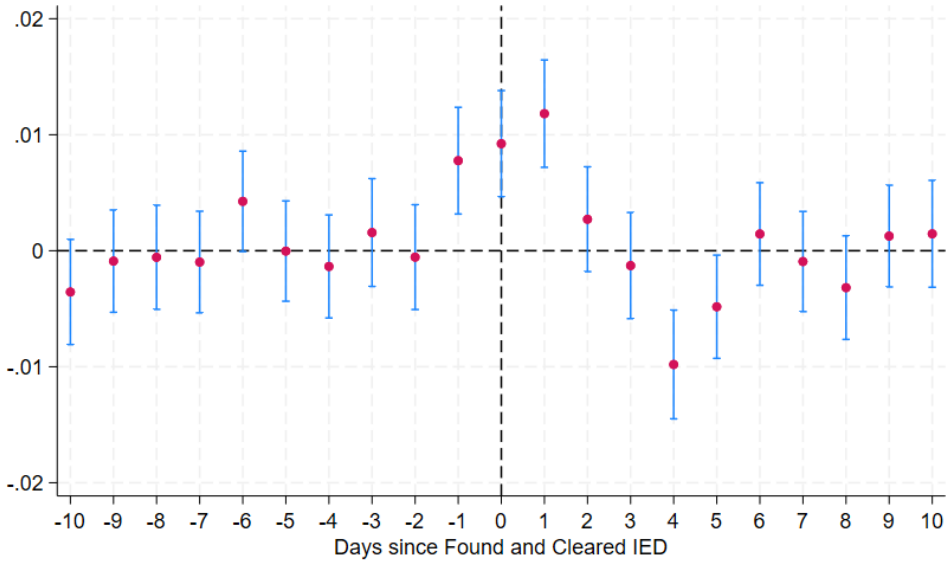


Figure 10: Likelihood of Phone Use

Notes: Each dot represents the estimated leads and lags coefficients of a regression that uses whether an individual has used their phone. Leads and lags are relative to the day where a found-and-cleared IED takes place within 5 kilometers of the individual

Table 11: Effect of Coverage on Likelihood of Attack near Villages and Roads

	Dep. Variable = $\mathbb{1}\{\text{Number of attacks} > 0\}$									
	Villages			Roads			Villages and Roads			Population
	$RD_{\leq 2km}$	$RD_{>2km}$	DD-RD	$RD_{\leq 1km}$	$RD_{>1km}$	DD-RD	$RD_{\leq 1km}$	$RD_{>1km}$	DD-RD	DD-RD
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Coverage	-0.05***	-0.01*	-0.01**	-0.03**	-0.01	-0.02**	-0.04***	-0.01	-0.02***	-0.03**
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Near village			0.05***							
			(0.02)							
Coverage×Near village			-0.02*							
			(0.01)							
Near road						-0.01**				
						(0.01)				
Coverage×Near road						-0.00				
						(0.01)				
Near road and village									0.03**	
									(0.01)	
Cov.×Near road and village									-0.01	
									(0.01)	
Population										0.04***
										(0.01)
Coverage×Population										-0.02***
										(0.01)
Observations	44228	53051	97279	26494	17973	44467	54708	42571	97279	44228
Number of districts	361	362	371	348	347	357	363	361	371	361
Mean Outside coverage	0.090	0.020	0.020	0.065	0.045	0.057	0.070	0.018	0.018	0.090
Bandwidth (dBm)	11.8	11.8	11.8	11.8	11.8	11.8	11.8	11.8	11.8	11.8

Notes: All specifications control for elevation and slope and standard errors clustered at the district level."Near village" means that grid cell is within 2 kilometers of a village's location. "Near village and road" means that grid cell is within 1 kilometer of a village's location and a road. Column (1) restricts the analysis to grid cells within 2 kilometers of a village's location. Column (4) restricts the analysis to grid cells within 1 kilometer from a village and a road. Column (2) restricts the analysis to grid cells beyond 2 kilometer of a village's location. Column (5) restricts the analysis to grid cells beyond 1 kilometer of a village's location or a road. "DD-RD" refers to a difference-in-differences design restricted to the optimal bandwidth used in the RD design. ***, **, * indicate 10, 5, and 1 percent significance, respectively.

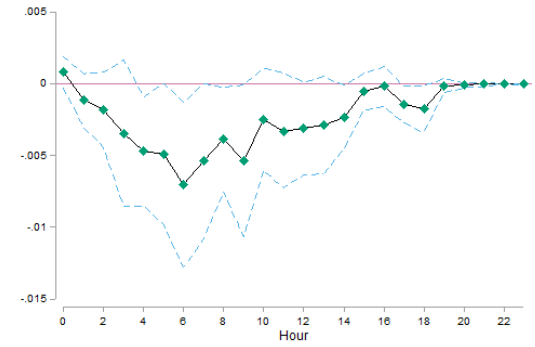
Table 12: Effect of Coverage by Time of Day

	Dependent variable: Any Attack=1		
	Daylight (7AM-5PM) (1)	Astronomical twilight (2)	Night (8PM-4AM) (3)
Coverage	-0.015**	-0.012***	-0.007**
	(0.006)	(0.004)	(0.003)
Observations	90993	90993	84333
Mean Outside coverage	0.030	0.017	0.017
Bandwidth (dBm)	10.7	10.6	9.19

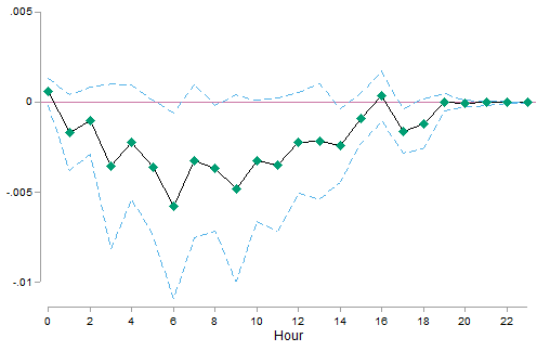
Notes: All specifications control for elevation and slope and standard errors clustered at the district level. ***, **, * indicate 10, 5, and 1 percent significance, respectively.

Table 12 splits our main RD results by time of day when attacks happen. If information sharing is a key channel, then the effect of coverage should be significant during hours when visibility is clear and therefore citizens can better observe insurgents. Column 1 shows that in coverage areas, the number of daytime attacks is significantly lower than in non-coverage areas. Column 3 shows a similar result for nighttime attacks, although the magnitude is lower than in column 1. Figure 11 shows the estimated RD coefficients by hour of day. Note that the effect of coverage is large and significant for attacks happening in the morning hours and less pronounced during the evening hours. Overall, these results suggest that information sharing can be an important channel explaining the drop in insurgent attacks due to the presence of coverage.

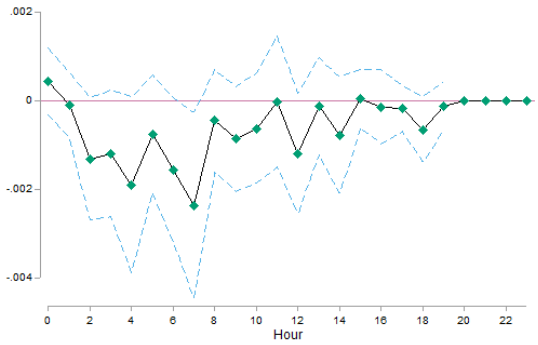
Lastly, coverage can shield civilians from retaliation if they engage in information sharing with counterinsurgency forces. In coverage areas, you can report on insurgent activity using your phone and shielded from their view. In non-coverage areas, information sharing might be riskier with civilians having to physically approach government teams. Coverage can thus spur information sharing via this shielding mechanism. We examine this by looking at how coverage affects reported incidents that can capture retaliation. Specifically, we estimate our baseline RD specification using an indicator for whether there is an assasination, kidnapping, murder threat, or intimidation report in a given grid cell. These results are presented graphically in Figure 12. We find no major evidence of a significant drop in these types of violent incidents. This may suggest that retaliation shielding is not a primary driver of information sharing.



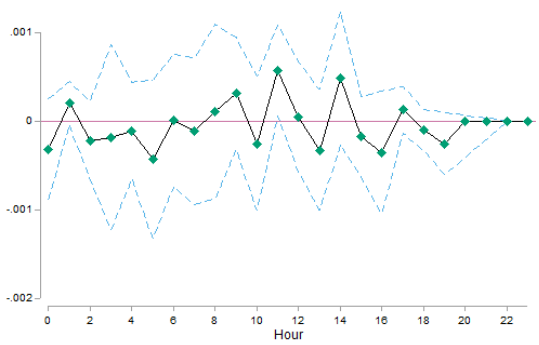
(a) Any attack



(b) Direct fire



(c) IED



(d) Indirect fire

Figure 11: Effect of Coverage by Time of Day

Notes: Estimated effect of coverage on insurgent violence by time of day when attacks happen.

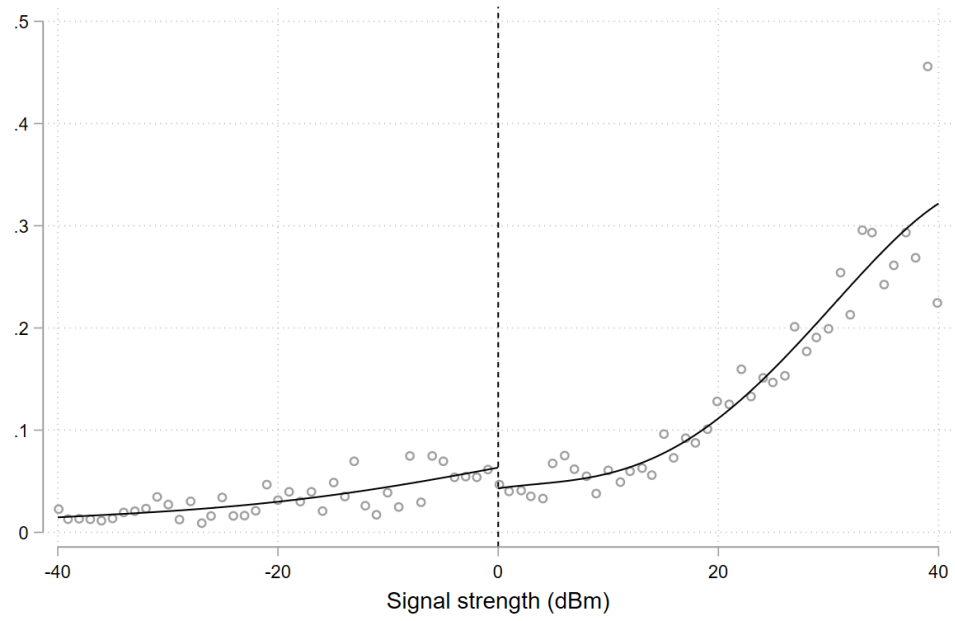


Figure 12: RD Plot of reported assassination, kidnapping, murder threat, or intimidation

Notes: RD plot of whether there is a reported assassination, kidnapping, murder thereat, or intimidation in a given grid cell.

7 Conclusion

This paper examines the impact of mobile phone coverage on insurgent violence. In theory, access to coverage can lower violence by increasing the flow of information from civilians to the government and by shielding informers from retaliation. On the other hand, cell phone access can increase violence by reducing the cost of producing violence (e.g., facilitating coordination among insurgents, remote detonation of IEDs). To answer this question, we propose a novel method that can be employed by researchers studying the impact of mobile phone coverage or any outcome of interest. Specifically, we estimate a high spatial resolution radio-wave propagation model that uses variations in terrain topography and the spatial distribution of mobile phone towers to predict signal strength on the ground for each cell of a $1X1$ kilometer grid of Afghanistan. The predicted signal strength is then used in a regression discontinuity design that compares grid cells within a small bandwidth around the signal strength threshold required for coverage. At this margin, access to coverage is mostly determined by minor exogenous changes in terrain features that lead to arbitrary diffraction and blocking of the signal. We find considerable evidence that the net effect of access to mobile phone technology is to *lower* insurgent violence. Specifically, grid cells with just enough coverage

experience a 2 percentage point drop in the likelihood of any attack and a 0.8 percentage point drop in the likelihood of an IED. Additionally, we employ a district panel of tower installations over time, allowing an alternative difference-in-differences approach that shows further evidence of the decreasing effect of coverage on violence. This effect remains robust even in areas where community norms are favorable to insurgents. We test the strength of potential channels through which coverage may affect violence. Changes in the spatial and temporal distributions of attacks could suggest insurgents respond to coverage through increased coordination, or produce more violence due to reductions in its cost. While we find evidence of a response through the temporal pattern of attacks, this does not identify improved coordination outright, given the overall drop in violence shown by our main results. Further analysis suggests that information gathering is likely a key mechanism. The deterring effect of coverage is significantly larger in cells where detection of insurgent activities by civilians is more likely: near populated areas, near primary roads, and during morning hours. Similarly, the effect of coverage on the failure rate of attacks–measured as the share of unsuccessful IEDs–significantly increases in these cells.

References

Barreca, A. I., M. Guldi, J. M. Lindo, and G. R. Waddell (2011). Saving babies? revisiting the effect of very low birth weight classification. *The quarterly journal of economics 126*(4), 2117–2123.

Berman, E., J. N. Shapiro, and J. H. Felter (2011). Can hearts and minds be bought? the economics of counterinsurgency in iraq. *Journal of Political Economy 119*(4), 766–819.

Blumenstock, J. E., M. Callen, T. Ghani, and R. Gonzalez (2024). Violence and financial decisions: Evidence from mobile money in afghanistan. *Review of Economics and Statistics 106*(2), 352–369.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Cattaneo, M. D., R. Titiunik, G. Vazquez-Bare, and L. Keele (2016). Interpreting regression discontinuity designs with multiple cutoffs. *The Journal of Politics 78*(4), 1229–1248.

Condra, L. N., J. D. Long, A. C. Shaver, and A. L. Wright (2018). The logic of insurgent electoral violence. *American Economic Review 108*(11), 3199–3231.

Crabtree, C. and H. L. Kern (2018). Using electromagnetic signal propagation models for radio and television broadcasts: An introduction. *Political Analysis 26*(3), 348–355.

Eppink, D. and W. Kuebler (1994). Tirem/sem handbook.

Esri (2024a). Cluster and outlier analysis (anselin local moran's i). *ArcGIS Pro Tool Reference*. Accessed: 2024-04-19.

Esri (2024b). Hot spot analysis (getis-ord gi*). *ArcGIS Pro Tool Reference*. Accessed: 2024-04-19.

Gonzalez, R. (2019). Cell phone access and election fraud: Evidence from a spatial regression discontinuity design in afghanistan. *Working Paper*.

Gonzalez, R. and E. M. Maffioli (2024). Is the phone mightier than the virus? cellphone access and epidemic containment efforts. *Journal of Development Economics 167*, 103228.

GSMA (2019). Operator Acceptance Values for Device Antenna Performance Version 4.0. Technical report.

JAXA (2016). Alos global digital surface model "alos world 3d-30m"(aw3d30), japan aerospace exploration agency.

Lazaridis, P., A. Bizopoulos, S. Kasampalis, J. Cosmas, and Z. D. Zaharis (2013). Evaluation of prediction accuracy for the longley-rice model in the fm and tv bands.

Lee, D. S. and T. Lemieux (2010). Regression discontinuity designs in economics. *Journal of Economic Literature 48*, 281–355.

Longley, A. G. and P. L. Rice (1968). Prediction of tropospheric radio transmission loss over irregular terrain. a computer method-1968.

MISTI (2013). Measuring impacts of stabilization initiatives (misti). (Available at: http://usaidmisti.com/gis-data).

NATO (2016). Nato/rs unclassified rel giroa. *NATO/RS UNCLASSIFIED REL GIRoA*.

Open Topography (2017). Alos world 3d - 30m.

Porter, J. and P. Yu (2015). Regression discontinuity designs with unknown discontinuity points: Testing and estimation. *Journal of Econometrics 189*(1), 132–147.

Qiu, P. (2011). Jump regression analysis. *International Encyclopedia of Statistical Science*.

Razally, F. (2015). Mobile Handset Testing: A Report for OFCOM, The UK Communication Regulator. Technical report.

Seybold, J. S. (2005). *Introduction to RF propagation*. John Wiley & Sons.

Shapiro, J. N. and D. A. Siegel (2015). Coordination and security: How mobile communications affect insurgency. *Journal of Peace Research 52*(3), 312–322.

Silva, J. M. C. S. and S. Tenreyro (2006, November). The Log of Gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Wooldridge, J. M. (2021, August). Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators.

Appendix A Additional Figures and Tables



Figure A1: Cell phone-triggered IED
Source: Shapiro and Weidmann (2015)

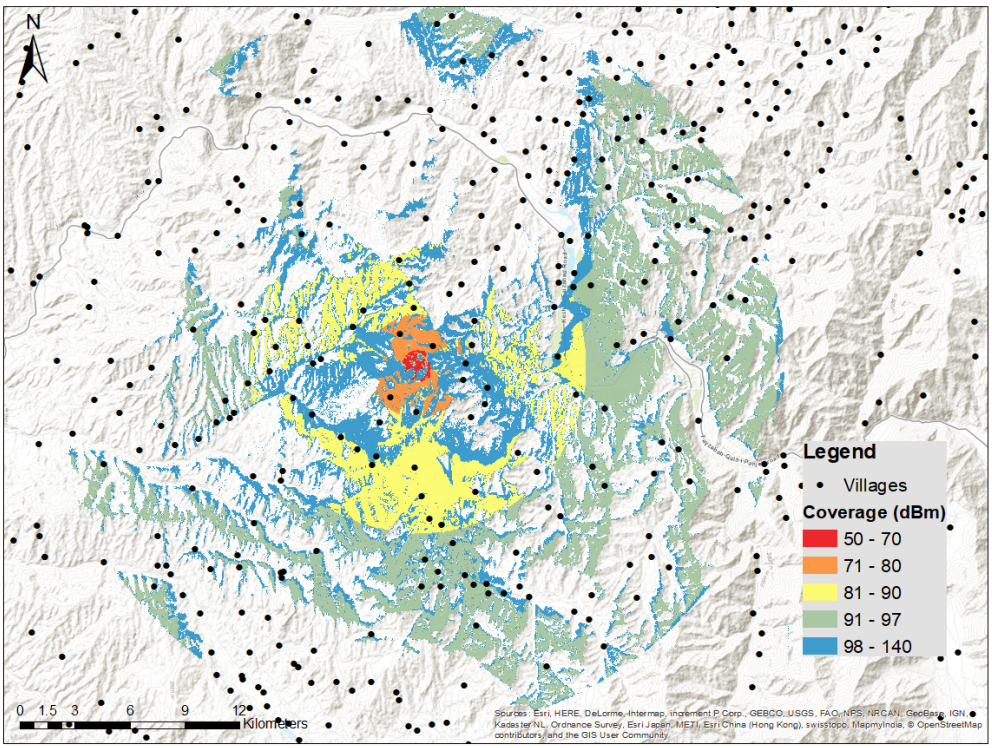


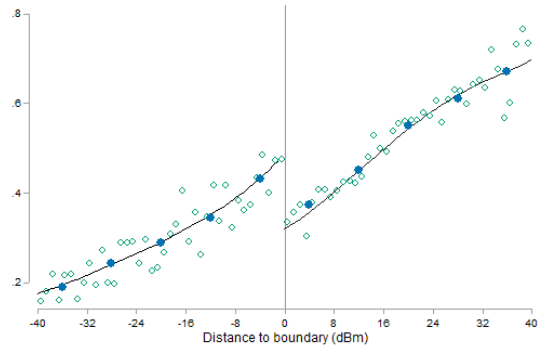
Figure A2: Sample ITM for a single tower

Notes:

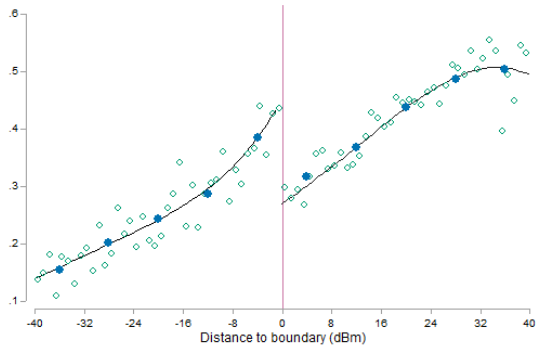
Table A1: Effect of Coverage on Hoax IEDs

Dependent variable: Hoax IED =1		
	Grid (1)	Villages (2)
Coverage	-0.001	0.002
	(0.001)	(0.022)
Observations	108471	8642
Mean Outside coverage	0.0027	0.057
Bandwidth (dBm)	13.9	11.6

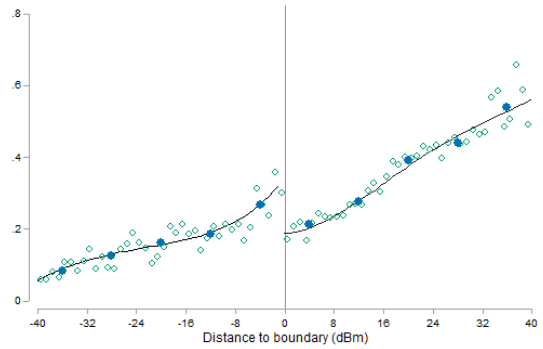
Notes:



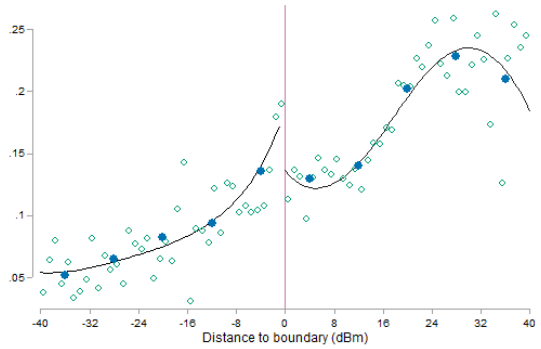
(a) Any attack



(b) Direct fire



(c) IED



(d) Indirect fire

Figure A3: Regression Discontinuity Plots

Notes:

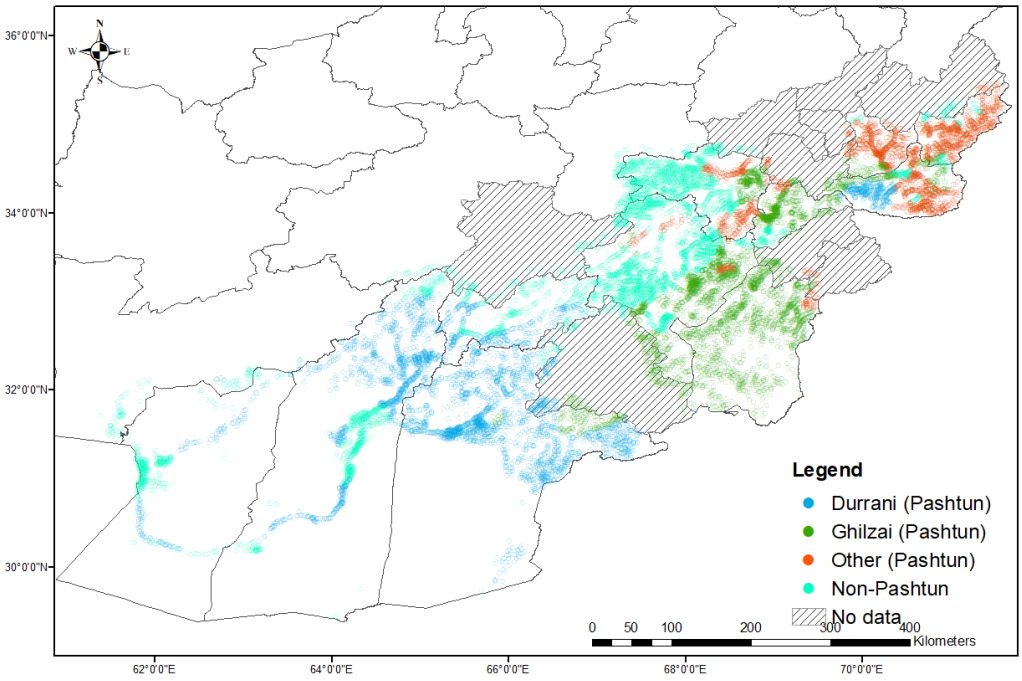


Figure A4: Pashtun Tribes of Afghanistan

Notes:

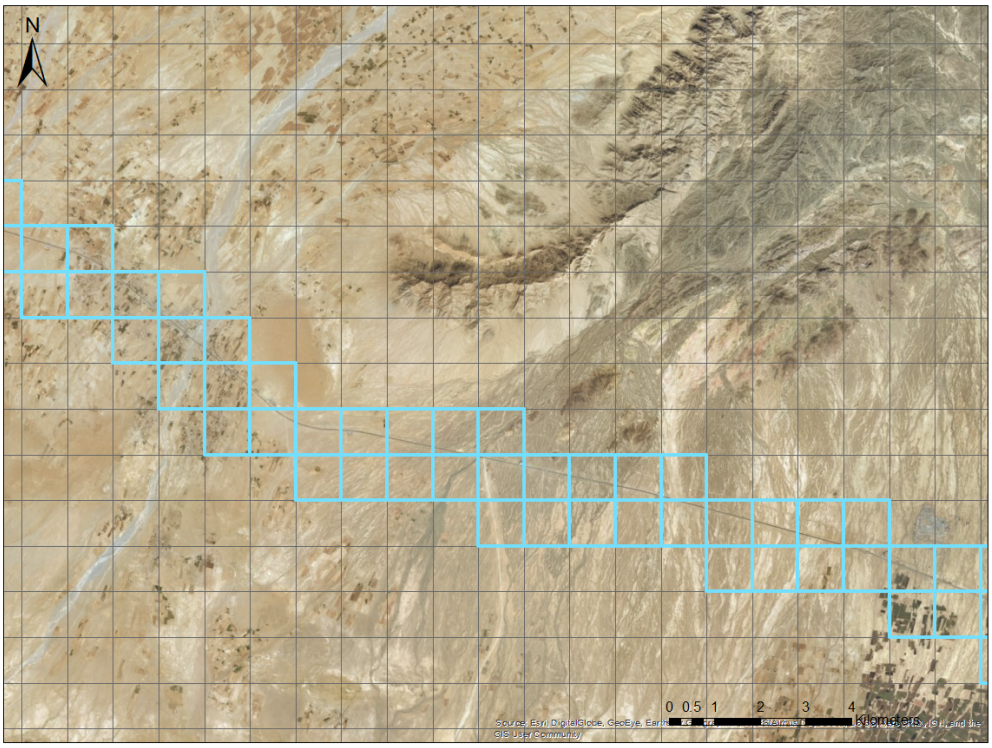
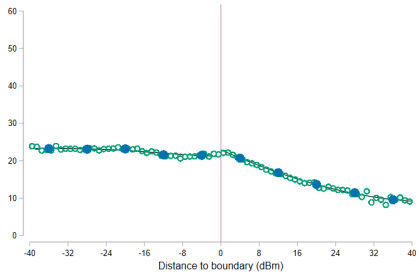
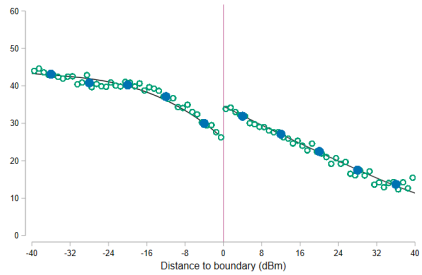


Figure A5: Sample Cell Selection

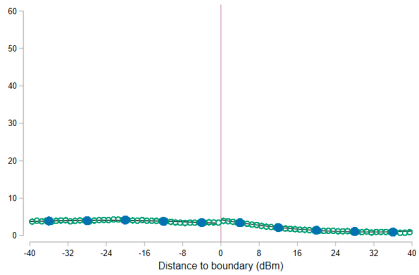
Notes:



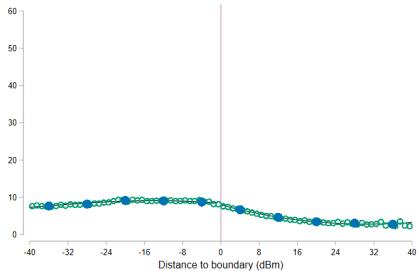
Health facility



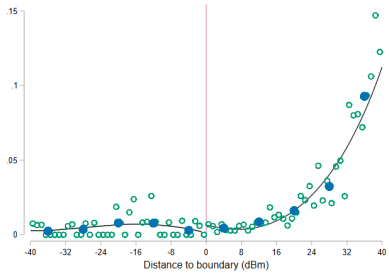
Road: primary



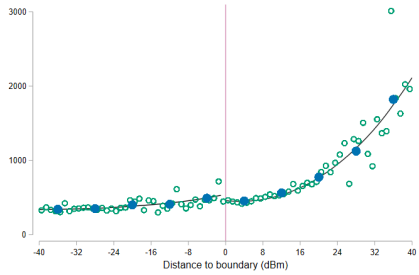
Road: secondary



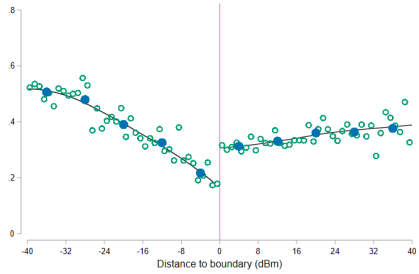
River



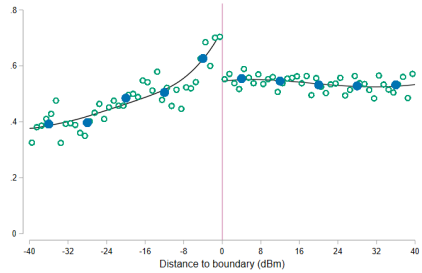
District center



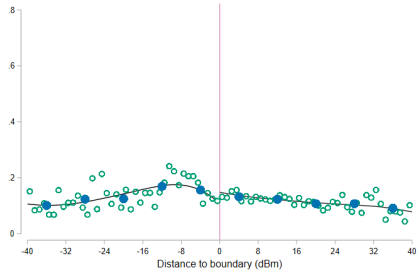
Population



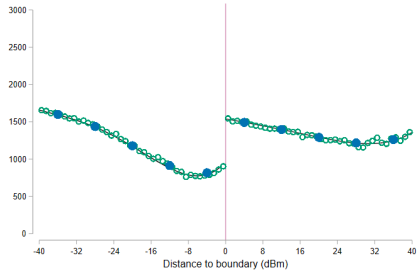
Language: Dari



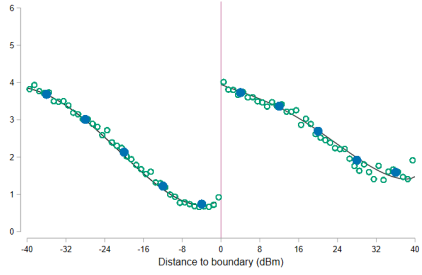
Language: Pashto



Language: Other



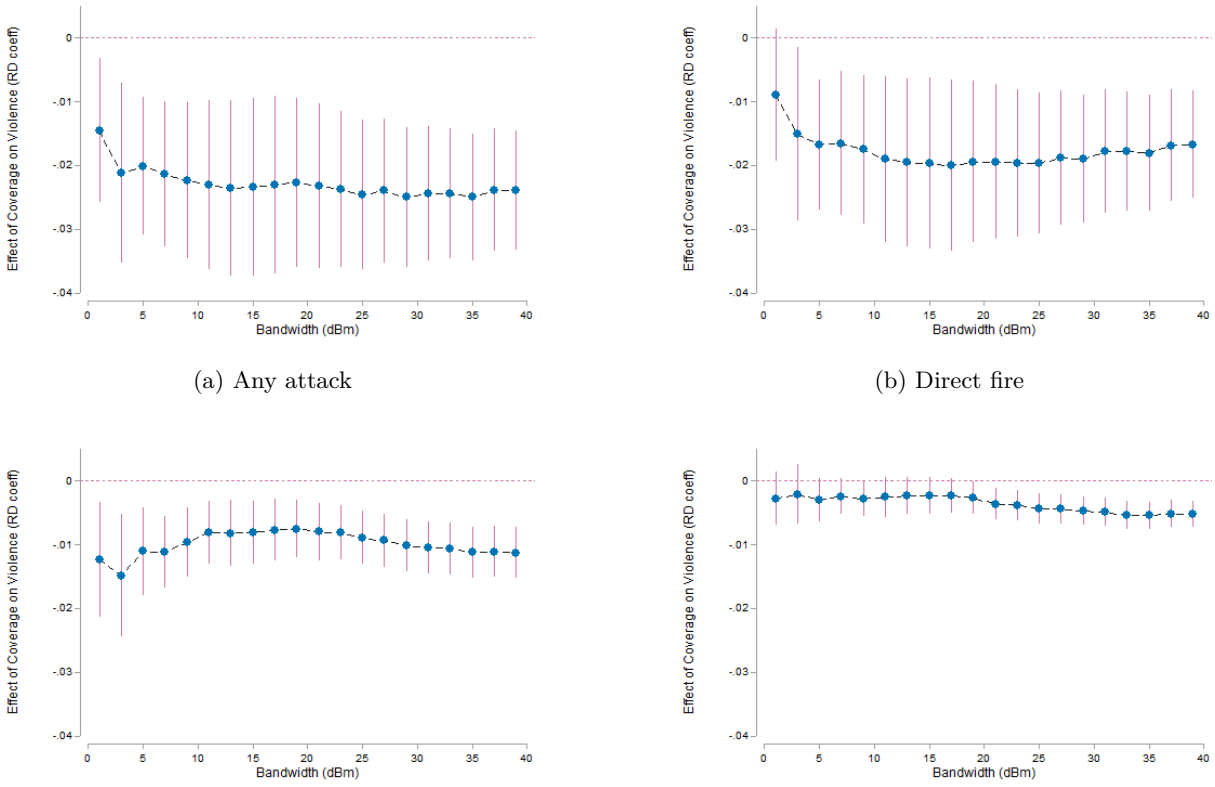
Elevation



Slope

Figure A6: Regression Discontinuity Plots, Covariate

Notes:



(a) Any attack

(b) Direct fire

(c) IED

(d) Indirect fire

Figure A7: Sensitivity of RD Results to Choice of Bandwidth

Notes:

Table A2: Effect of Coverage on Insurgent Violence, RD Estimates, Villages Sample

	Any attack (1)	Direct fire (2)	IED (3)	Indirect fire (4)
Panel A: Baseline RD Estimates + Controls				
Coverage	-0.090**	-0.100**	-0.073*	-0.034
	(0.043)	(0.044)	(0.042)	(0.048)
Observations	8642	9488	8642	8642
Number of districts	344	346	344	344
Mean Outside coverage	0.41	0.36	0.25	0.13
Bandwidth (dBm)	11.7	12.8	11.3	11.6
Panel B: Baseline RD Estimates + Topography Polynomial				
Coverage	-0.076*	-0.083**	-0.071*	-0.018
	(0.039)	(0.040)	(0.037)	(0.045)
Observations	8642	9488	8642	8642
Number of districts	344	346	344	344
Mean Outside coverage	0.41	0.36	0.25	0.13
Bandwidth (dBm)	11.7	12.8	11.3	11.6
Panel C: Probit Estimates				
Coverage	-0.309***	-0.296***	-0.257**	-0.095
	(0.100)	(0.106)	(0.117)	(0.165)
Marginal effect	-0.12	-0.11	-0.077	-0.020
Observations	8642	9488	8642	8642
Bandwidth (dBm)	11.7	12.8	11.3	11.6

Notes:

Table A3: Summary Statistics for Tribal Affiliation

	Number (1)	Share of all groups (2)	Share of Pashtuns (3)
Durrani confederation	4,343	0.24	0.37
Popalzai	200	0.01	0.02
Ghilzai confederation	4,469	0.25	0.38
Other Pashtun	3,052	0.17	0.26
All Pashtuns	11,864	0.66	1.00

Notes: