Households in Conflict Network

www.hicn.org

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

Robert Gonzalez*

HiCN Working Paper 370

May 2022

Abstract
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 of 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 diffractions 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. This effect remains robust even in areas where community norms are favorable to insurgents. 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.

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

*Department of Economics, Moore School of Business, University of South Carolina, Columbia, SC 29208
(e-mail: robert.gonzalez@moore.sc.edu)

1 Extended Abstract

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 (e.g., facilitating coordination among insurgents, remote detonation of 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 diffractions and blocking of the signals.

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. This effect remains robust even in areas where community norms are favorable to insurgents. The results are robust using a spatial first-differences design recently proposed in the literature. 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.

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 2 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 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

[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.

location of an antenna, one can therefore assess the strength of coverage at a given point on the 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 1a 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 1b. 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.

[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.

4 Regression Discontinuity Design (RD)

The ability to transfer information–voice calls, SMS, etc.–using a mobile device is a continuous function of signal strength (i.e., received power). However, this can only take place as long as the received power is higher than the receiver sensitivity threshold. This is essentially a network-specific threshold that determines whether the mobile device is receiving enough signal to be able to make voice calls, send SMS, etc.[5] In the case of a GSM network such as that of Afghanistan, receiver sensitivity typically ranges between -95 to -105 dBm. Appendix Figure A2 depicts the received signal on the ground obtained from the ITM model described in section 3.2 for a sample tower in Badakhshan province. The tower is at the center of the red shaded area, where received power is strongest. Areas in blue are receiving power below -97dBm which is under the typical mobile phone sensitivity threshold. These areas are essentially "dead zones".

This paper estimates the effect of cell phone coverage on violence by employing a regression discontinuity (RD) design that uses received power on the ground as the forcing variable and the receiver sensitivity threshold as the treatment cutoff. Specifically, we estimate:

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

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).[6] To ensure that the estimated effect on violence is not explained by topographic characteristics, all results also include a polynomial in elevation and terrain slope. β estimates the causal effect of coverage access on insurgent violence under certain assumptions. These assumptions are discussed next.

4.1 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.

[5] In more familiar terms, received power is the continuous version of the bars one sees in the phone screen while the receiver sensitivity threshold is the point where the phone goes from a single bar to "no-service".

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

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 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 seems purely determined by minor exogenous variations in topography, which we can readily 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.

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.[7] 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.

[7] 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.

5 Results

5.1 Graphical Analysis

Figure 3 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 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. In this paper, the sensitivity threshold used for the cutoff is 97 dBm. We chose this value for two reasons: first, it is within the 95-105dBm interval specified in the signal propagation literature, and second, graphical examination of the outcome variables as in Figure 3 gave a clear discontinuous change at that value.

Figure 3 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 Regression Discontinuity 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. 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).[8] 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.[9]

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 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.

5.2.1 Spatial First Differences

We assess the sensitivity of the main results to the empirical strategy used. Specifically, we use the Spatial First Differences method presented in Druckenmiller and Hsiang (2018). This method is particularly designed for gridded datasets such as the one used in this paper. Broadly speaking, the method classifies the latitude of the grid as the panel variable while the grid cells within a given latitude act as the within-panel observations. Identification is obtained by comparing outcomes for consecutive cells within a given latitude. Similarly, one can reverse the analysis and use longitude

[8] 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: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$

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

as the panel variable. We estimate:

$$\Delta v_{ij} = \beta_{SFD} \Delta D_{ij} + \Delta \tilde{R}_{ij} + \Delta \mathbf{G}_{ij} + \Delta \epsilon_{ij} \tag{6}$$

where variables are defined as in Equation (5). Subscript i denotes the panel variable. Δ denotes the first difference operator for the specified variable (e.g., $\Delta D_{ij} = D_{i,j} - D_{i,j-1}$). Table 5 presents the estimates of β_{SFD} in Equation 6. Panel A uses latitude as the panel variable and compares cells across longitudes. Consistent with the RD findings, all violence categories exhibit a significant drop in cells with coverage. Switching the analysis by comparing cells across latitudes within a given longitude (Panel B) reveals qualitatively similar results. These results suggest that the reported findings are robust to the empirical methodology used.

5.3 Insurgent Affinity and the Coverage Effect

The findings in section 5.2 provide evidence that, on average, coverage deters the likelihood of insurgent violence. This section explores whether the coverage effect remains robust in the presence of norms in the population favoring insurgents, or in the context of the theoretical model, areas with high n. For a budget constrained government engaging in counter-insurgency, it is important to know whether policies that favor coverage expansion are effective in areas with high anti-government (pro-rebel) attitudes.

We proceed by estimating Equation (5) separately in areas with high versus low anti-government (pro-rebel) attitudes. We use information on the ethnic composition of the population to assess attitudes. Specifically, we georeference tribal maps of the southeastern region of Afghanistan obtained from the Culture and Conflicts Studies program at the Naval Postgraduate School and based on the Tribal Hierarchy and Dictionary of Afghanistan (2007).[10] We then overlay the village coordinate data from the Measuring Impacts of Stabilization Initiatives project (MISTI, 2013) with the georeferenced maps to construct village-level indicators of primary tribe. Our analysis focuses on two tribal confederations: the Ghilzai and the Durrani.[11]

The Durrani confederation is the confederation of then-president Hamid Karzai and historically, members from this confederation form the bulk of government jobs and the bureaucracy. Durrani areas therefore correspond to low n areas in the theoretical model. We also present results for

[10]Refer to the Culture and Conflicts Studies website (available at: `https://my.nps.edu/web/ccs/afghanistan1` for more information on the tribal maps.

[11]Confederations are typically formed by groups of tribes with common origin or historical alliances.

Popalzai areas. The Popalzai are the tribe within the Durrani confederation of then-president Hamid Karzai. We compare these results with the analysis focusing on villages of the Ghilzai confederation. The Ghilzai confederation are historic geopolitical rivals of the Durrani. Most importantly, most of the historical Taliban leadership, including founder Mullah Omar, come from the Ghilzai confederation. With this in mind, Ghilzai areas correspond to high n areas in the theoretical model. Note in Appendix Table A3 that the Durrani (37% of all Pashtun villages) and the Ghilzai (38% of all Pashtun villages) are the two largest Pashtun confederations. Refer to Appendix Figure A4 for a map of these confederations.

Table 6 presents the RD estimates using whether there has been any attack within a 2-km radius of a village's location as the outcome variable. Panel A presents results for potentially high n, Ghilzai villages. Note that, for all attack types, coverage has no meaningful impact on the likelihood of an attack. The RD coefficients are both small in magnitude and imprecise. Panel B presents a contrasting picture. For low n Durrani villages; we find large and statistically significant drops in all attack types. For instance, the likelihood of a direct fire attack or an IED explosion declines by 12 and 23 percentage points, respectively, in villages with coverage. The magnitudes of the coefficients are even higher when restricting the analysis to Popalzai villages (Panel C), although we note that the sample size is much smaller and thus the results are less precise.

The findings in this section suggest that the deterring effect of coverage is trivial in the presence of norms favoring insurgents. In other words, even when cell phone coverage is available, citizen monitoring levels are significantly low when sympathy for insurgents is high. Alternatively, coverage may be increasing citizen engagement in monitoring, however, this is only offsetting whatever beneficial effect from coverage insurgents are simultaneously getting (e.g., coordination of attacks, IED remote detonation, etc.). The next section explores these channels in more detail.

6 Coverage-Violence Channels

Coverage can affect violence through two primary channels: (1) it reduces the cost of producing violence (e.g., efficient coordination of attacks, remote detonation of IEDs), and (2) it facilitates intelligence gathering by state actors engaging in counterinsurgency. The intelligence gathering channel is primarily driven by two sub-channels: the transfer of insurgent-related information from citizens to the government, and the collection of signals intelligence by the government (e.g., tracking insurgents' movements or intercepting their communications). Broadly speaking, we consider the former to be a bottom-up approach to intelligence gathering while the latter is a top-down approach.

This section starts by exploring the relative importance of the violence production channel and the intelligence gathering channel. However, we note that the drop in violence documented in section 5.2 provides considerable evidence that the intelligence channel more than offsets any gains to insurgents resulting from lower costs of producing violence. Thus, the remaining of the section primarily focuses on disentangling the relative importance of the top-down and bottom-up sub-channels in reducing violence.

6.1 Violence Production Costs and Intelligence Gathering

Although we do not observe coordination because of coverage, we try to assess it by testing whether coverage makes insurgents more successful/more efficient. To do this, we use the failure rate of IEDs as performance measure of whether good coordination and also whether coverage makes IED more efficient. Failure rate refers to the rate of IEDs that result in explosion out of total IEDs (including those that were found and cleared).

Table 7 Column (1) tests if more efficient. Evidence suggest that no. However, information sharing can also negatively impact the effectiveness. Columns 2 and 3 explore this. Evidence that coverage increases the failure rate if close to a village. This suggests that information sharing works. Therefore we conclude that there is no evidence that coverage improves the effectiveness , and if it does improve effectiveness somehow, the benefit is entirely offset by the effect of information sharing on the

The findings in this section suggest that the deterring effect of coverage is trivial in the presence of norms favoring insurgents. In other words, even when cell phone coverage is available, citizen monitoring levels are significantly low when sympathy for insurgents is high. Alternatively, coverage may be increasing citizen engagement in monitoring, however, this is only offsetting whatever beneficial effect from coverage insurgents are simultaneously getting (e.g., coordination of attacks, IED remote detonation, etc.).

6.2 Intelligence Gathering: Top-down *versus* Bottom-up Information

The intelligence gathering channel is primarily driven by two sub-channels: the transfer of insurgent-related information from citizens to the government, and the collection of signals intelligence by the government (e.g., tracking insurgents' movements or intercepting their communications). Broadly speaking, we consider the former to be a bottom-up approach to intelligence gathering while the latter is a top-down approach. For convenience, throughout this paper we refer to the former as human intelligence (HUMINT) and to the latter as signals intelligence (SIGINT). we cannot observe the sharing of information, however, we can deduce/provide evidence

Appendix Figure A5

Table 8

Table 9

Figure 4

References

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.

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

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.

Druckenmiller, H. and S. Hsiang (2018). Accounting for unobservable heterogeneity in cross section using spatial first differences. Technical report, National Bureau of Economic Research.

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

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

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*.

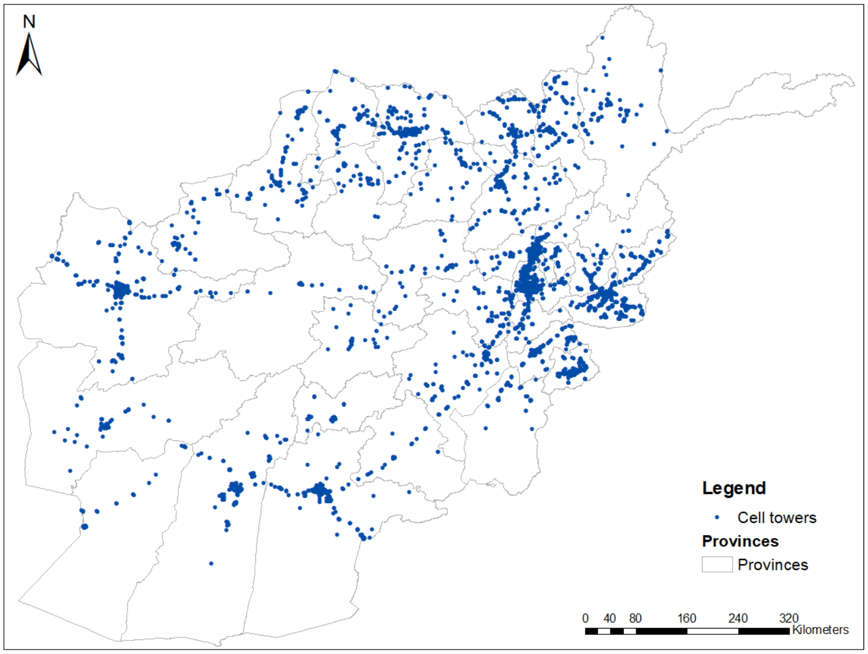
Open Topography (2017). Alos world 3d - 30m.

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

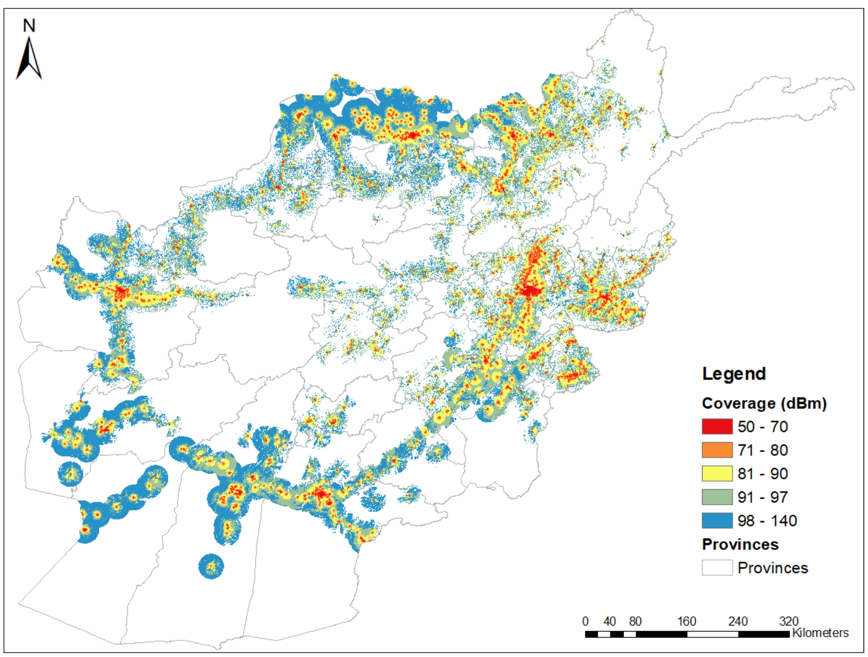
Shapiro, J. N. and D. A. Siegel (2015). Coordination and security: How mobile communications affect insurgency. *Journal of Peace Research 52*(3), 312–322.

Tribal Hierarchy and Dictionary of Afghanistan (2007). *Tribal hierarchy and dictionary of Afghanistan: a reference aid for analysts*. Courage Services Inc.

Figures and Tables



(a) Cell Towers, Afghanistan (2012)



(b) ITM Model Output

Figure 1: 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 B. Lower dBm values mean higher signal strength (i.e., more coverage).

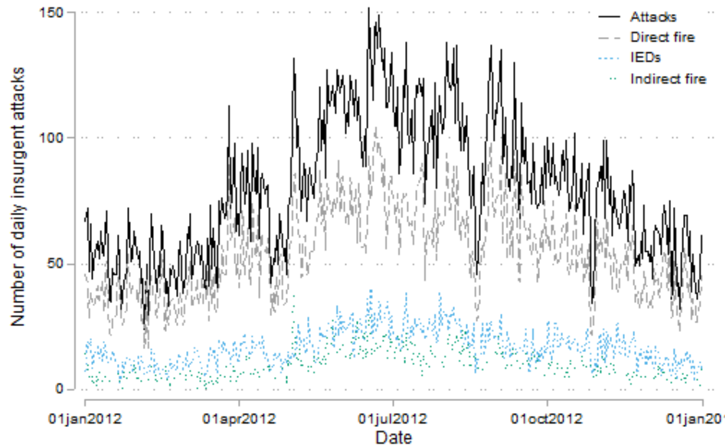
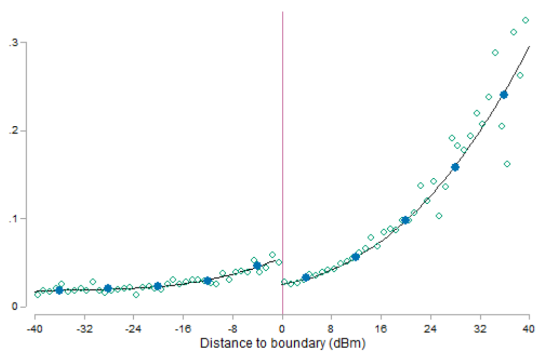
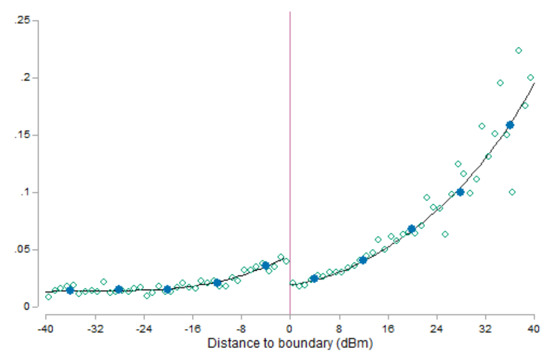


Figure 2: Insurgent Violence, 2012

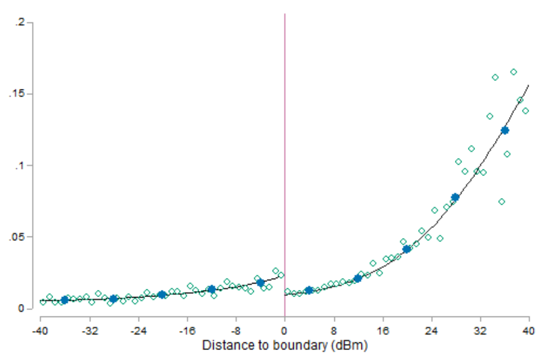
Notes:



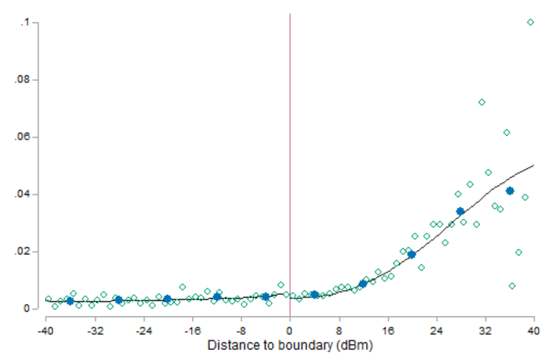
(a) Any attack



(b) Direct fire



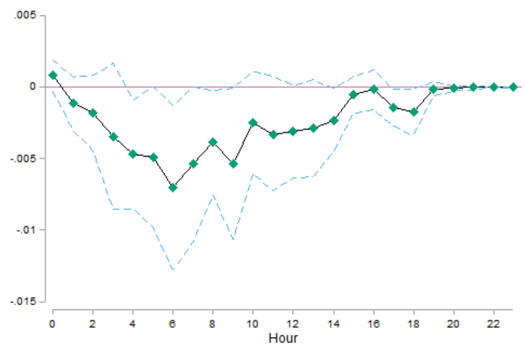
(c) IED



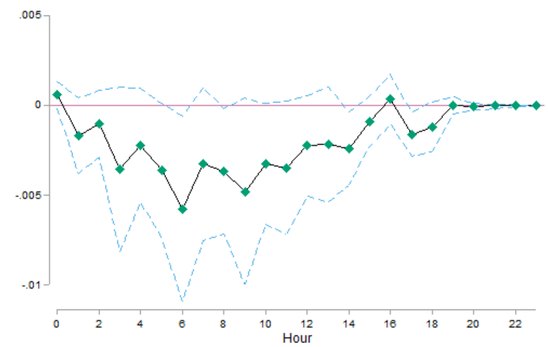
(d) Indirect fire

Figure 3: Regression Discontinuity Plots

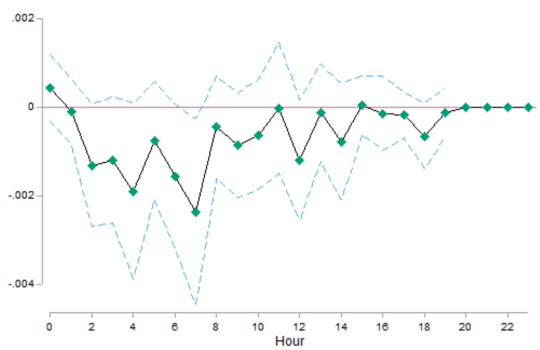
Notes:



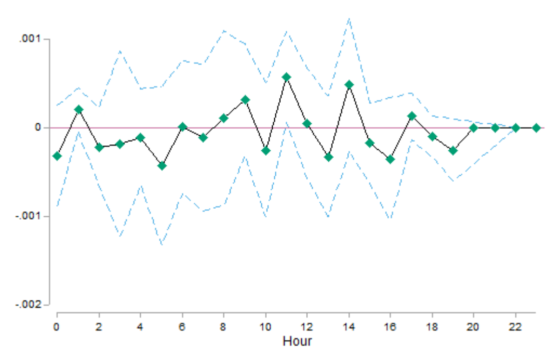
(a) Any attack



(b) Direct fire



(c) IED



(d) Indirect fire

Figure 4: Effect of Coverage by Time of Day

Notes:

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:

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.

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.

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.

Table 5: Effect of Coverage on Insurgent Violence, Spatial First Differences

	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: Latitude as Panel variable				
Coverage	-0.007***	-0.006***	-0.006***	-0.003**
	(0.003)	(0.002)	(0.002)	(0.001)
Observations	157,769	157,769	157,769	157,769
Number of districts	371	371	371	371
Mean Outside coverage	0.0083	0.0058	0.0029	0.0012
Panel B: Longitude as Panel variable				
Coverage	-0.009***	-0.006***	-0.008***	-0.001
	(0.003)	(0.002)	(0.002)	(0.001)
Observations	158,060	158,060	158,060	158,060
Number of districts	371	371	371	371
Mean Outside coverage	0.0083	0.0058	0.0029	0.0012

Notes: Results present estimates of β using a Spatial First Differences design (Druckenmiller and Hsiang, 2018). Refer to section 5.2.1 for more detail on the design. All specifications use standard errors clustered at the district level. *, **, and *** indicate 10, 5, and 1 percent significance respectively.

Table 6: Effect of Coverage on Insurgent Violence, By Tribal Affiliation

	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: Effect within Guilzai villages (High n)				
Coverage	0.012	-0.097	-0.012	0.096
	(0.084)	(0.085)	(0.078)	(0.066)
Observations	922	922	922	824
Mean Outside coverage	0.58	0.51	0.31	0.15
Bandwidth (dBm)	8.31	8.07	8.16	7.16
Panel B: Effect within Durrani villages (Low n)				
Coverage	-0.144**	-0.118*	-0.226***	-0.128*
	(0.065)	(0.066)	(0.070)	(0.073)
Observations	935	935	935	839
Mean Outside coverage	0.68	0.63	0.53	0.33
Bandwidth (dBm)	8.31	8.07	8.16	7.16
Panel C: Effect within Popalzai villages (Low n)				
Coverage	-0.294**	-0.278**	-0.204*	-0.011
	(0.127)	(0.115)	(0.117)	(0.039)
Observations	182	182	182	182
Mean Outside coverage	0.31	0.23	0.23	0

Notes: Outcome variables are an indicator for whether a village has experienced at least one attack within 2 kilometers of village's location. Classification of tribal confederation done using the Tribal Hierarchy and Dictionary of Afghanistan (2007). Refer to Section 5.3 for more details on the classification and the choice of tribes analyzed. All specifications include controls for elevation and slope. Standard errors clustered at village level given low number of districts within each subsample. Panel C does not include any bandwidth restrictions given low sample size. ***, **, * indicate 10, 5, and 1 percent significance, respectively.

Table 7: Effect of Coverage on Unsuccessful IEDs

	Dependent variable= IED Failure Rate			
	RD (1)	$RD_{\leq 2km}$ (2)	$RD_{>2km}$ (3)	DD-RD (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.

Table 8: Effect of Coverage on Likelihood of Attack near Villages and Roads

	Dep. Variable = $\mathbb{1}\{$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 9: Effect of Coverage by Time of Day

	Dependent variable: Any Attack=1		
	Daylight (7AM-5PM)	Astronomical twilight	Night (8PM-4AM)
	(1)	(2)	(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.

Appendix A Additional Figures and Tables

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

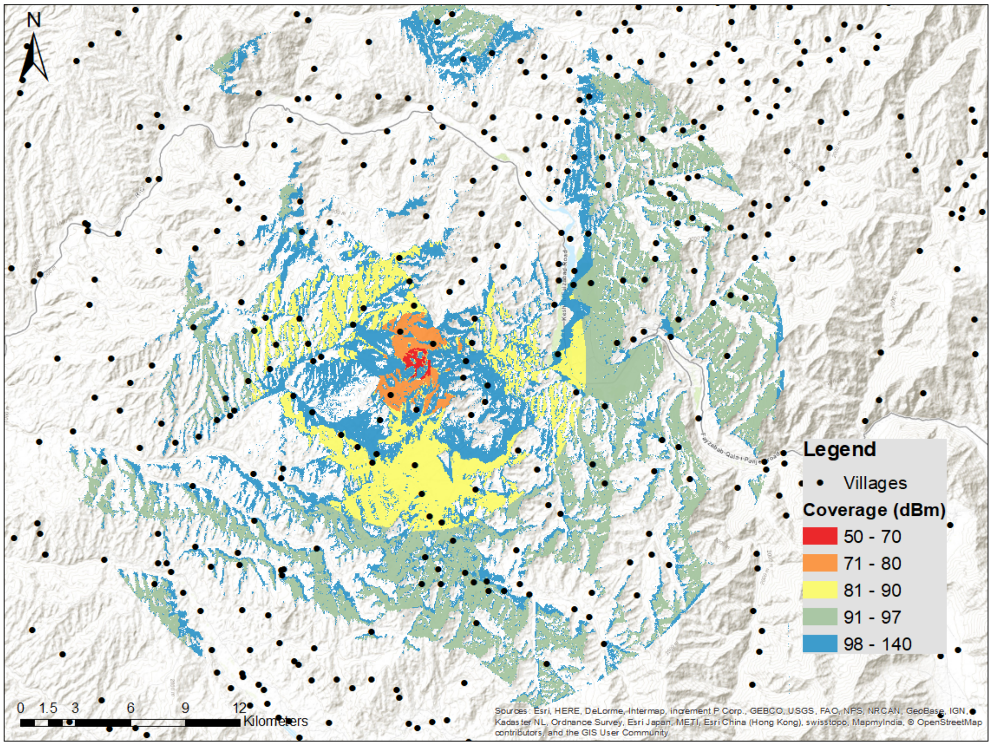


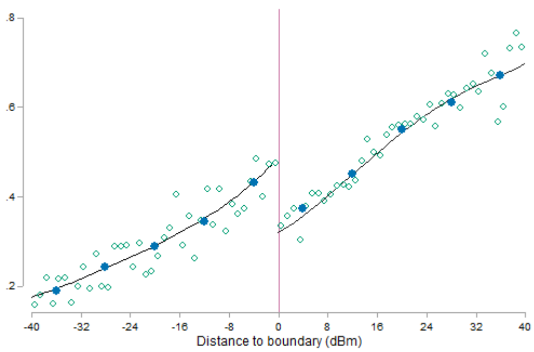
Figure A2: Sample ITM for a single tower

Notes:

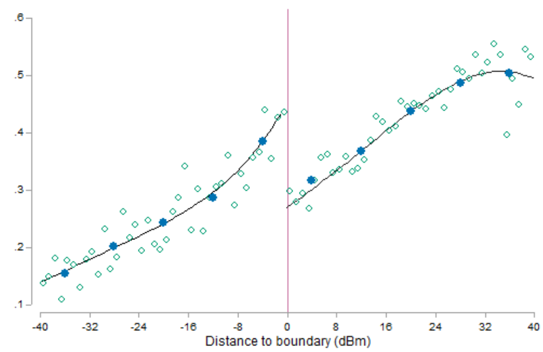
Table A1: Effect of Coverage on Hoax IEDs

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

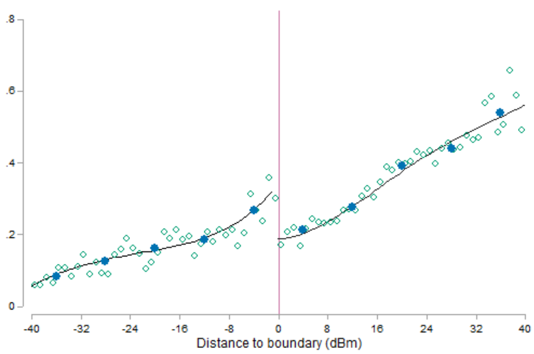
Notes:



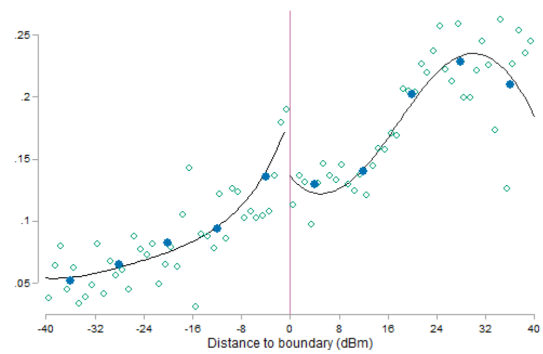
(a) Any attack



(b) Direct fire



(c) IED



(d) Indirect fire

Figure A3: Regression Discontinuity Plots

Notes:

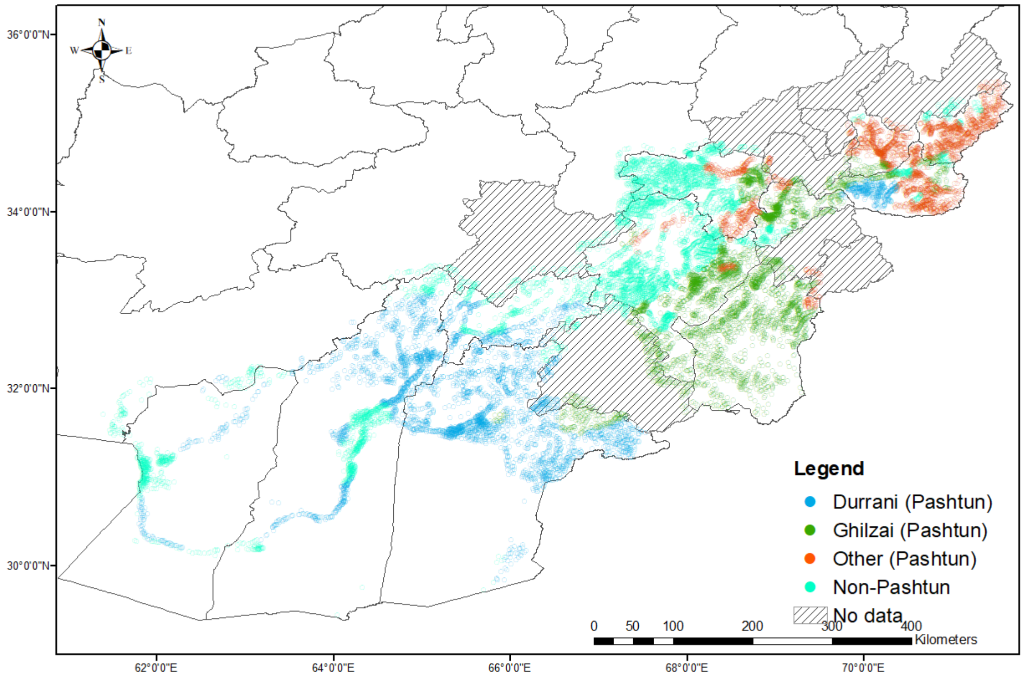


Figure A4: Pashtun Tribes of Afghanistan

Notes:

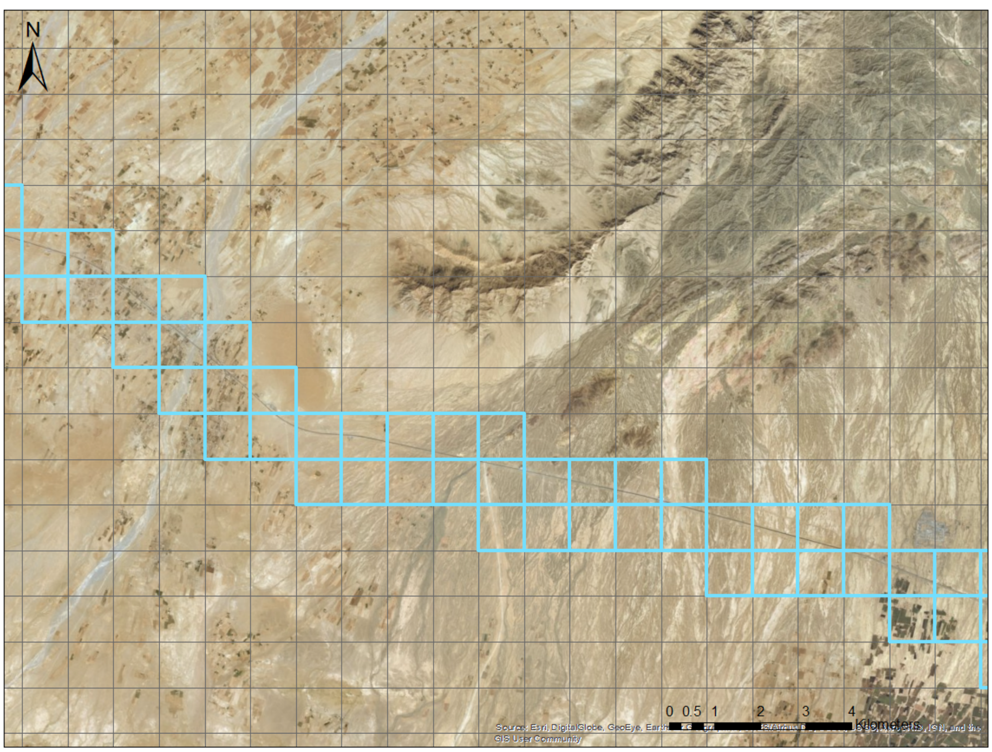
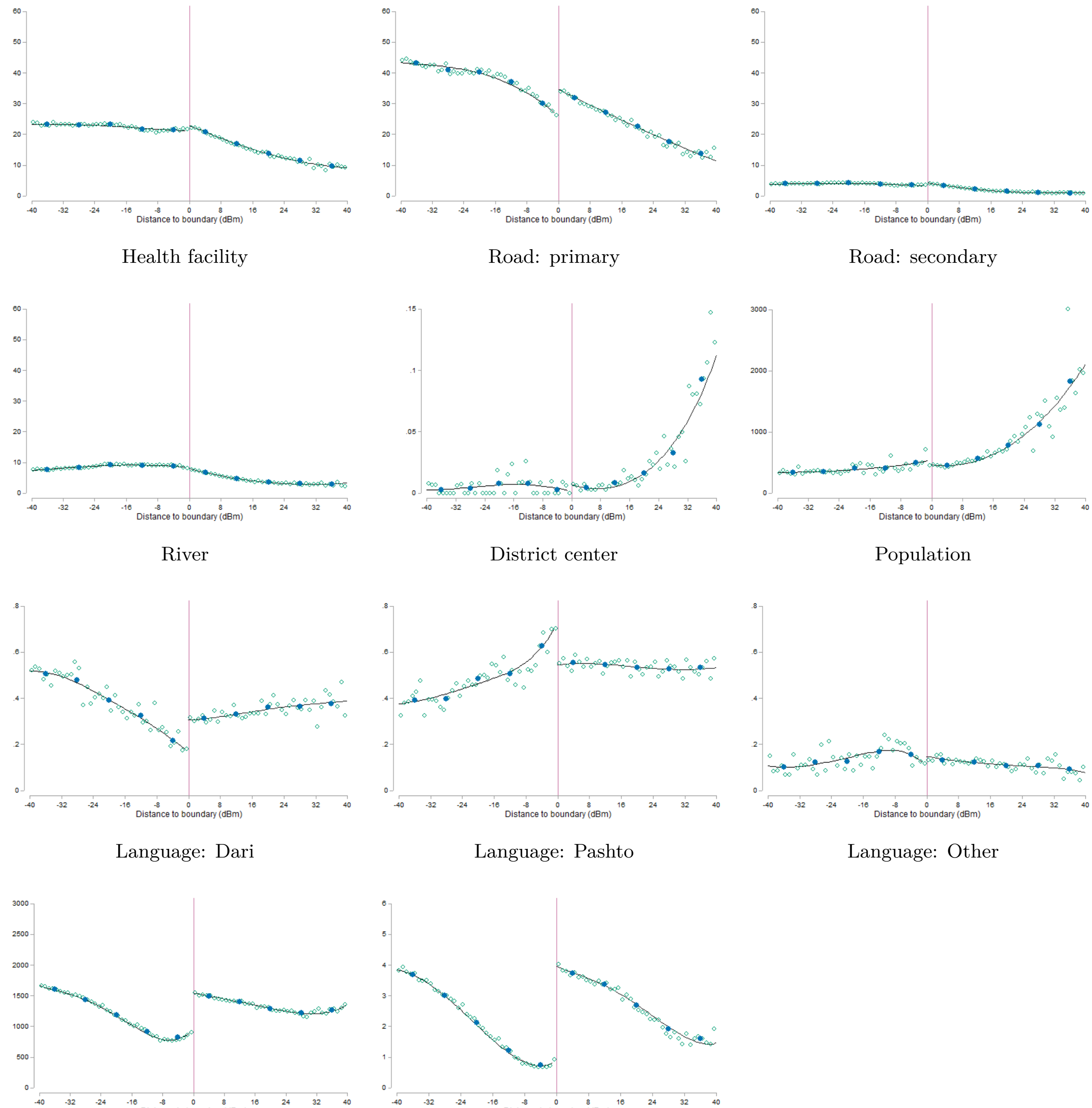


Figure A5: Sample Cell Selection

Notes:



Health facility | Road: primary | Road: secondary

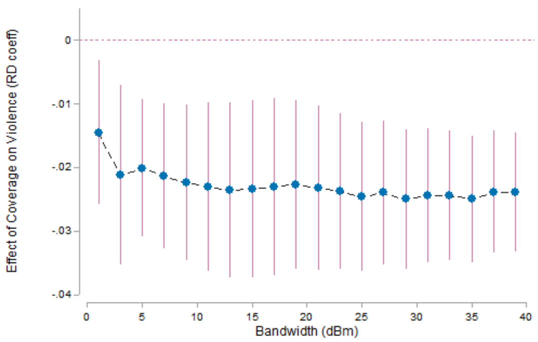
River | District center | Population

Language: Dari | Language: Pashto | Language: Other

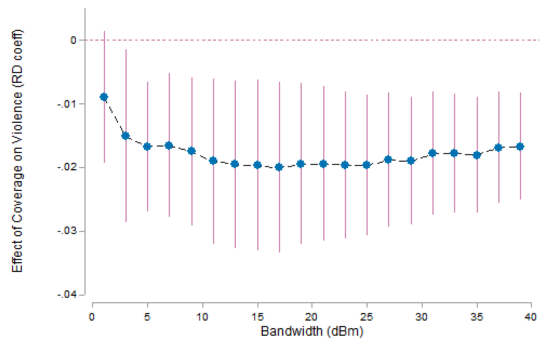
Elevation | Slope

Figure A6: Regression Discontinuity Plots, Covariate

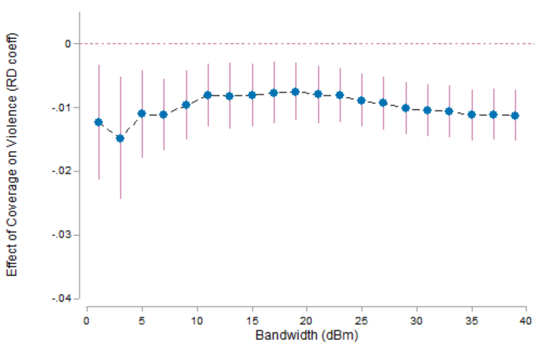
Notes:



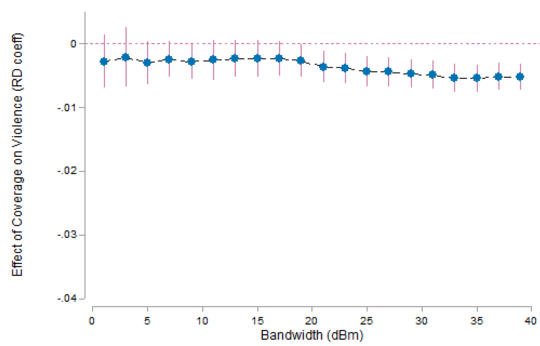
(a) Any attack



(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:

Appendix B Estimation of the Irregular Terrain Model

Other model inputs:

ITM reliability rel 99 Output type out 2 Distance unit dis m Knife-Edge difraction ked 0 Antenna code 39

Estimation of the ITM model uses the SLEIPNIR engine in cloudRF.

Table B4: Main variables and Parameters of Irregular Terrain Model (ITM)

Model variables	Description	Parameters
Transmitter characteristics		
Transmitter power	Transmission power (Watts)	5
Frequency	Radio wave frequency (MHz)	900
Latitude	Latitude of cell tower	e.g., 37.158
Longitude	Longitude of cell tower	e.g., 70.765
Transmitter height	Height of cell tower above ground (meters)	30
Radius	Maximum coverage radius (kilometers)	20
Antenna gain	Transmitter antenna gain (dBi)	2.14
Receiver characteristics		
Receiver sensitivity	Minimum power received threshold (dBm)	-140
Receiver height	Receiver height above ground (meters)	1.5
Antenna gain	Receiver antenna gain (dBi)	2.14
Geographic characteristics		
Resolution	Topographic model	DSM30
Radio climate	1: Equatorial	e.g., 5
	2: Continental Subtropical	
	3: Maritime Subtropical	
	4: Desert	
	5: Continental Temperate	
	6: Maritime Temperate, over land	
	7: Maritime Temperate, over sea	
Terrain conductivity	Salt water : 80	e.g., 13
	Fresh water : 80	
	Good ground : 25	
	Marshy land : 12	
	Farmland, forest : 15	
	Average ground : 15	
	Mountain, sand : 13	
	City : 5	
	Poor ground : 4	

Notes: