Journal of Development Economics 167 (2024) 103228

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

Is the phone mightier than the virus? Cellphone access and epidemic containment efforts[☆]

Robert Gonzalez [a], Elisa M. Maffioli [b,*]

[a] School of Economics, Georgia Institute of Technology, Atlanta, GA 30313, United States of America
[b] Health Management and Policy, University of Michigan School of Public Health, Ann Arbor, MI 48109, United States of America

ARTICLE INFO

JEL classification:
I15
I18
O22

Keywords:
Ebola virus disease
Cellphones
Technology
Information
Care

ABSTRACT

This paper examines the impact of cellphone access on the containment of an epidemic. We study this question in the context of the 2014 Ebola Virus Disease (EVD) outbreak in Liberia. Combining novel data on cellphone towers and EVD cases, we estimate a high-resolution radio-wave propagation model that uses variations in terrain topography and the spatial distribution of cellphone towers to predict signal strength on the ground. We then employ a regression discontinuity design that compares villages at the margin of the signal strength threshold required for coverage. We find that having access to cellphone coverage leads to a 10.8 percentage point reduction in the likelihood that a village has an EVD case. Results from a novel survey collected following the epidemic suggest that this may be explained by cellphone access facilitating treatment provision. However, we cannot fully exclude that improving access to preventive care or information could have also contributed to containment.

1. Introduction

Infectious disease outbreaks are a significant burden to low and middle-income countries (Holmes et al., 2017). Therefore, assessing the effectiveness of tools that can prevent or contain these outbreaks is a first-order policy issue. Given their widespread availability, cellphones have the potential to be one such tool. A growing literature shows that cellphone technology can be used to improve the delivery of health care (Braun et al., 2013; Agarwal et al., 2015; Obasola et al., 2015), to predict the spread of infectious diseases by studying mobility patterns (Bengtsson et al., 2011, 2015; Wesolowski et al., 2015; Milusheva, 2020), to help diagnose disease (D'Ambrosio et al., 2015), and as a tool for information sharing, reporting, and surveillance (Yang et al., 2009; Freifeld et al., 2010; Sacks et al., 2015). While this literature offers guidance on the design and use of specific tools that can be deployed during outbreaks, the broader question of whether access to cellphone technology itself has an impact on the spread (or containment) of disease during emergency situations remains largely unexplored.

Ex-ante, the impact of cellphone access on the spread of an infectious disease is ambiguous. Cellphone technology can facilitate access to preventive care (*preventive care* channel) and treatment resources (*treatment care* channel). It also enables individuals to more efficiently interact with a potentially larger network of friends and family and to improve within-network collective action during emergencies.[1] However, improved coordination may result in an increase in person-to-person interactions within the network (e.g., caring for sick members, attending funerals), potentially increasing the likelihood of spread along the network (*network* channel). Finally, cellphone access can increase exposure to outbreak-related information while, simultaneously, providing a platform for the spread of misinformation (*information* channel).

This paper explores the causal effect of cellphone access–proxied by cellphone coverage–on the spread (or containment) of an infectious disease, namely the Ebola Virus Disease (EVD, hereafter) in the context of the 2014 West Africa epidemic in Liberia. We investigate whether cellphone coverage affects the likelihood that a village reports an EVD case, by employing several novel sources of data. First, we use data on EVD cases compiled by the authors from primary records obtained from the Liberia's Ministry of Health (MOH). This dataset encompasses the entire set of villages in the country for the whole

☆ We thank participants at the 2019 Econometric Society Winter Meeting, 2020 Barcelona GSE Summer Forum, PacDev 2020, the University of Michigan, UC-Santa Cruz, University of Notre Dame, University of Chicago, Chan School of Public Health at Harvard University, Mailman School of Public Health at Columbia University, European Economic Association meeting, Milan, Italy; The Econometric Society Africa Meeting and North American Meeting, the Federal Communications Commission (FCC), Georgia Tech Global Development Week, and University of Southern Denmark for invaluable comments and suggestions.
* Corresponding author.
E-mail addresses: robert.gonzalez@gatech.edu (R. Gonzalez), elisamaf@umich.edu (E.M. Maffioli).
1 See, for example, Hampton et al. (2011), Pew Research Center (2011, 2019), Blumenstock et al. (2016).

https://doi.org/10.1016/j.jdeveco.2023.103228
Received 26 May 2023; Received in revised form 25 November 2023; Accepted 27 November 2023
Available online 14 December 2023

duration of the epidemic. Second, we gather data on the location and characteristics of cellphone towers across Liberia in the year 2013–just prior to the outbreak–obtained from the Liberia Telecommunications Authority (LTA). Third, we explore mechanisms using original survey data conducted six months after the end of the epidemic on about 2000 respondents across Liberia, along with more than 233 million anonymized call detail records (CDR) obtained from one of the major mobile network operators in the country.

Research studying the impact of cellphone coverage on any outcome of interest faces two hurdles: how to accurately measure coverage, and how to address the endogenous selection of locations and people into coverage. This paper addresses these issues in two ways.

First, we measure cellphone coverage by estimating a high spatial resolution radio-wave propagation model—the Irregular Terrain Model (ITM). This model is widely used by regulators and mobile network operators for its accuracy (Crabtree and Kern, 2018). It combines cell tower footprint, terrain topography, and other factors to provide a measure of signal strength at each point on the ground. To the best of our knowledge, this is the most accurate measure of cellphone coverage available in the literature. While previous research has used radio wave models to predict radio and TV signal strength (Olken, 2009; Enikolopov et al., 2011; Durante et al., 2019; Walsh, 2022), the cellphone literature has lagged. Existing approaches use the presence of a tower in a given location (Jensen, 2007; Aker, 2010) or a fixed radius around a tower (Shapiro and Weidmann, 2015), but these ignore the effects of topography on signal diffusion. More recent approaches use GSMA coverage maps that implicitly use radio wave models (Guriev et al., 2020; Manacorda and Tesei, 2020; Gonzalez, 2021). However, the reliability of these maps is questionable.[2]

Second, we employ a regression discontinuity (RD) design that uses the signal strength obtained from the propagation model as the forcing variable and the minimum signal needed for coverage as the cutoff. Our RD design differs from related work in that we use signal strength as the forcing variable rather than geographic distance to the coverage boundary (Gonzalez, 2021). This allows us to better account for selection into coverage by limiting our analysis to villages within a margin of the signal strength cutoff. Within this margin, whether a village receives *just* enough signal strength is determined by minor variations in topography (e.g., small hill, rock outcrop, vegetation clutter, etc.) that lead to arbitrary diffraction and blocking of the signal. We confirm this empirically by exploring a rich set of ex-ante village characteristics to predict coverage: topographic characteristics–not demographic or socioeconomic characteristics–are the sole predictors of coverage within a close margin of the coverage cutoff. Further analysis of these village characteristics also suggests a smooth transition across the cellphone coverage cutoff, and thus little indication that the likelihood of an EVD case is explained by these characteristics jumping at the cutoff.

We find evidence that cellphone coverage helped contain the spread of the disease. Our RD estimates show a 10.8 percentage point reduction in the likelihood that a village with just enough coverage reports an EVD case relative to villages that are just under the cutoff. Results are robust when exploring measures of disease at the intensive margin such as the number of cases. Additional results using a panel-RD specification that exploits monthly variation in EVD incidence, provide significant evidence of containment effects. We find that the likelihood that a village reports an EVD case given past EVD incidence within the district, is reduced by 1.9 percentage points if the village has cellphone coverage. Villages without coverage are not as shielded, reporting instead a 1.6 percentage point increase in the likelihood of EVD given past exposure to EVD within the district. This suggests that coverage potentially overcomes the contagion effect.

In a similar panel-RD specification, we explore how past exposure to EVD *within* a given village affects the likelihood of an additional case. We find consistent evidence that cellphone coverage reduces the contagion effect within a village. Finally, we complement our panel-RD findings with results from a Cox proportional hazards model that estimates the hazard ratio as a function of cellphone coverage (and geographic controls). We show that having cellphone coverage decreases the hazard rate of having a first EVD case as well as an additional case, after experiencing the first one. The survival curves show that, throughout the study period, the survival probabilities at any given month (i.e., experiencing no first or no additional cases) remain statistically significantly higher in villages with cellphone coverage than in villages without coverage. The results are similar when exploring suspected EVD deaths.

Our main findings are robust to a battery of robustness and falsification checks. First, our results do not change after controlling for "free-space" signal strength (signal strength in the absence of topography) (Olken, 2009). By comparing locations that would have received the same signal strength in the absence of topography, this exercise ensures that the identification relies solely on variation due to minor topographic characteristics. In addition, our findings hold when: (i) using alternative measures of the outcome (likelihood and number of suspected deaths, and number of months affected by EVD), (ii) using coverage one year after the epidemic as a falsification test, (iii) using both geographic distance and signal strength distance as forcing variables, (iv) accounting for potential non-compliance in access to cellphone coverage ("walking into coverage"), (v) assessing the sensitivity of our RD estimates to near-cutoff observations, and (vi) estimating spatial standard errors using 10, 20, and 50 kilometer distance cutoffs (Conley, 1999).

We explore several channels underlying the relationship between cellphone coverage and EVD. We explore the *preventive* and *treatment care* channels using survey data. However, since the data were collected using a mobile phone survey, we cannot estimate the effect of coverage at the extensive margin since, potentially, all respondents have access to coverage. Instead, we use signal strength (i.e., intensive margin) to estimate how the quality of coverage affects several mechanisms. Specifically, we test whether villages with higher signal strength are more likely to receive preventive care (e.g., whether health workers, officials, and community task-forces came to their village to explain EVD, hold hygiene meetings, bring information, teach safe burial procedures, or bring preventive materials). We find no statistically significant evidence on this channel. Second, we test whether stronger coverage is associated with improved treatment care during the epidemic. We find that villages with higher signal strength are more likely to report that ambulances arrived on time and that care centers were placed near their villages. Putting together the preventive and treatment care outcomes into summary index measures (Kling et al., 2007), we report a statistically significant effect for treatment care, but not for preventive care. Overall, these findings as well as the of additional results from the panel-RD within a village and the Cox proportional hazards model point to treatment care playing a significant role in explaining how coverage helped contain the disease. However, we cannot fully exclude that prevention may have also played a role.[3]

We test the *network* channel using two alternative measures of a village's network. First, we use original data on more than 233 million anonymous CDR for the universe of mobile subscribers of one of the

[2] GSMA maps are self-reported by the mobile network operators; they do not include all operators in a country; they are not uniformly reported across countries (some countries use radio wave models while others report tower locations); the majority of maps use outdated data for certain providers; maps use a single signal strength cutoff to assign coverage (−100 dBm), however this cutoff varies across different technologies (2G versus 3G, frequency of transmission, etc.) which vary significantly across countries and years.

[3] Our main specification relies on an extensive measure of disease, while we explore measures at the intensive margin as robustness because of the potential measurement error in the number of cases.

largest mobile network operators in Liberia. Information on date and time of calls, and the location of originating and receiving cell towers allow us to create tower-to-tower measures of connectedness based on day-to-day call behavior. We define networks of villages based on whether they are within the service areas of connected towers. Second, we classify all villages across Liberia based on clans using the latest available pre-outbreak (2008) census data. Clans are groups of villages that, although currently considered administrative units, correspond to historical tribal chiefdoms that were gradually fused into the state (Nyei, 2014). We then test whether the likelihood of an EVD case in villages with cellphone access increases if an affected village within the network also has coverage. We find that a "coverage match" between villages of the same network does not lead to any meaningful change in the likelihood of EVD spread.

Finally, we test whether respondents in villages with higher signal strength are more likely to receive EVD-related information and whether they are more likely to be (mis)informed about the origin of the epidemic. We do not find strong evidence that coverage quality has an effect on the likelihood of being (mis)informed. We attribute this finding to both the widespread availability of radio across Liberia and the lack of Internet access at the time of the epidemic.

This paper fits into the economics literature that investigates the economic impact of cellphones and other information and communication technologies (ICTs) in developing economies (Aker and Mbiti, 2010). These studies explore the effects on price dispersion (Jensen, 2007; Aker, 2010; Aker and Fafchamps, 2014), education (Aker et al., 2012; Aker and Ksoll, 2018), agriculture (Fabregas et al., 2019; Cole and Fernando, 2021), and the role of mobile money in financial transactions (Jack and Suri, 2011, 2014; Blumenstock et al., 2016). To the best of our knowledge, this is one of the first papers in the economics literature that looks at the effects of cellphone coverage on health outcomes.[4]

Our paper contributes to past research in public health exploring cellphones and a number of health-related outcomes: the management of health records (Agarwal et al., 2015), maternal and child health indicators (Obasola et al., 2015), the remote diagnosing of diseases (D'Ambrosio et al., 2015), the quality and efficiency of care and services (Braun et al., 2013). We advance this literature in three areas. First, we present an improved way of measuring cellphone access at a large, country-wide scale. Previous studies typically focus on limited settings such as access to study-specific phone tools or applications. Second, our empirical design addresses the issue of selection into cellphone technology. Third, while this literature focuses on health-related outcomes in regular, day-to-day settings, our paper explores whether the technology is effective in the midst of a health crisis–a sudden-onset epidemic–in a setting characterized by general mistrust towards local and international institutions. Our paper shows that cellphones can be effective even in such settings.

This paper also contributes directly to the strand of the literature related to cellphone technology and infectious diseases. Studies within this area focus on cellphone technology as a tool to predict future outbreaks (e.g., using mobility patterns estimated from phone usage to predict the spread of disease (Lu et al., 2012; Bengtsson et al., 2015; Wesolowski et al., 2015)), or evaluating phones as a "participatory epidemiology" tool (e.g. for information sharing, reporting, and tracking of cases within communities (Yang et al., 2009; Freifeld et al., 2010; Sacks et al., 2015; Feng et al., 2018)).[5] While this literature evaluates specific tools that can be deployed during outbreaks, our paper answers a broader question: can access to cellphones among the general population have an impact on spreading (or containing) outbreaks? We also go beyond the evaluation exercise by exploring mechanisms that can explain the relationship between cellphone access and epidemic containment. Our findings show how something as simple and ubiquitous as a cellphone can have positive impacts.

2. Background on the 2014 EVD outbreak in Liberia

The first case of EVD in West Africa occurred in Guinea near the border with Sierra Leone and Liberia in December 2013, but EVD was not confirmed in Liberia until March 2014. After the first case was recorded in the country on March 20, 2014, the Government of Liberia (GOL) responded to the epidemic with social mobilization, case management, treatment and surveillance, water sanitation, and hygiene activities. The MOH led relief efforts supported by several international institutions such as the World Health Organization (WHO), Medicines sans Frontiers, and Samaritan's Purse. The first wave of EVD was quickly contained and by April 9, the last EVD case for almost two months was confirmed.

However, on May 25, 2014, a second wave of the epidemic started when a new EVD case was recorded in Lofa county near the border with Guinea. By the end of June, the disease had spread to the capital city, Monrovia. By August 2014 the situation was out of control. The GOL urgently called on the international community for a massive response, by declaring a State of Emergency on August 6. Schools and Liberia's land borders were closed, and strict control measures, including quarantines of neighborhoods and a nightly nationwide curfew, were imposed. On August 8 the WHO declared the Ebola outbreak a "Public Health Emergency of International Concern", the highest level of international alert. By the end of that month, there was a growing awareness of the need for more decentralized control and involvement of local communities: the GOL created county-level taskforces to strengthen local coordination in the fight against EVD, and an Incident Management System (IMS) devoted exclusively to the national management of the epidemic (Nyenswah et al., 2016, Hymowitz, 2017).

As the number of EVD cases continued to rise, international funding started being poured into Liberia. The United States government committed US$319 million for the response in West Africa. Other institutions, such as the World Bank, approved an additional US$105 million, with US$52 million specifically for Liberia (World Bank, 2014). Overall, about 62 countries committed US$2.3 billion to respond to the epidemic in West Africa, including US$806 million to Liberia (White House, 2014). Over time, the GOL was able to open Community Care Centers (CCCs), Ebola Treatment Units (ETUs), and coordinate safe burials and the removal of dead bodies from communities, through teams of governmental health workers. Following an assessment of the major areas of intervention during the EVD outbreak, Kirsch et al. (2017) concluded that no single intervention stopped the epidemic, rather all interventions likely had reinforcing effects. In fact, the epidemic's turning point–September 2014–coincided with a reorganization of the response, the emergence of community leadership in control efforts, and changing beliefs and practices within the population. While in the following months the epidemic was rapidly slowing down, the GOL efforts kept securing additional funding, constructing the planned ETUs and coordinating the activities of the international partners involved. By early 2015, 31 ETUs were constructed and more than 70 CCCs opened.

In January 2015 there were fewer than 15 weekly confirmed cases with the last EVD case being reported in mid-March 2015. The country was initially declared EVD-free on May 9, 2015. However, a small number of cases reported in July and December of the same year led to the official EVD-free declaration to be postponed to January 14, 2016. Along with Sierra Leone and Guinea, Liberia was among the most affected countries by EVD in West Africa. In Liberia, 10,675 confirmed, probable, or suspected cases were recorded, while the cumulative number of deaths reached 4809—the highest number in West Africa (World Health Organization, 2016).

[4] A contemporaneous paper by Mensah et al. 2022 uses plausibly exogenous variations in lightning intensity and (sub)regional convergence in mobile penetration as instrumental variables for mobile network expansion, and it finds that cellphone coverage is associated reduction in infant mortality.

[5] Cellphones were also used during the 2014 West Africa Ebola epidemic to collect and share data, to create and share digital maps of the diseases, to track contacts and the spread of the disease within a community (Sacks et al., 2015), and to track health seeking behavior (Feng et al., 2018). A more recent literature focuses on the role of cellphones during COVID19 (Grantz et al., 2020; Oliver et al., 2020).

3. Data

3.1. Ebola data

The data on EVD cases are primarily constructed from the MOH patient database containing more than 19,000 patients tested for EVD from March 2014 to July 2015. The data are considered to be the most comprehensive database to date, since every organization taking part in the response to the outbreak was required to report cases to the MOH (Liberian Ministry of Health, 2017). Furthermore, we supplement these with a database from Global Community, a development organization that managed all the burials after July 2014. Since the database records the village where the person resided when suspected to have contracted EVD, we were able to manually code and match the data with the entire list of 9686 villages in the 15 counties of Liberia.

For each village, we construct the main outcome of interest as an indicator equal to 1 if at least one (probable, confirmed, or death) case was recorded in the village during the study period (January, 2014–July, 2015). We primarily rely on this measure of EVD at the extensive margin because the MOH database might underestimate the total number of cases as it records only patients tested for the disease.[6] Despite the potential measurement error, we also investigate the number of cases as an outcome in a robustness check. As alternative outcomes, we also explore the total number of months a village was affected by EVD and whether a village recorded a suspected EVD death.

3.2. Cellphone coverage data and measures

This paper uses a high resolution radio wave propagation model–the Irregular Terrain Model (ITM)–to determine coverage strength across Liberia. Regulatory agencies and businesses around the world rely on the ITM as their workhorse model to model coverage and signal propagation due to its high accuracy, its ability to capture terrain topography, and being repeatedly validated via on-the-ground measurements (Longley and Rice, 1968; Eppink and Kuebler, 1994; Seybold, 2005; Lazaridis et al., 2013).[7]

The ITM provides a measure of signal strength at a given point on the ground, taking as inputs three sets of information: (1) the characteristics of the transmitter or cell tower (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, terrain conductivity). Appendix A provides a detailed discussion of the variables and parameters used in the estimation of the model.

We obtain the geolocation of cellphone towers in the year 2013, just prior to the outbreak, for the two largest network providers in Liberia, MTN Lonestar and Cellcom. These account for 91% of all mobile subscribers during the year of study (LTA, 2014).[8] The data are obtained from the Liberia Telecommunications Authority (LTA). Appendix Fig. B.1 provides a map of the towers' footprint. We combine these data with the highest resolution global-scale elevation data model publicly available, the 30-meter resolution ALOS Global Digital Surface Model (JAXA, 2016).

Fig. 1 presents the model output on a map of Liberia along with the location of cellphone towers. Areas of Liberia with no cellphone coverage in 2013 roughly correspond to non-populated areas covered by forest.[9] Received power on the ground is measured in decibel-milliwatts (dBm) and typically ranges between −50 and −140 dBm with values closer to zero representing higher signal strength. For ease of interpretation, our measure of coverage uses the absolute value of received power. For 2G GSM networks using a 900 MHz frequency such as the one used in Liberia at the time, sufficient coverage to make a call or send an SMS entails a signal strength below 95 dBm in absolute value (GSMA, 2019), therefore areas shaded in red in Fig. 1 are receiving sufficient cellphone coverage.[10]

3.3. Survey data

We use novel survey data gathered about six months after the end of the epidemic (Maffioli, 2020) to explore potential mechanisms. Phone numbers from 2265 respondents in 571 villages across all of Liberia were selected through random dialing of phone numbers. These respondents were then interviewed through a combination of an Interactive Voice Response (IVR) survey to find out about their location before the beginning of the outbreak and a mobile phone survey conducted by a local NGO. We note that the survey sample is not representative of the national population, instead it is biased towards respondents with access to a cellphone during the time of the survey, i.e. male and educated individuals from urban areas.[11]

3.4. Village location and census data

We obtain GPS coordinates of each village from the Liberia Institute of Statistics and Geo-Information Services (LISGIS).[12] This allows over-imposing the village location data with the spatial radio-wave propagation model in order to pinpoint signal strength for each village in Liberia. We also obtain data on road networks to construct other determinants of EVD such as travel distance to the origin point of the epidemic and Monrovia.

In addition, we gather data from the 2008 National Population and Housing Census (LISGIS, 2008) on population characteristics, such as education, household size, working status, occupation, tribe, religion, housing and asset ownership, which we use to create proxies of village wealth. We aggregate census data at the village level and merge it with our village-level EVD and cellphone coverage measures. We use this

[6] Confirming the exact number of cases by testing was challenging during the epidemic: while the database records about 8308 cases, WHO reports 10,675 probable, confirmed, or suspected cases by the end of the epidemic. The difference in the number of deaths is even starker.

[7] Refer to Crabtree and Kern (2018) for a detailed discussion of the ITM. Other propagation models specifically designed to model cellphone coverage exist. However, these were mainly designed for urban and suburban environments where obstacles to propagation come from building footprints rather than topography. We also note that the ITM has been used in related literature to measure radio coverage Adena et al. 2015, Gagliarducci et al. 2020, Armand et al. 2020 and television coverage (Olken, 2009).

[8] Other four operators existed in the country. The biggest by market share was Comium (less than 8%), while LiberCell and Libtelco had less than 1% and WAT's share was negligible. These companies had mostly users in the capital city where cellphone coverage is existing for the other major companies. Note also that Cellcom is currently Orange given its acquisition by Orange in 2016.

[9] Please see the location of villages in Liberia in Maffioli (2021), Fig. 2; see map of Liberia on forest land at http://www.fda.gov.lr.

[10] Refer to Section 4.1.1 for more information on this cutoff. Additionally, one concern might be that cellphone coverage estimated from the ITM might not always correspond to actual cellphone ownership on the ground. Unfortunately, our data does not allow testing this relationship as we do not have information on cellphone ownership. Instead, we use the Demographic and Health Survey 2013 to assess the relationship between phone ownership and predicted coverage from the ITM, at the district level. Appendix Fig. B.2 shows that there is a positive and statistically significant correlation (0.55) between the proportion of individuals reporting owning a cellphone and the proportion of villages within each district predicted to have coverage.

[11] We refer to Maffioli (2020) for more details on the methodology used to sample and screen respondents and to gather data, and for more details on the sample characteristics and how it compares to a nationally representative sample.

[12] We refer to village for both urban and rural locations. For instance, the capital city entails a unique geographic location in our dataset.

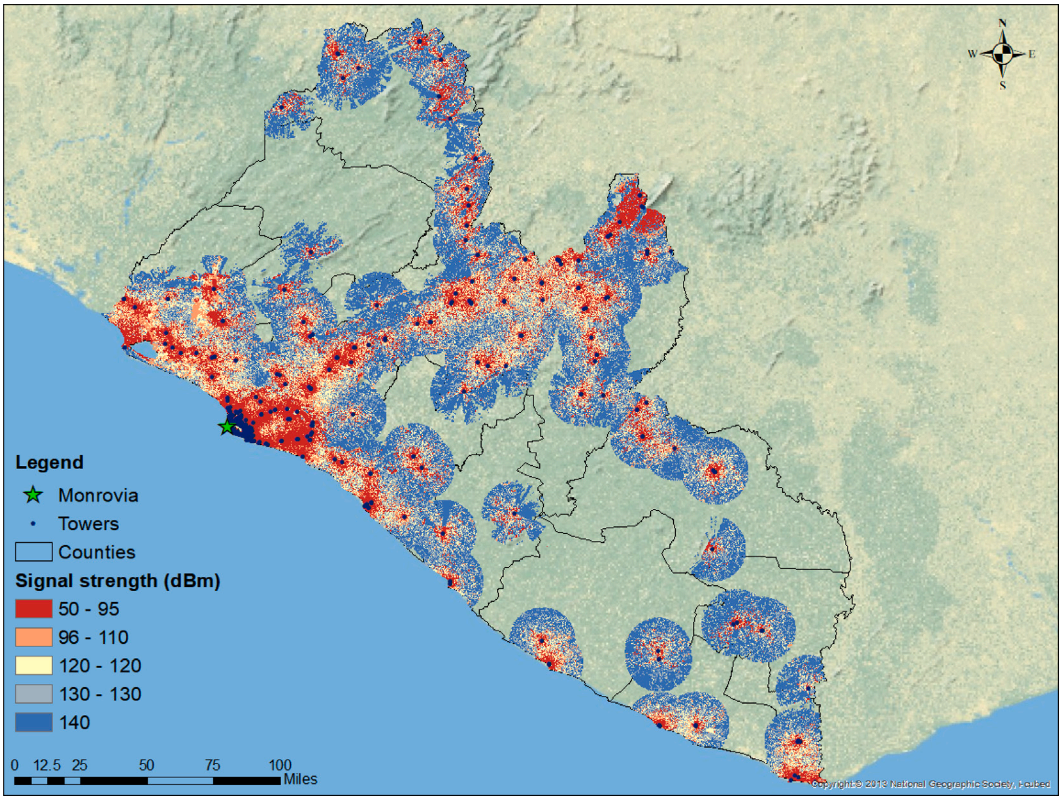


Fig. 1. Irregular Terrain Model, Liberia (2013).
Notes: Cellphone towers' location obtained from the Liberia Telecommunications Authority (LTA). Estimates of the ITM model described in Section 3.2. Lower dBm values mean higher signal strength (i.e., more coverage).

dataset primarily as a source of covariates for the main analysis and to assess the validity of the empirical design. Finally, we use publicly available data on various measures of village exposure to relief efforts, such as the location of CCCs.[13]

3.5. Call detail records (CDR) database

In order to create measures of each village's network, we obtain anonymous CDR data from on of the largest mobile network operators (MNO) in the country at the time of the study. Each record includes a unique (anonymized) device identifier, date and time of call/SMS, identifier for cell tower where call/SMS originated, and identifier for cell tower receiving the call/SMS. We use close to 234 million CDR from the universe of subscribers in the country–more than one million unique users–for June and July of 2015. For reference, Panel (a) of Appendix Fig. B.3 maps all calls during a randomly drawn day of operation (June 15, 2015). Panel (b) maps outgoing calls for the tower servicing Ganta city, Nimba county, as an example. The line color indicates the frequency of calls between the two locations, and the polygons depict Voronoi cells around each tower location.

4. Methods

4.1. Regression discontinuity design (RD)

We estimate the effect of cellphone coverage on the spread of EVD, by employing a regression discontinuity (RD) design that uses signal strength as the forcing variable and the receiver's sensitivity threshold as the treatment cutoff. A receiver's or cellphone's sensitivity threshold is essentially the minimum signal strength required for that phone to be able to send or receive a voice call or SMS. Our baseline RD specification is given by the following equation:

$$EVD_i = \alpha + \beta \text{Coverage}_i + f(\tilde{R}_i) + h(\mathbf{G}_i) + \epsilon_i \quad (1)$$

where EVD_i is an indicator for whether village i was affected by at least one EVD case (probable, confirmed, or death) within our study period (January, 2014–July, 2015). $\tilde{R}_i = -1 \times (R_i - c)$ is the signal strength (measured in dBm) obtained from the ITM in village i net of the receiver sensitivity cutoff c.[14] Values of $\tilde{R}_i$ greater than zero mean that cellphone coverage is available in village i, while negative values mean that the location is below the sensitivity threshold and thus no cellphone coverage is available. $\text{Coverage}_i = \mathbb{1}\{R_i \geq c\} = \mathbb{1}\{\tilde{R}_i \geq 0\}$ is an indicator for whether village i has coverage (i.e., signal strength is higher than cutoff c). $f(\tilde{R}_i)$ is the RD polynomial. Our main analysis uses a local linear specification with a bandwidth h around c, optimally determined as in Calonico et al. (2014), and a triangular weighting kernel.[15] $h(\mathbf{G}_i)$ is a flexible polynomial in topographic characteristics of village i. Standard errors are clustered at the district level. Coefficient β identifies the causal effect of cellphone coverage under certain assumptions that are discussed in Section 4.1.2.

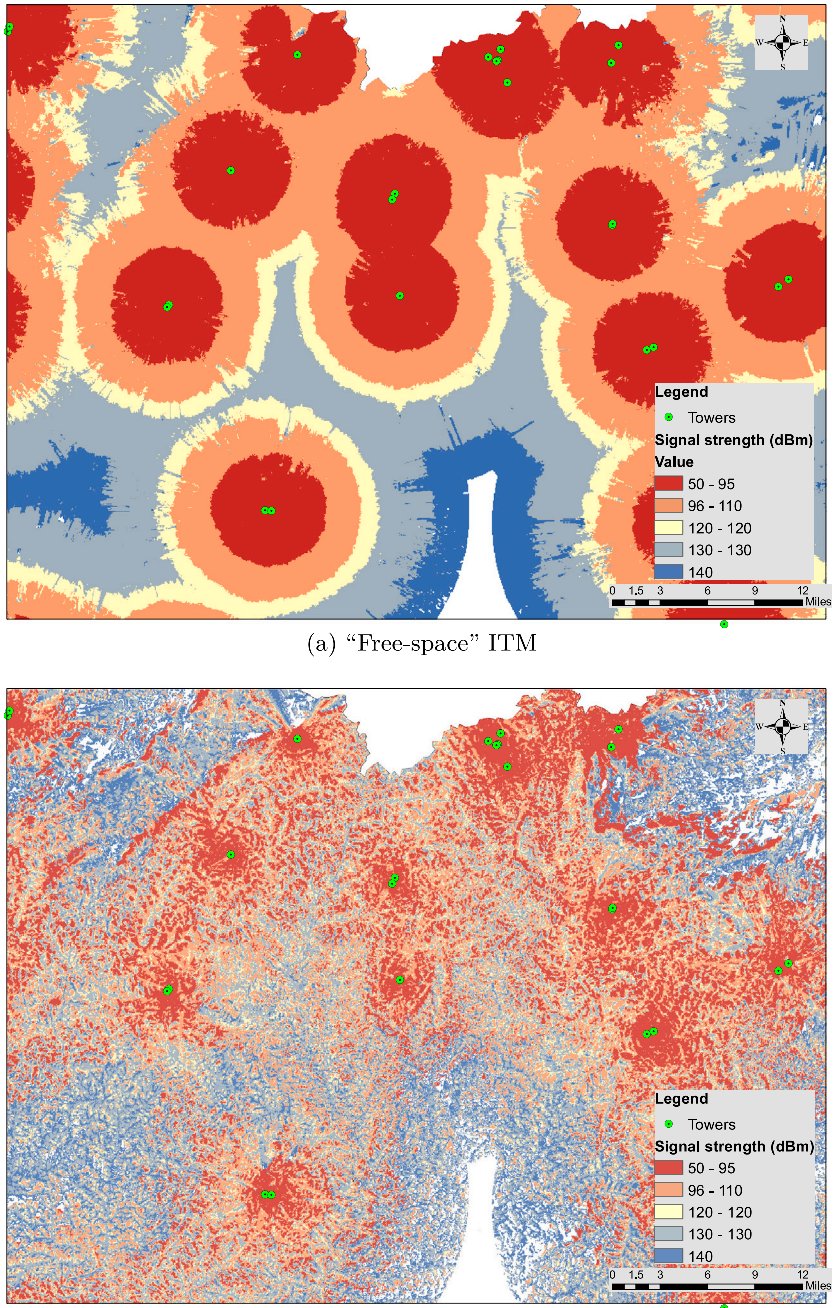
4.1.1. Determining the sensitivity cutoff

We define the minimum required signal strength or sensitivity cutoff c based on the industry standard for the specific network used in Liberia

13 See the Ebola crisis page at Humanitarian Data Exchange, https://data.humdata.org/ebola.

14 For convenience, we multiply the normalization by −1 simply to make positive values mean more coverage while negative values no coverage.

15 This is the equivalent of running a weighted regression that sets $f(\tilde{R}_i) = \tilde{R}_i + \text{Coverage}_i \times \tilde{R}_i$ where weights are obtained using a triangular kernel $K(u) = 1 - |u_i|$ when $|u_i| \leq 1$ and $K(u_i) = 0$ when $|u_i| > 1$, with $u_i = \tilde{R}_i/h$. In our estimation, we use the bias-corrected estimator proposed in Calonico et al. (2014).



(a) "Free-space" ITM

(b) ITM with Topography

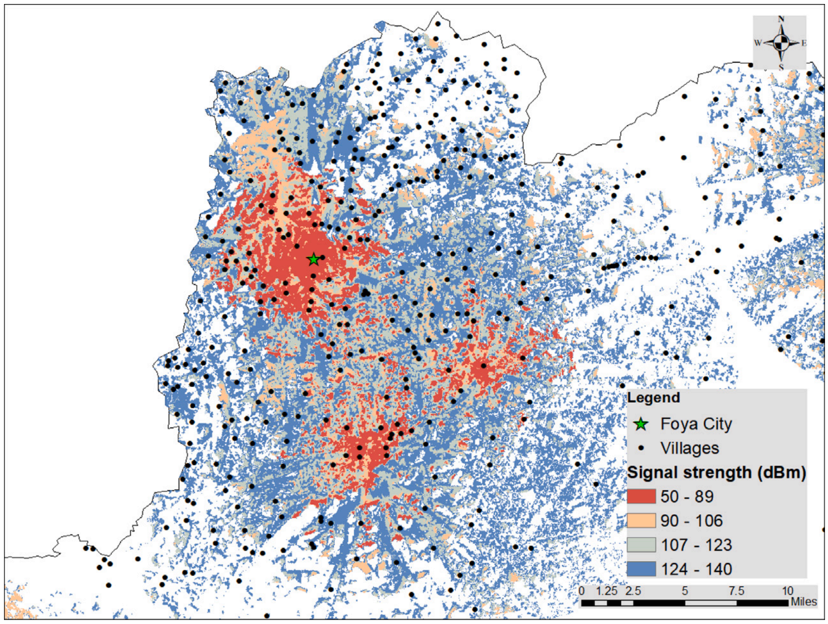
Fig. 2. ITM "Free-space" versus ITM with Topography, Liberia (2013).
Notes: Green dots give the location of towers. Sample along the border of Nimba and Bong counties in Northern Liberia. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

at the time of the study: GSM900, a 2G network using a 900 MHz frequency band.[16] Assignment into coverage based on this approach is more accurate than existing work relying on GSMA coverage maps ((Guriev et al., 2020), Gonzalez (2021)) that implicitly use a uniform cutoff of −100 dBm for all networks and countries despite this cutoff varying significantly across different types of networks.

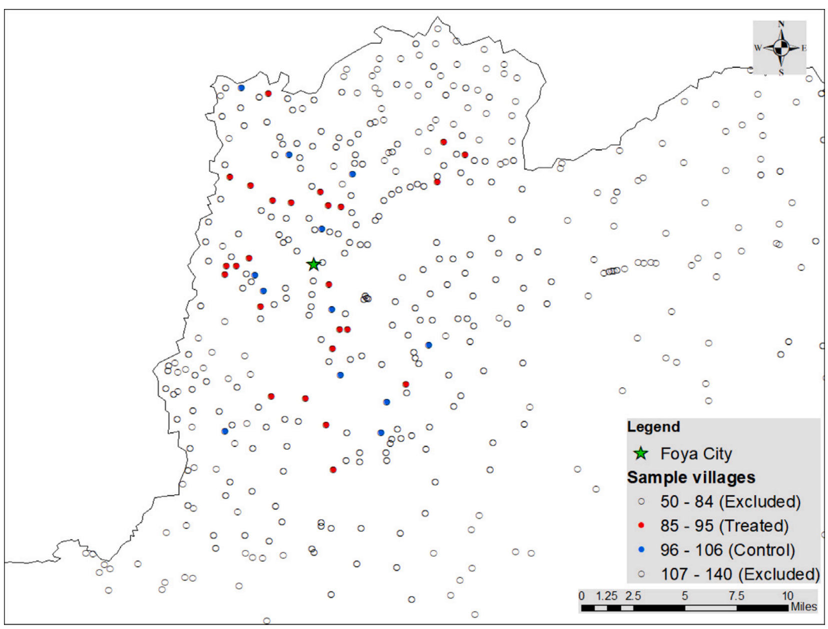
Specifically, we use a cutoff of −95 dBm (i.e., $c = 95$). As mentioned, this is the standard for GSM900 networks proposed by the GSMA and recommended by the Cellular Telephone Industries Association (CTIA) for industry-standard test procedures (Razally, 2015).[17] Independent

[16] More information specific to Liberia found in the Liberia Telecommunications Authority (LTA) spectrum management page: https://www.lta.gov.lr/spectrum-management. About 93% of the towers in Liberia at the time used a 900 MHz band (GSM900), while under 7% of the towers used 1800 MHz (GSM1800). The latter towers were exclusively located within the Monrovia area and therefore are not part of our estimation sample. These numbers are computed by the authors and come from MTN–the largest MNO in Liberia at the time–which provide information on the frequency bands of their towers.

[17] The cutoff of −95 dBm is the accepted standard for "phone between the head and hand (BHH)" position which is the most realistic position (phone over the ear). Other cutoffs are for free-space (i.e., signal propagating in a vacuum) or browsing position (i.e., phone in hand for browsing). This latter



(a) ITM detailed



(b) Villages within 10dBm Bandwidth of Cutoff

Fig. 3. Irregular Terrain Model with Village Sample.
Notes: Cellphone towers' location obtained from the Liberia Telecommunications Authority (LTA). Estimates of the ITM model described in Section 3.2. Lower dBm values mean higher signal strength (i.e., more coverage). Dots indicate the location of villages. Red dots indicate villages with just enough coverage (treatment). Blue dots denote villages missing coverage (control). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

tests by industry regulators empirically back our choice of cutoff. In these tests, phones on a GSM900 network had a mean cutoff of −95 dBm with some variation in tested cutoffs resulting from device characteristics, positioning, etc.[18] We note that the $c = 95$ cutoff should be interpreted as the *expected* cutoff and the effect of coverage documented in this paper should therefore be interpreted as an intent-to-treat (ITT) given the degree of variability in this cutoff.[19]

position is more relevant for networks with data/internet access and therefore more advanced than the one used in Liberia in 2013.

[18] Razally (2015) and GSMA (2019) performed independent tests on the sensitivity cutoffs on BHH position for GSM900 networks for OFCOM (the UK's communication regulator) and the GSMA, respectively. The sensitivity cutoffs for the tested devices in the GSMA study were on average −94.45 (95% CI = −(−95.04, −93.86)). In the OFCOM study, the overall sensitivity cutoffs averaged −94.46 (95% CI = (−95.10,−93.82)) with the cutoffs for lower channels of the band (925.2 MHz) averaging −94.93 (95%CI = (−96.21, −93.65)) and −94.83 (95% CI = (−95.89, −93.77)) for medium channels (942.6 MHz). Calculations performed by the authors.

[19] One can interpret this as the typical noncompliance issue in a fuzzy RD setup where some villages might be wrongly assigned into treatment/control

Table 1
Determinants of coverage.

	Dep. variable: Signal strength (dBm)		
	Full sample (1)	Within 20 dBm (2)	Within 10 dBm (3)
Topographic controls			
Elevation (m)	−0.082***	−0.032**	0.005
	(0.024)	(0.014)	(0.009)
Slope (%)	−0.283	1.129***	0.828***
	(0.455)	(0.379)	(0.256)
Demographic controls			
Household size	−0.194	0.014	0.118
	(0.171)	(0.131)	(0.087)
Population (log)	0.880***	0.225	−0.013
	(0.166)	(0.162)	(0.117)
Female (%)	−0.367	1.805	1.474
	(2.456)	(3.029)	(1.775)
Married (%)	−4.267*	−1.750	1.236
	(2.170)	(1.679)	(1.481)
Christian (%)	1.974	3.553**	−0.340
	(2.593)	(1.543)	(1.603)
Muslim (%)	1.655	−0.263	−1.721
	(3.180)	(2.473)	(1.982)
African religion (%)	6.102	6.790	−0.579
	(5.602)	(6.518)	(7.128)
Kpelle (%)	−0.040	−1.410	−1.190
	(1.159)	(1.430)	(0.946)
Bassa (%)	3.648	1.031	0.497
	(2.229)	(1.771)	(1.163)
Economic controls			
Primary education (%)	−0.347	0.142	−1.199
	(1.665)	(1.269)	(1.200)
Secondary education (%)	5.027**	2.097*	0.925
	(2.091)	(1.107)	(1.137)
Owns house (%)	1.084	−0.498	−0.403
	(1.189)	(0.858)	(0.568)
House condition: Good (%)	9.476***	4.908***	0.858
	(2.175)	(1.332)	(0.941)
Assets ownership (%)	2.350	1.413	0.505
	(3.844)	(2.828)	(1.621)
Distance to Monrovia (km)	−0.098**	−0.013	−0.020
	(0.047)	(0.049)	(0.024)
Distance to closest city (km)	−0.306***	−0.064	−0.035
	(0.113)	(0.088)	(0.049)
Observations	7014	3839	1913
P-value for F-test for joint significance:			
Topographic controls	0.00	0.00	0.00
Demographic controls	0.00	0.01	0.35
Economic controls	0.00	0.00	0.79

Notes: Standard errors clustered at district level. All specification include district fixed effects. *, **, and *** indicate 10, 5, and 1 percent significance, respectively. Sources of data: ALOS Global Digital Surface Model; 2008 National Population and Housing Census.

We further corroborate our choice of −95 dBm using three separate methods designed to empirically detect unknown discontinuities points. First, we use the Difference in Kernels estimator described in Qiu (2011) and recommended in Porter and Yu (2015). Second, we employ the maximum R^2 strategy, described in Card et al. (2008) that estimates a polynomial between our outcome Y_i and the estimated signal strength R_i for several potential cutoffs r within a fixed interval $[\underline{R}, \overline{R}]$. Lastly, we use a modification of the method proposed by Spokoiny (1998) adapted to our RD setting. All these three methods consistently point to a value of 95 dBm as the sensitivity cutoff (see Appendix C for more details).

4.1.2. Internal validity of the RD design

Coefficient β in Eq. (1) identifies the causal effect of cellphone coverage under the assumption that potential outcome functions $E[EVD(1)|\tilde{R}]$ and $E[EVD(0)|\tilde{R}]$ are continuous at the coverage cutoff c, where one and zero denote assignment and non-assignment into treatment, respectively. This entails that observable and unobservable characteristics must transition smoothly across the coverage cutoff, so that villages with signal strength just below the cutoff can serve as a valid counterfactual for villages just above the cutoff. Within a reasonable bandwidth around this cutoff, whether a village receives just enough signal strength is determined by minor variations in topography. Fig. 2 compares signal strength obtained from estimating the ITM without accounting for topography (Panel a) and accounting for it (Panel b). Note how terrain topography leads to arbitrary blocking and diffraction of the signals.

Fig. 3 provides a visual depiction of our empirical strategy. In panel (a), we provide a closer look at the estimated signal strength taking as example three cellphone towers near Foya city along with the surrounding villages.[20] Panel (b) highlights villages that are part of

due to variability in the cutoffs for those locations. We do not observe actual coverage at the village level to implement a fuzzy design, therefore we have to rely on estimating an ITT using expected coverage.

[20] Foya is a city within Lofa county and is one of the largest cities in Liberia close to the border with Sierra Leone and Guinea, where the outbreak originated.

a hypothetical RD design that uses a bandwidth of 10 dBm around the coverage cutoff. First, note that at the margin of coverage, there is rich spatial variation in treatment and control villages. Most importantly, at this margin, treatment status is determined by minor topographical differences that dictate whether enough signal reaches the ground.

Table 1 presents results from a linear regression of signal strength, measured in dBm, on a rich set of ex-ante village-level covariates. If selection is not an issue near the cellphone coverage cutoff, then we should expect ex-ante demographic and economic characteristics to not predict signal strength. There is significant evidence of selection into cellphone coverage when considering the entire sample of villages (column (1)): elevation, population size, and some socioeconomic indicators are strongly correlated with signal strength while the set of topographic, demographic, and economic controls are all jointly statistically significant. However, as we restrict our analysis to villages that are within a close window of the coverage cutoff (column (3)), only the topographic controls remain jointly significant. This suggests that at the margin of coverage, what largely determines cellphone availability are minor differences in topography, and not endogenous village characteristics.

We also linked villages to the closest Demographic Health Survey (DHS, 2013) enumeration units to predict signal strength with covariates that are measured closer to the onset of the epidemic. There are 322 enumeration units in DHS 2013, linked to the closest 244 villages. However, it is important to keep in mind that because of the GPS coordinate displacement process carried out (urban clusters are displaced a distance up to two kilometers (0–2 km) and rural clusters are displaced a distance up to five kilometers (0–5 km), with a further, randomly-selected 1% (every 100th) of rural clusters displaced a distance up to 10 kilometers (0–10 km)), clusters could be wrongly assigned across the cut-off to coverage and non-coverage areas. With this caveat in mind, we replicated Table 1 and found no evidence of selection into cellphone coverage within the 10 dBm bandwidth (Appendix Table B.1).

Appendix Fig. B.4 further confirms that selection of villages near the coverage cutoff does not seem to be an issue as there is no significant jump in the density of the forcing variable. Appendix Table B.2 also provides summary statistics for village-level characteristics for several bandwidths around the cellphone coverage threshold. Columns (1) and (2) report the mean of these variables by coverage status for the entire sample. Columns (4) and (5) repeat the exercise for villages within 20 dBm on each side of the sensitivity cutoff. Columns (7) and (8) narrow the window of analysis to a 10 dBm bandwidth. Columns (3), (6), and (9) report the clustered standard errors of the difference in means between villages with and without cellphone coverage. Comparing columns (1) and (2) confirms that, among other things, villages in areas with coverage tend to be at lower elevation and on a smoother terrain, have a smaller average household size, higher levels of primary and secondary education, higher asset ownership and quality housing, and they live much closer to the capital Monrovia and to the closest main city. As we restrict our analysis to villages at the margin of cellphone coverage, however, most statistically significant differences disappear (columns (6) and (9)). Overall, these results provide support for the continuity assumption discussed above.[21]

4.1.3. External validity of the RD design

Given the localized nature of the RD design, we explore how our estimation sample differs from the rest of the Liberian population to broadly characterize which sub-population our design speaks to.

Appendix Table B.3 presents summary statistics for villages within our estimation sample (column 1), and outside (columns 2 and 3). We define the estimation sample for this comparison as the sample within the (Calonico et al., 2014)-estimated bandwidth for our main RD design (column 1 of Table 2) at 8.9 dBm. Column 2 presents summary statistics for the sample below the lower limit of the bandwidth (i.e. < -8.9 dBm) while column 3 presents summary statistics for the sample above the upper limit of the bandwidth (i.e. $>$ 8.9 dBm). Broadly speaking, locations with robust signal strength (column 3) are generally closer to the towers and therefore more likely to be urban and developed, while locations far from any viable signal (column 2) are far away from the towers and therefore more likely to be isolated and rural. For ease of interpretation, Appendix Fig. B.6 also presents the results by standardizing all the variables in Table B.3 and showing how the means of villages below (red) and above (blue) the bandwidth limits deviate from the estimation sample mean.

Overall, this exercise points to clear distinctions between our sample of analysis and locations outside the sample. Relative to villages with strong signal strength (above the bandwidth), our sample villages are quite similar in terms of topographic characteristics and most demographic indicators.[22] However, they are significantly poorer and less educated while being as centrally located as richer villages. This latter fact, in particular, points to our sample villages likely including peri-urban locations and the urban poor—a subpopulation of key policy relevance. With this in mind, results from our design speak to this likely group and not to more richer/urban or more isolated/rural locations.

4.2. Panel-regression discontinuity design (RD)

We also employ a more dynamic model that explores whether cellphone coverage helped contain the disease by exploiting the (monthly) time variation of the EVD epidemic. Specifically, we disaggregate our EVD measure in Eq. (1) and create a village-by-month panel database. We are interested in learning whether the likelihood that EVD spreads into a village from surrounding affected villages diminishes with cellphone coverage. We use the following empirical specification:

$$EVD_{ijt} = \alpha + \beta \text{Coverage}_{ij} + \gamma EVD_{j(i),t-1} + \delta \text{Coverage}_{ij} \times EVD_{j(i),t-1} + f(\tilde{R}_{ij}) + \lambda_j + \nu_t + \epsilon_{ijt} \quad (2)$$

where EVD_{ijt} is an indicator for whether a (probable, confirmed, or death) EVD case was recorded in village i in district j in month t. $\tilde{R}_{ij}$, $\tilde{R}_{ij}$, and Coverage_{ij} are defined as in Eq. (1) since these variables do not vary by month. $\text{EVD}_{j(i),t-1}$ is an indicator for whether district j, where village i is located, was affected by EVD in the previous month $t-1$. λ_j and ν_t are district and month fixed effects, respectively. The district fixed effects account for any time-invariant unobservables that may lead to endogenous selection into EVD within a village's district.

To account for endogenous selection into cellphone coverage, we integrate into our panel study an RD design that uses a linear specification in $\tilde{R}_{ij}$, while restricting our analysis to the same bandwidth as the baseline specification in Eq. (1).[23] Eq. (2) estimates the likelihood of an EVD case in village i given that there was at least one EVD case within that village's district in the last month. Coefficient δ estimates how this contagion effect differs by whether village i has coverage or not.

[21] For a graphical depiction of the continuity of time-invariant covariates across the coverage cutoff, refer to Appendix Fig. B.5.

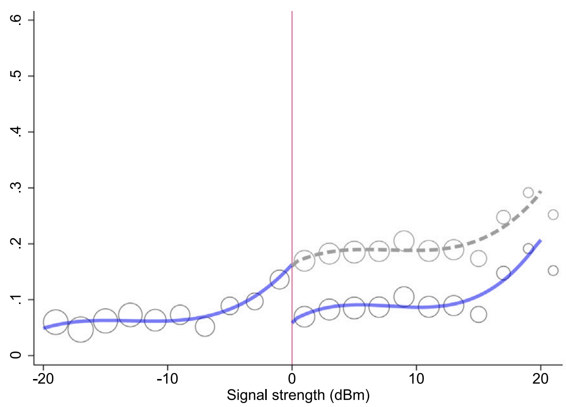
[22] In terms of topographic characteristics, our sample villages are no different than villages with strong signal strength (above the bandwidth) and significantly less rugged than more isolated villages (below the bandwidth). In terms of demographic characteristics, our sample villages are quite similar to richer villages in terms of ethnicity, religion, and household size, among other characteristics. However, they are closer to more isolated villages in terms of population size, and educational attainment. In terms of economic characteristics, the sample villages tend to own less assets and have lower quality housing than villages with strong signal strength (above the bandwidth), but tend to have significantly better housing than isolated villages. They also tend to be as close to main cities and Monrovia as the richer villages and significantly closer than unconnected villages (below the bandwidth).

[23] Specifically, we let $f(\tilde{R}_{ij}) = \theta_1 \tilde{R}_{ij} + \theta_2 \text{Coverage}_{ij} \times \tilde{R}_{ij} + \theta_3 EVD_{j(i),t-1} \times \tilde{R}_{ij} + \theta_4 \text{Coverage}_{ij} \times EVD_{j(i),t-1} \times \tilde{R}_{ij}$.

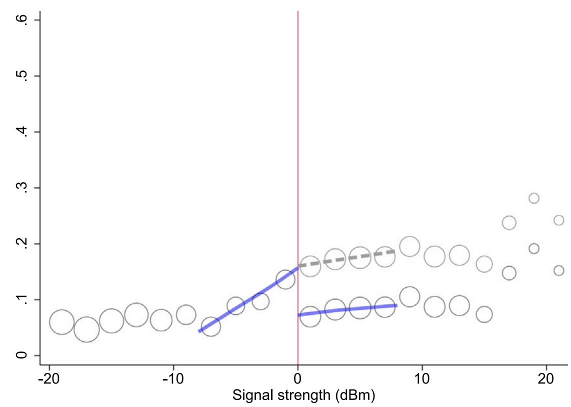
Table 2
Effect of coverage on likelihood of EVD case.

	Dep. Variable = $\mathbb{1}${Number of Reported EVD Cases > 0}					
	Baseline (1)	Controls (2)	Topography Poly (3)	Polynomial RD (4)	Kernel choice (5)	Probit model (6)
Coverage	−0.108**	−0.100**	−0.110**	−0.096***	−0.106***	−0.429**
	(0.048)	(0.042)	(0.048)	(0.035)	(0.041)	(0.206)
Mean outside coverage	0.09	0.09	0.09	0.06	0.09	0.06
Bandwidth (dBm)	8.99	8.17	9.07	50.00	8.21	8.99
Observations	1547	1547	1741	7014	1547	1547
Districts	83	83	84	115	83	83
Marginal Effect	–	–	–	–	–	−0.070

Notes: Columns (1) and (2) present estimates of β using a local linear regression specification of Eq. (1). All specification include controls for elevation and slope. Column (2) adds controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Optimal bandwidth chosen as in Calonico et al. (2014). Column (3) uses topography polynomial: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. It also includes distance to tower as additional control. Column (4) uses a wider bandwidth of 50 dBm and a third-degree polynomial specification for $f(\tilde{R}_i)$. Column (5) uses a rectangular kernel. Column (6) estimates a probit specification of Eq. (1). Optimal bandwidths chosen as in Calonico et al. (2014) except for column (4) which uses a fixed, wider bandwidth. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.



(a) Polynomial RD



(b) Local Linear Regression

Fig. 4. Regression Discontinuity (RD) Plots for Likelihood of EVD Case.
Notes: Dots give EVD likelihood for each bin of signal strength (dBm). Dot size weighted by the number of observations within each bin. "Signal strength (dBm)" refers to the received power or signal strength, measured in decibel milliwatts, at the village level. Signal strength is normalized so that negative values mean no coverage (i.e., not enough signal strength). Bin width is 2 dBm. Window of analysis is −20 to 20 dBm. Solid lines give the predicted values from a regression of the outcome variable on a third-degree polynomial in distance to threshold that uses a triangular kernel (panel a) and a linear trend using a triangular kernel (panel b). Light gray dots and dashed lines give a representation of the potential outcomes by shifting upwards the graph for positive dBm values to the point where the trends on each side of zero intersect. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

5. Results

5.1. Graphical analysis

Fig. 4 presents regression discontinuity plots for the outcome variable EVD_i in Eq. (1). The solid vertical line indicates the cellphone coverage cutoff, and the signal strength is normalized so that positive (negative) dBm values represent coverage (no coverage). The circles give the averages of the outcome variable for 2 dBm signal strength bins, while the circle size is weighted by the number of villages within each bin.[24] The solid trend lines predict EVD_i using a third-degree polynomial in the normalized signal strength (panel a) and a linear specification (panel b). The gray dots and lines represent potential EVD_i by shifting upwards the predicted trend and binned averages of treated villages to the point where they intersect the observed trends in non-treated locations.

Panel (a) shows continuity in potential EVD in the absence of coverage. As soon as coverage is available, the likelihood of an EVD case drops by about 10 percentage points. In addition, there is no clear change in the number of villages near the cutoff consistent with our density analysis in Section 4.1.2. Panel (b) provides a graphical depiction of our baseline results presented in Section 5.2 below. Specifically, the difference at the cutoff gives the RD coefficient β in the local linear regression specification of Eq. (1).

5.2. Main RD estimates

Table 2 presents estimates of the RD coefficient β in Eq. (1). Given the relationship between signal strength and topography (Table 1), all specifications include controls for terrain elevation and slope. Columns (1)–(3), and (5) use a linear specification of the RD polynomial while column (4) uses a third-degree polynomial in signal strength given the

[24] The relationship between signal strength and distance to tower is complex due to topography and the nonlinear nature of the relationship, but as a reference, in our sample, moving one kilometer away from a tower decreases signal strength by about 2.6 dBm, on average.

wider bandwidth. In column (1), we document a reduction of about 10.8 percentage points in the likelihood that a village has an EVD case relative to villages that are just under the coverage cutoff. The optimal bandwidth used in this empirical specification is about 9 dBm.[25] The estimates remain very similar after including a set of socio-economic and demographic characteristics (column (2)), confirming that the estimated drop in the likelihood of EVD at the cellphone coverage cutoff is not explained by these covariates.

Columns (3)–(6) show that the results are robust to a set of alternative specifications. Column (3) includes a flexible polynomial in elevation and slope to capture whether the effect on EVD is simply driven by changes in topography captured by the ITM. The effect remains robust suggesting that this is unlikely. Column (4) estimates a parametric RD specification that uses almost the entire sample of villages (50 dBm bandwidth) and a flexible third-degree polynomial in signal strength. The results are consistent with the optimal bandwidth estimates in columns (1)–(3), although there is a gain in precision given the larger number of observations. Column (5) confirms that the estimated effect is robust to the choice of kernel (uniform) and that are not driven entirely by observations near the cutoff. Given our binary outcome variable, column (6) estimates a Probit model within a specified bandwidth around the coverage cutoff. The marginal effect (at 7 percentage points) is not far from our previous estimates. We also probe the sensitivity of our baseline results to the choice of bandwidth. Appendix Fig. B.7 confirms that the coverage effect on EVD remains negative and statistically significant for a wide set of bandwidths.

Appendix Table B.4 shows that villages with cellphone coverage experienced about one less EVD case relative to villages without coverage. However, the magnitude of the coefficients should be interpreted with caution due to the potential measurement error in this outcome. Appendix Tables B.5–B.7 also show that the findings are robust to alternative measures of EVD: whether a suspected death from EVD was recorded in the village, total number of suspected deaths from EVD, and the total number of months the village was affected by the epidemic.[26] Lastly, it is important to highlight that if access to cellphone coverage arbitrarily led to more reporting of cases in villages with coverage relative to villages without coverage, then our estimates would be a lower bound on the actual magnitude of the drop in EVD cases due to coverage.

5.2.1. Panel RD estimates

Columns (1) and (2) in Table 3 present panel-RD estimates of the effect of cellphone coverage on the likelihood that a village has an EVD case. Column (3) presents estimates of the contagion effect, namely the association between a village's district having an EVD case in the last month and the likelihood that village subsequently has an EVD case in month t. Columns (4) and (5) present estimates on how this contagion effect varies by whether village i has cellphone coverage or not. In line with previous findings (Table 2), we find that cellphone coverage leads to a 0.74 percentage point drop in the likelihood of an EVD case in any given month (column (1)). Column (3) provides strong evidence of contagion effects within districts: the likelihood of a village reporting an EVD case in a given month increases by 0.45 percentage points if there was at least one EVD case in the previous month within that village's district. Column (4) disaggregates this contagion effect by whether a village has cellphone coverage or not. The results provide evidence that the spread of the disease is considerably undermined by the presence of cellphone coverage. In fact, being in a village with coverage overcomes the contagion effects, even if the village is surrounded by at least one other village with EVD in its district. The estimates are quantitatively similar after adding controls (column (5)).

Table 4 estimates Eq. (2) using EVD exposure within village i (rather than in village i's district as in Table 3). This aims to explore how past exposure to EVD within a given village affects EVD outcomes for that same village and how this effect varies with coverage status. Column (1) shows that coverage in village i decreases the likelihood of an EVD case in any given month by about 0.7 percentage points. Column (2) estimates the within-village contagion effect: given an EVD case in village i, the likelihood of a future EVD case in that village increases by about 21 percentage points. Columns (3)–(6) present the effect of past exposure in the village on different EVD measures. Column (3) shows that experiencing an EVD case increases the likelihood of a future case by about 24 percentage points in villages without coverage. However, if village i has cellphone coverage then that contagion effect significantly drops by about 15 percentage points. Similarly, columns (4)–(6) show that after experiencing an EVD case, villages with coverage can expect 2.7 less EVD cases (column 4), 0.43 lower suspected EVD deaths (column 6), and a 15 percentage point lower likelihood of a suspected EVD death (column 5) than villages without coverage.

We complement our panel-RD findings with results from a Cox proportional hazards model that estimates the hazard ratio as a function of cellphone coverage (and geographic controls). To better account for selection into coverage, we limit the analysis to villages within the same bandwidth as the baseline specification in Eq. (1). We structure our data as one observation per village, with coverage status and months until the first EVD case as our key variables. We assume our model to be censored since most villages did not experience a case, therefore we define a failure as whether a village has had any EVD case within the study period. Fig. 5 presents the survival curves by coverage status as well as the estimated hazard ratio for coverage. Panel (a) shows that having cellphone coverage decreases the hazard rate by 60% (HR = 0.404). The survival curves in fact shows that throughout the study period the survival probabilities at any given month (i.e., experiencing no cases) remain statistically significantly higher in villages with cellphone coverage than in villages without coverage.

Panel (b) explores the survival probabilities and hazard ratio conditional on having experienced a case. Our aim is to explore the effect of coverage on how villages fare after being exposed to a first EVD case. We restructure our data such that a failure is defined as having an additional EVD case, and time is defined as months until the additional case after experiencing a first EVD case. Again, we assume our model to be censored since many villages do not experience multiple cases. The sample size is significantly smaller since few villages within our small bandwidth have experienced a case. Note that, conditional on experiencing a first EVD case, having cellphone coverage decreases the hazard rate of having an additional EVD case by 47% (HR = 0.533). Similarly, the survival probabilities remain higher (even though not statistically significant) in villages with cellphone coverage than in villages without coverage. Panels (c) and (d) of Fig. 5 repeat the exercise defining a first suspected death (panel (c)), and an additional suspected death (panel (d)) as the failure. Again, we find a lower hazard (and consistently higher survival probabilities) for villages with coverage relative to villages without coverage.

5.3. Robustness Checks

5.3.1. "Free-Space" ITM

We proceed by presenting further evidence supporting our main results. We use the ITM to estimate signal strength without accounting for topography. This is typically called a "Free-Space" model and it assumes that there is a direct line-of-sight between the tower and the receiver (Olken, 2009). Fig. 2 compares the output from the two

[25] As a reference, within this bandwidth, the average distance between a treated village and the closest control village is about 2.6 km with a standard deviation of 2.2 km. This suggests that control villages are also a good counterfactual in terms of geographic distance.

[26] Unfortunately, the Ebola data do not allow us to distinguish between EVD and non-EVD deaths since all cases reported in the patient database from the MOH were suspected with EVD and not every case was tested before dying.

Table 3
Effect of coverage on likelihood of EVD, by past EVD exposure.

	Dep. Variable = $\mathbb{1}\{$Number of Reported EVD cases > 0$\}$				
	(1)	(2)	(3)	(4)	(5)
Coverage	−0.007**	−0.007**		0.000	0.001
	(0.003)	(0.003)		(0.002)	(0.002)
$\text{EVD}_{j(i),t-1}$			0.005*	0.019**	0.018**
			(0.002)	(0.009)	(0.008)
Coverage $\times \text{EVD}_{j(i),t-1}$				−0.021***	−0.020***
				(0.008)	(0.007)
Mean	0.007	0.007	0.015	0.015	0.015
Observations	29 393	29 393	27 846	27 846	27 846
Bandwidth (dBm)	8.99	8.99	8.99	8.99	8.99
Districts	83	83	83	83	83
District FE	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes

Notes: Each observation is a village-month. Coverage_i equal 1 if village i has coverage. $\text{EVD}_{j(i),t-1}$ equals 1 if there is one or more EVD cases in at least one village within village i's district in the past month. Standard errors clustered at district level. Columns (2), (3), and (5) include controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table 4
Effect of coverage on EVD measures, by past EVD exposure.

	Reported EVD (= 1) (1)	Reported EVD (= 1) (2)	Reported EVD (= 1) (3)	Reported EVD (No.) (4)	Suspected death (= 1) (5)	Suspected death (No.) (6)
Coverage	−0.007**		−0.005**	0.009	−0.008**	−0.011
	(0.003)		(0.002)	(0.015)	(0.003)	(0.007)
EVD_{t-1}		0.213***	0.244***	1.367***	0.221***	0.500***
		(0.047)	(0.083)	(0.487)	(0.065)	(0.186)
Coverage $\times \text{EVD}_{t-1}$			−0.146**	−2.770*	−0.147	−0.434**
			(0.067)	(1.561)	(0.096)	(0.217)
Mean	0.007	0.249	0.249	1.662	0.219	0.403
Observations	29 393	27 846	27 846	27 846	27 846	27 846
Bandwidth (dBm)	8.99	8.99	8.99	8.99	8.99	8.99
Districts	83	83	83	83	83	83
District FE	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each observation is a village-month. Outcome variables are an indicator for whether village i had any EVD case in month t (columns 1–3), the number of EVD cases reported in village i in month t (column 4), an indicator for whether village i had a suspected death in month t (column 5), and the number of suspected EVD deaths in village i in month t (column 6). EVD_{t-1} is a binary variables equal to one if village i had any EVD case in month $t-1$. Coverage is an indicator for whether village i has coverage. All specifications include controls for elevation, slope, average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Standard errors clustered at district level. Mean refers to the mean of the outcome variable for villages without coverage and within the specified bandwidth (column 1) and the mean of the outcome variable for villages with at least one EVD case and within the specified bandwidth (columns 2–6). All specifications use the optimal bandwidth of our baseline specification (column 1 of Table 2. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

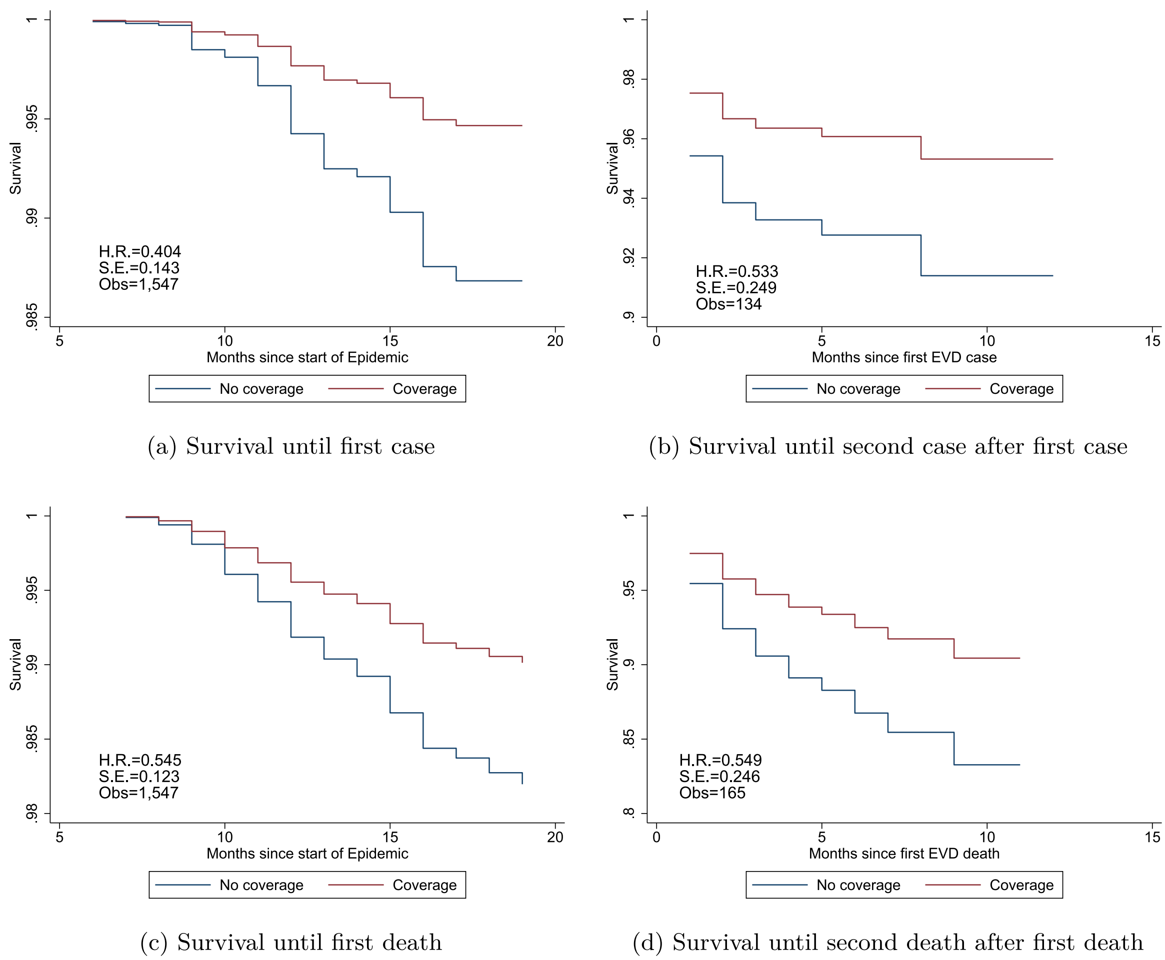
models. Note that minor changes in terrain topography lead to arbitrary blocking or diffraction of the signals. Following Olken (2009), we use the free-space predicted signal strength and coverage as additional controls in Eq. (1). This allows identifying our coverage effect using the variation in signal strength that is solely due to exogenous changes in topography. Intuitively, by controlling for free-space signal strength, we are comparing two villages that are observationally identical in terms of the signal strength they would have received in "free-space", but that differ in their actual signal strength due to exogenous terrain characteristics. We estimate an expanded version of Eq. (1):

$$EVD_i = \alpha + \beta \text{Coverage}_i + f(\tilde{R}_i) + \delta \text{Coverage}_i^{Free} + f(\tilde{R}_i^{Free}) + h(\mathbf{G}_i) + \epsilon_i \quad (3)$$

where $\tilde{R}_i^{Free} = R_i^{Free} - c$ is the received power (measured in dBm) using the "free-space" ITM in village i net of the receiver sensitivity cutoff c. $\text{Coverage}_i^{Free} = \mathbb{1}\{\tilde{R}_i^{Free} \geq 0\}$ is an indicator for whether village i has coverage (i.e., received power is higher than the cutoff c). Coefficient β gives the effect of coverage after controlling for free-space coverage Coverage_i^{Free} and a flexible form of free-space signal strength $f(\tilde{R}_i^{Free})$. We estimate Eq. (3) within a bandwidth of the coverage cutoff c that is optimally determined following (Calonico et al., 2014). The remaining terms are defined as in Eq. (1). Note from Eq. (3) that we are combining our RD strategy comparing villages at the margin of coverage with the free-space strategy. This allows obtaining results that are potentially more robust to any concerns of endogenous selection into coverage.

Columns (1)–(4) of Table 5 present the estimates for different versions of Eq. (3). All specifications include topographic, economic, and demographic controls.[27] The effect of coverage remains statistically significant and similar in magnitude to our main results in Table 2 even after controlling for free-space coverage (column (1)) and free-space signal strength (column (2)). Columns (3) and (4) present the fully specified version of Eq. (3). Again, we find that our coverage effect remains robust even after including a flexible polynomial in topography (column (4)). We perform two falsification tests designed to rule out the possibility of the modeling details of the ITM driving our results. First, column (5) presents results from estimating Eq. (1) using free-space

[27] Refer to the notes on Table 5 for a list of these controls.



(a) Survival until first case

(b) Survival until second case after first case

(c) Survival until first death

(d) Survival until second death after first death

Fig. 5. Survival probabilities and Hazard Ratios by Coverage Status.
Notes: Results from a Cox proportional hazards model that estimates the hazard ratio as a function of coverage, and geographic controls and restricts the analysis within the signal strength bandwidth used in our baseline specification in Table 2 (8.99 dBm). Failure is given by whether a village had a first EVD case (panel a) or EVD death (panel c) and whether they had an additional EVD case (panel b) or death (panel d). Model is assumed to be censored since some villages do not have a case. Figures present the estimated hazard ratio (HR) for cellphone coverage, the standard error clustered at the district level, and the number of observations.

coverage instead of actual coverage. As expected, we find no effect when using the free-space measure of coverage. Second, we estimate signal strength around towers built in 2015 when the epidemic was practically over. We limit the analysis to districts that did not have any coverage in 2013 so that we do not pick up any effect from the coverage footprint in 2013. In all, that leaves us with 383 villages, of which 77, that did not have coverage in 2013, obtained coverage from the towers built in 2015. Given the low number of observations, we cannot employ a RD design that limits observations within a bandwidth. Therefore, we run a regression of EVD likelihood in 2013 on the measure of coverage in 2015 for all villages within districts that did not have any coverage in 2013. We should find no effect as this is essentially a placebo test that uses a boundary that did not exist in 2013. Column (6) presents no statistically significant effect of 2015 coverage on the likelihood of EVD in the previous years.

5.3.2. RD estimates using geographic and technological distance

The results in Section 5.2 compare villages within a close window of the signal strength cutoff. This section adds geographic distance as another dimension by which villages are compared. Specifically, we construct the geographic distance from each village to its closest point on the two-dimensional coverage boundary. We then restrict the analysis to villages that are both, technologically (i.e., within a close bandwidth of the signal strength cutoff) and physically (i.e., within a close bandwidth of the spatial coverage boundary) close to each other.

Appendix Table B.8 presents the results for the baseline RD (columns 1–3) and the RD specification that adds controls (columns 4–6). Columns (1) and (4) present estimates without putting any restrictions on physical distance to the coverage boundary and thus replicate columns (1) and (2) of Table 2. Instead, columns 2 and 5, and columns 3 and 4, consider a physical distance of 2 and 4 km, respectively. The results remain quantitatively similar to the main results (columns 1 and 4) after restricting the analysis to villages that are geographically close to each other. This is not surprising since there is a significant correlation between geographic and technological distance, i.e., villages close to each other in terms of signal strength tend to be geographically close.

5.3.3. Sensitivity of results to near-cutoff observations

Appendix Fig. B.8 examines the sensitivity of our main results to observations near the signal strength cutoff. We start by performing 250 replications where, within each replication, we randomly drop 5% of

Table 5
Controlling for "free-space" signal strength and falsification tests.

	Dep. Variable = $\mathbb{1}$\{Number of Reported EVD cases > 0\}					
	Controlling for Free-space signal strength				Falsification tests	
	(1)	(2)	(3)	(4)	(5)	(6)
Coverage	−0.087**	−0.084*	−0.077**	−0.081**		
	(0.038)	(0.041)	(0.037)	(0.040)		
Free-space coverage	−0.013		−0.044*	−0.045*	−0.003	
	(0.020)		(0.024)	(0.026)	(0.040)	
Free-space signal strength		0.000	−0.000	−0.000	0.002	
		(0.001)	(0.001)	(0.001)	(0.006)	
Free-space coverage × Free-space signal strength			0.003	0.003	−0.008	
			(0.002)	(0.002)	(0.007)	
Coverage (2015)						0.084
						(0.112)
Signal strength (2015)						0.001
						(0.002)
Mean outside coverage	0.093	0.101	0.094	0.093	0.094	0.057
Bandwidth (dBm)	8.35	7.99	9.14	8.93	9.36	–
Observations	1547	1352	1720	1528	1540	383
Districts	83	81	84	83	84	53
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Topography Polynomial	No	No	No	Yes	No	No

Notes: Columns (1)–(5) present estimates using a local linear regression specification of Eq. (1). Column (1) adds controls for coverage under free-space. Column (2) adds controls for free-space signal strength. Columns (3) and (4) add the full interaction of free-space coverage and strength. Column (5) estimates the local linear regression specification of Eq. (1) using free-space coverage. Column (6) estimates the effect of coverage in 2015 excluding districts that had coverage in 2013, including district fixed-effects. All specifications include controls for elevation, slope, average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Optimal bandwidth in Columns (1)–(5) chosen as in Calonico et al. (2014). Column (4) adds topography polynomial: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. It also includes distance to tower as additional control. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

the observations within a 1 dBm window of the cutoff and estimate our model (Eq. (1)) using the restricted sample. We then repeat this exercise dropping 10, 15, …, 95% of the sample within the 1 dBm window. Appendix Fig. B.8 plots the set of 250 estimated RD coefficients for each of the dropped sample levels (5 to 95%) along with 95% confidence intervals for each set of estimates. As we drop a larger share of the sample, the estimates tend to get smaller in magnitude relative to our baseline RD estimate presented in column (1) of Table 2. However, the estimates are consistently negative and significantly different from zero even after dropping up to 95% of the observations near to the cutoff. This suggests that, although observations nearest the cutoff understandably influence our results to some degree, they are not the sole drivers of the negative effect of coverage on the likelihood of having EVD.

5.3.4. "Walking" to coverage boundary

Recall that our analysis estimates intention-to-treat (ITT) effects. In principle, individuals in a village without coverage can own phones and travel to nearby villages with coverage to take advantage of the technology. If that is the case, then our main estimates would be attenuated since villages assigned to control (non-coverage) areas would arguably receive some degree of treatment (coverage).[28]

For all villages without cellphone coverage, we calculate travel distance to the nearest village with coverage. We rely on the Open Source Routing Machine (OSRM) engine for calculations and OpenStreetMap as the base map.[29] Appendix Table B.9 presents the average calculated travel distances. We were able to estimate travel distances for about 59% of the full sample of villages without coverage (7,830 villages). Since the calculated travel distances for 2,816 villages (36%) came out shorter than the Euclidean distance, we drop these observations for this analysis.[30] Overall, we find that the average travel distance from villages without coverage to the nearest village with coverage is about 26 kilometers (column 1). When we reduce our analysis to villages closer to our main estimation sample–i.e., within 10 dBm of signal strength threshold (column 3)–the average distance is reduced to about 5.7 kilometers with a standard deviation of about 10 kilometers.

We cannot directly observe whether individuals in villages without coverage traveled to nearby villages with coverage. However, we can assess how our baseline RD estimates respond to restricting the comparison sample to villages without coverage spatially close to villages with coverage, where individuals could have traveled to make a phone call. By restricting the sample to villages without coverage within 2, 4, 6, 8, and 10 kilometers of a village with coverage, we find no evidence that being spatially close to a village with coverage significantly attenuates our estimates (Appendix Table B.10).

The fact that we find little evidence of people "walking" to coverage is contextually plausible for several reasons. First, there were significant restrictions on mobility. All affected counties were under full quarantine and strict curfews at the height of the epidemic (IFRC, 2014). Furthermore, villagers informally enforced strict protocols to limit outsiders from entering their villages (Ruble, 2015). Second, cellphones are a technology that is inherently dependent on a network effect. If other members of the network do not have access, then there is little incentive for a single member to select into the technology

28 Note that this is a problem only if non-coverage villages that are physically near coverage villages also happen to fall within the signal strength bandwidth used in the analysis. However, since physical distance and signal strength across villages are correlated, this can still attenuate our main estimates.

29 We use "georoute" package (Weber and Péclat, 2017) for travel distance calculations.

30 The calculated travel distance can be shorter than the Euclidean distance if a village is relatively far from, or not directly connected to, the road network used in the calculation. We use the best available road segments but this leads to an arbitrarily short travel distance for part of the sample. Also note that the travel distances for 375 non-coverage villages (5%) could not be calculated.

(e.g., incur costs of purchasing handset, call/data plan, SIM cards, etc.). To explore this possibility, Appendix Fig. B.9 presents a histogram using 5 percentage point bins of the share of villages within a clan that have access to coverage. Clans can serve as a proxy for a village's network based on historical links between villages within clans. A key takeaway is that in most clans (about 63%) only a small fraction of villages actually have coverage (less than 5%). This suggests that individuals within these clans, even if they live near a coverage village, might be less willing to incur the costs of the technology until a larger mass of villages within their own clans receives coverage. Lastly, financial uncertainties inherent to epidemics can lead individuals in villages without coverage to delay investments in the technology (handset, call/data plans, SIM cards, etc.).

5.3.5. Spatial standard errors

We also experiment with Conley (1999) spatially clustered standard errors using 10, 20, and 50 kilometer distance cutoffs and a linear decay in the correlation weights (Appendix Table B.11). We find that these standard errors tend to be smaller or very similar to the district-level clustered standard errors which is our preferred specification.

6. Channels of impact

This section explores potential channels underlying the relationship between cellphone coverage and EVD outcomes. Cellphone access can better connect individuals to preventive care resources (e.g., prevention education, hygiene practices guidance), as well as facilitate access to treatment resources (e.g., taking sick and dead people, ambulances). It is possible that individuals living in villages with cellphone coverage not only were more likely to directly call to ask for more information or an ambulance (e.g., a hotline number was set-up by the government to this purpose), but also to report cases and better coordinate with political representatives to lobby for additional resources with the central government (Maffioli, 2021). We refer to the former as the *preventive care* channel and the latter as the *treatment care* channel.

Cellphone coverage also enables individuals to more efficiently interact with a potentially larger network of friends and family. This can improve within-network collective action during emergencies (Hampton et al., 2011, Pew Research Center, 2011, Pew Research Center, 2019, Blumenstock et al., 2016). For example, individuals with available cellphone coverage can more easily share information, rely on their network if they need care, mobile transfers, or if they want to gather family for funerals, one of the main factors contributing to EVD spread (Alexander et al., 2015, Fallah et al., 2015). The overall effect on spread (or containment) is not clear: within-network person-to-person interactions may decrease if cellphone interactions are a good substitute to in-person interactions. On the other hand, an unintended consequence could be the increase in the number of person-to-person interactions–potentially with sick (or dead) members within the network–via improved coordination and attendance of super-spreader events such as funerals and burials. This is especially concerning during the early stages of the epidemic when knowledge about transmission was low and government-provided alternatives to in-network care were scarce. We refer to this mechanism as the *network* channel.

Cellphone coverage can also decrease the cost of information access. This can increase exposure to outbreak-related information (e.g., latest recommendations, preventive measures, treatment resources, etc.) while also increasing exposure to misinformation. This potentially dampens any benefits from increased access to quality information.[31] We refer to this mechanism as the *information* channel. In principle, information sharing within a network (another potential component that raises the complexity of the network channel described above) can also be a source of (mis)information. Therefore, we think of the information channel here as capturing top-down information (e.g., information provided by the government, NGOs, and other organizations) while within-network information sharing can be thought as an additional component of the network channel.

Overall, we expect the preventive and treatment care channels to decrease the likelihood and intensity of EVD (containment). The effect of the network and information channels will depend on the degree to which cellphone interactions substitute for in-person interactions (network channel) and on the type of information provided—quality information *versus* misinformation (information channel).

Finally, it is also possible that access to cellphones generated benefits pre-pandemic for individuals living in villages with coverage (e.g. improvements in market efficiency (Jensen, 2007; Aker, 2010), literacy (Aker et al., 2012; Aker and Ksoll, 2018), access to mobile banking (Jack et al., 2010; Jack and Suri, 2011), and risk sharing (Jack and Suri, 2014; Blumenstock et al., 2016)). This may enable individuals with access to coverage to better cope with the epidemic shock and thus leading potentially to lower incidence within these areas. Yet, our evidence suggests that households in villages at the margin of coverage (Appendix Table B.2) are similar across the signal strength cutoff in terms of socio-demographic characteristics including wealth, job opportunities, and education. Overall, this suggests that the income channel may not play a major role in reducing EVD likelihood near the coverage margin.

6.1. Preventive and treatment care channels

We explore the preventive and treatment care channels using the survey data described in Section 3.3. Since the survey was conducted using cellphones, all respondents potentially have access to coverage and thus we cannot observe outcomes in locations without coverage. As a result, we cannot estimate the effect at the extensive margin as we did in our RD results. Instead, we use signal strength as our treatment variable controlling for "free-space" signal strength as in Eq. (3). This allows identifying the effect of signal strength based on minor differences in topography between villages since we are comparing villages that in the absence of topography (i.e., in "free-space") would have received the same signal strength. Specifically, we estimate the following model:

$$Y_{ij} = \alpha + \beta R_{ij} + \gamma R_{ij}^{Free} + \mathbf{X}_{ij}\delta + \lambda_j + \epsilon_{ij} \tag{4}$$

where Y_{ij} refers to a preventive or treatment care outcome for an individual in village i in county j. R_{ij} is the signal strength received in the village and R_{ij}^{Free} is the free-space signal strength that village i would have received in the absence of topography. $\mathbf{X}_{ij}$ is a vector of individual level controls that include sex, age, urban status, secondary education, religion (Christian), and ethnicity (Kpelle, Bassa, and other). λ_j is a county fixed effect.

We construct self-reported measures of access to *preventive* and *treatment* care. We use a set of outcomes that capture whether cellphone access allowed communities to be more exposed to either care efforts. The survey asked whether during the Ebola epidemic anyone from the government, health workers, NGO, or international organizations came to the community, and if so, what was the purpose of the visit: explain what EVD was, hold hygiene meetings, distribute prevention material, do contact tracing, explain how to conduct safe burials, take sick people or dead bodies. Respondents were also asked directly whether a taskforce came to their villages as the taskforce was directly set-up to bring information about EVD. In addition, we asked respondents to report how long it took, on average, for ambulances to get to their communities, whether the ambulance came very late, or never came. Finally, we use publicly available data on the location of CCCs to

[31] The spread of misinformation during epidemics is significant (Venkatraman et al., 2016; Ortiz-Martínez and Jiménez-Arcia, 2017; Carey et al., 2020). The 2014 West Africa Ebola epidemic was no exception (Krishna and Thompson, 2019). Refer to World Health Organization (2014), Oyeyemi et al. (2014), Onyeonoro et al. (2015), Allgaier and Svalastog (2015), Pathak et al. (2015) and Roberts et al. (2017) for information on how misinformation can disrupt epidemic response.

Table 6
Effect of coverage on preventive care.

	Contact tracing (1)	Explain EVD (2)	Hygiene meetings (3)	Bring info (Taskforce) (4)	Explain burials (5)	Prevention material (6)	Preventive care index (7)
Signal strength (SD)	−0.031	−0.030*	0.020	0.050*	−0.027	0.044	0.000
	(0.021)	(0.016)	(0.035)	(0.027)	(0.025)	(0.029)	(0.035)
Mean	0.135	0.884	0.457	0.229	0.159	0.714	0.009
Observations	2064	2064	2064	2064	2064	2064	2064
Counties	15	15	15	15	15	15	15
Villages	536	536	536	536	536	536	536

Notes: Results present estimates of β in Eq. (4) using the variables specified in each column as outcomes. Sample are respondents in the survey collected by authors (Section 3.3). Outcome variables are indicators for whether someone from the government, health workers, local or international NGOs came to the village to: do contact tracing (column (1), explain what EVD was (column (2)), teach EVD-related hygiene practices (column (3)), whether a community taskforce came to share information on EVD and how to prevent it (column (4)), explain how to conduct safe burials (column (5)), and bring preventive material such as buckets for chlorine (column 6). The care index (column 7) is constructed following (Kling et al., 2007). All specifications include controls for "free-space" signal strength, sex, age, urban, and education level, and categories for Kpelle, Bassa, and other tribes. County fixed effects are included. Standard errors clustered at the village level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

construct a variable equal to 1 if the village had a CCC within a 10 km radius.[32] We classify preventive and treatment care by using measures that allowed for limiting EVD cases from happening or that addressed the needs of individuals affected by the disease, respectively.

Specifically, for the *preventive care* channel (Table 6), we study whether individuals report that someone (from government, health workers, local or international NGOs) came to their village to do contact tracing (column (1), explain what EVD was (column (2)), to teach EVD-related hygiene practices (column (3)), whether a community taskforce came to share information on EVD and how to prevent it (column (4)), or someone came to explain how to conduct safe burials (column (5)) and brought preventive material such as buckets for chlorine (column 6).[33] For the *treatment care* channel (Table 7), we explore indicators for whether someone came to take sick people or dead bodies (columns (1)–(2)). We further examine if individuals report that ambulances arrived within 4 h after being requested (column (3)), and whether a CCC was built within 10 kilometers of the village (column (4)).[34]

Table 6 presents results for the *preventive care* channel. We find mixed evidence: a one standard deviation increase in signal strength is associated with a statistically significant 3 percentage point drop in the likelihood that someone came to their village to explain EVD (column (2)) and a 5 percentage point increase in the likelihood that a community taskforce came to teach preventive measures (column (4)). This might be consistent with the fact that in areas with cellphone coverage simple information about the virus can be channeled through cellphones or radio, instead of using and sending already limited personnel. On the other hand, even in areas with higher coverage, it might still be necessary to send health workers to explain preventive measures or to build trust with the community. Estimates on other outcomes have mixed directions and lack statistical significance. Thus, we cannot fully exclude that improving access to preventive care through better quality coverage could have also contributed to reduced likelihood of the disease.

Table 7 explores the *treatment care* channel. We find evidence that a one standard deviation increase in signal strength is associated with a statistically significant 5 percentage point increase in the likelihood that ambulances arrived on time when requested (column (3)), and 22.8 percentage point increase in the likelihood that a CCC was placed near their village (column (4)). While we cannot test this, it is possible that higher signal strength allowed individuals to report cases and communicate local needs to their political representatives who could lobby for the allocation of resources with the MOH at the central level (Maffioli, 2021). We do not find statistically significant effects on whether someone came to take sick or dead people (columns (1) and (2)).

Comparing the results in Tables 6 and 7, we find suggestive evidence that treatment care measures may play a bigger role in explaining how coverage helped contain EVD. This is also reflected when we explore summary indexes of each of the two care channels constructed following (Kling et al., 2007). The index of all preventive care outcomes results in a null coverage effect (Table 6, column (7)), while we find a positive and statistically significant effect in the case of the combined treatment care outcomes (Table 7, column (5)). These results are contextually plausible. In the midst of a crisis, the impact of on-time ambulance service or a CCC near a village is likely large and immediately observed. The effects of preventive care may take longer to materialize as they entail behavioral changes.

We note that although treatment care measures cannot directly explain the results presented in Table 2 (i.e., treatment is unlikely to affect EVD at the extensive margin), a set of additional results point to treatment care playing a significant role. Recall that conditional on having an EVD case, villages with coverage are less likely to experience an additional case, a suspected death, as well as a lower number of additional cases and deaths (Table 4), and have a higher survival probability (i.e., report no additional cases or deaths) in any given month after experiencing a case or a death (panels (b) and (d) of Fig. 5). Villages with coverage also report a lower number of cases overall (Table B.4), and a lower number of suspected deaths (Table B.6) than villages without coverage. Different types of treatment care such as receiving ambulances on time, taking sick people or dead bodies, and CCCs nearby can explain these findings by limiting exposure after a first EVD case. However, note that treatment care does not exclusively explain these additional results, preventive care and information access can also play a part.

[32] It may be the case that some CCCs fall in non-coverage areas. However, the main purpose of this exercise is to assess whether villages with stronger coverage are more likely to be located near treatment centers regardless of whether the actual CCC falls in a coverage area or not.

[33] Compared to the COVID19 pandemic for example, the government or other institutions did not provide strong direction of staying indoors or avoiding contacts with infected people, thus we did not measure those outcomes in the survey.

[34] In the survey data, 4 h is the median time taken from an ambulance to reach the village of destination from the dispatch.

Table 7
Effect of coverage on treatment care.

	Take sick (1)	Take dead (2)	Ambulance on-time (3)	CCCs within 10 km (4)	Treatment care index (5)
Signal strength (SD)	−0.024	0.018	0.050**	0.228***	0.085*
	(0.019)	(0.016)	(0.021)	(0.081)	(0.048)
Mean	0.083	0.057	0.858	0.664	0.002
Observations	2064	2064	2064	2064	2064
Counties	15	15	15	15	15
Villages	536	536	536	536	536

Notes: Results present estimates of β in Eq. (4) using the variables specified in each column as outcomes. Sample are respondents in the survey collected by authors (Section 3.3). Outcome variables are indicators for whether: someone came to take sick people (column 1), someone came to take dead bodies (column 2), ambulances arrived within 4 h after being requested (column 3), and a CCC was built within 10 kilometers of the village (column 4). The care index (column 5) is constructed following (Kling et al., 2007). All specifications include controls for "free-space" signal strength, sex, age, urban, and education level, and categories for Kpelle, Bassa, and other tribes. County fixed effects are included. Standard errors clustered at the village level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table 8
Likelihood of EVD within the network.

	Dep. Variable = $\mathbb{1}\{\text{Number of Reported EVD cases} > 0\}$					
	Network: Call Detail Records (CDR)					
	Top tower (1)	Top 5 towers (2)	50% call share (3)	75% call share (4)	90% call share (5)	Network: Clan (6)
$EVD_{j,t-1}$	0.040***	0.031***	0.041***	0.018**	0.009**	0.034
	(0.007)	(0.009)	(0.007)	(0.008)	(0.004)	(0.028)
$Match_{ij} \times EVD_{j,t-1}$	0.006	0.006	0.011	0.011	0.0085	−0.003
	(0.033)	(0.024)	(0.036)	(0.016)	(0.006)	(0.045)
Mean	0.004	0.004	0.004	0.004	0.004	0.007
Observations	18 864	18 864	18 864	18 864	18 864	19 372
Villages	9432	9432	9432	9432	9432	9686
Clusters	96	96	96	96	96	631
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Quarter FE	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No

Notes: Dependent variable is a dummy indicator for whether there is at least one EVD case in village i in quarter t. $EVD_{j,t-1}$ equals 1 if there is at least one EVD case within village i's network $j(i)$ in the past quarter. $Match_{ij}$ equals 1 if there is a cellphone coverage match between village i and any of the affected villages in the past quarter within village i's network (i.e., village i has coverage and at least one EVD-affected village within i's network also has coverage). Columns 1–5 define village i's network using call detail records (CDR) data between village i's cell tower and other cell towers across Liberia. Columns 1 and 2 define i's network as all villages within the service area of the top and top-5 towers receiving calls originating from i's tower. Columns 3–5 define i's network as all villages within the service area of cell towers receiving 50%, 75%, and 90%, respectively, of all calls originating from i's tower. CDR data obtained from Cellcom Liberia. Column 6 defines i's network as all villages within the same clan as village i. Standard errors clustered at the network level. All specifications include village and quarter fixed effects. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table 9
Effect of coverage on (Mis)information.

	Informed (1)	Misinformed (2)	Gov (3)	Foreign (4)	Other (5)
Signal strength (SD)	0.054*	−0.038	0.002	−0.006	−0.026
	(0.032)	(0.034)	(0.023)	(0.007)	(0.023)
Mean	0.442	0.320	0.173	0.023	0.128
Observations	2064	2064	2064	2064	2064
Counties	15	15	15	15	15
Villages	536	536	536	536	536

Notes: Results present estimates of β in Eq. (4) using the variables specified in each column as outcomes. Sample are respondents in the survey collected by authors (Section 3.3). All specifications include controls for "free-space" signal strength, sex, age, urban, and education level, and categories for Kpelle, Bassa, and other tribes. County fixed effects are included. Standard errors clustered at the village level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

6.2. Network channel

We are unable to directly observe an individual's or village's social network. However, we use two different approaches to potentially assess a village's network.

First, we use CDR data to define each village's network using call patterns between the cell tower servicing that village and other cell towers across Liberia. Broadly speaking, we consider two villages to be connected if there is significant call traffic between their corresponding cell towers (Appendix Fig. B.3 shows for example that the core network for Ganta city likely lies in villages to the North and East of this tower). Specifically, for each tower k, we calculate the number of outgoing calls from that tower to all other towers across Liberia (including itself). We use this to formulate several definitions of a network: (1) all villages within the catchment areas of tower k plus the tower receiving the highest number of calls from k; (2) all villages within the catchment areas of tower k plus the towers receiving the top five highest number of calls from k; (3) all villages within the catchment areas of tower k plus the towers receiving 50% of all calls from k; (4) all villages within the catchment areas of tower k plus the towers receiving 75% of all calls from k; and (5) all villages within the catchment areas of tower k plus the towers receiving 90% of all calls from k.[35]

Second, we use historic Liberian clans to define a village's closest social network. Officially, clans in Liberia are a third-tier administrative division but correspond to historical tribal chiefdoms that were merged into the state throughout Liberian history, with chiefs simply assuming the role of agents of the central government (Nyei, 2014). Therefore, villages within the same clan are more likely to be socially interconnected. Based on the Census data (LISGIS, 2008), there are 631 clans in Liberia with an average of about 38 villages within each clan.

To assess whether the network channel is relevant, we estimate the following empirical model:

$$EVD_{ijt} = \alpha + \gamma EVD_{j,t-1} + \delta Match_{ij} \times EVD_{j,t-1} + \lambda_i + \nu_t + \epsilon_{ijt} \quad (5)$$

where j indexes a network (CDR or clan-based definition). EVD_{ijt} is an indicator for whether there is an EVD case in village i of network j, in quarter t. $EVD_{j,t-1}$ is an indicator for whether there is at least one EVD case within village i's network j in the previous quarter. We define $Match_{ij}$ as a "coverage match" between village i and any of the affected villages in the past quarter within village i's network, i.e., $Match_{ij}$ equals 1 if village i has cellphone coverage and at least one of the villages within i's network j having an EVD case in the past quarter also has coverage. λ_i is a village fixed effect. ν_t is a quarter fixed effect. The village fixed effects λ_i account for any time-invariant unobservables that may lead to endogenous selection of villages into EVD.[36]

In the presence of a network channel, we should expect the likelihood of an EVD case in village i in quarter t to decrease or increase if there is a "coverage match" between village i and a village within i's network with an EVD case at time $t-1$. Thus, we should expect coefficient δ to be nonzero if the network channel is important. We restrict the analysis to the early-outbreak period when access to centralized relief efforts was limited and individuals likely relied on their network for help.

Table 8 presents the results. First, we find significant evidence of contagion within networks: Villages where another village within their network had an EVD case in the previous quarter are significantly more likely to report an EVD case in the current quarter. Although the evidence is stronger when using the CDR-defined networks (columns (1)–(5)), there is still some evidence of contagion within clans (column (6)). Not surprisingly, the magnitude of the contagion effect diminishes as we expand our definition of a network (0.0399 in column 1 *versus* 0.0310 in column 2, and moving from column 3 to column 5). This is expected if one considers that the likelihood of an interaction (and thus EVD transmission) is higher within a tighter definition of a network compared to a wider one. More importantly, we find no conclusive evidence of a network channel: for various definitions of a network, the coefficient estimates on the interaction term are small and statistically insignificant. This suggests that although contagion within networks is prevalent, it does not significantly increase if two villages happen to be connected. Note that this does not necessarily imply that the network effect is trivial. It can very well be that any network benefits (e.g. collective action, transfers, etc.) get canceled by negative network externalities (e.g., contagion due to increased person-to-person interactions).

6.3. Information channel

We finally test whether access to (mis)information varies with cellphone coverage. As part of our survey, we asked individuals whether they received daily information related to EVD (transmission, prevention, treatment). The survey also documented each individual's beliefs about the origin of the outbreak. We use this latter information to classify individuals as *informed* if they reported that the epidemic originated at the border with Guinea and Sierra Leone or with traders from these areas. We classify individuals as *misinformed* if they reported that the origin of the outbreak was other institutions, distinguishing between government, foreign organizations/people (white people, UNMIL, foreign NGOs) and others (god, witches, the Fula, Mandingo, or Kissi) (see Maffioli and Gonzalez (2022) for more information on the classification).

We find almost no variation in the likelihood of receiving EVD-related information on a daily basis in our sample: more than 93% of respondents reports receiving information and this proportion does not change with coverage status. Table 9 presents the results on the degree of knowledge about the origin of the epidemic. We find some evidence that individuals with higher coverage are more likely to be informed (column (1)), but results are statistically insignificant if we restrict our measure of informed individuals only to those who reported that the epidemic originated at the border with Guinea, the actual origin of the outbreak. We do not find any evidence that individuals are more likely to be misinformed as signal strength increases (columns (2)–(4)).

The fact that we find no large impacts of cellphone coverage on access to EVD-related information is not surprising in our context. First, access and use of radio in Liberia is much more ubiquitous than access to cellphone coverage (International Media Support, 2007). In our sur-

[35] As an example, Appendix Fig. B.10 depicts the networks around the cell tower located in Ganta city, Nimba county, using these different definitions. Appendix Fig. B.11 also shows that most outgoing calls go to relatively few towers. Notice in panel (a) that for a given tower, 10 to 60% of all outgoing calls go to only one tower (the top called tower). However, only 0 to 5% of outgoing calls go to the 5th ranked tower. This indicates that the networks are generally small in terms of geographic extent. We use Voronoi cells to draw catchment areas around each of the 152 unique cell tower locations. Voronoi cells define the area for which any village within the cell is closest to that cell tower. Refer to Blumenstock et al. (2016) for earlier work using Voronoi cells to denote the catchment area of cell towers. Note that being within a tower's catchment area does not necessarily mean that the location gets coverage. It simply means that the location is closest to that specific tower.

[36] Given data sparsity in terms of EVD cases at the village-month level, we decide to aggregate cases to the village-quarter level. Note that we cannot estimate a coefficient for the $Match_{ij}$ variable by itself as it is collinear to the village fixed effect since coverage status does not change within our sample period.

vey 80% of individuals report owning a radio, while in a national media survey 94% of urban dwellers and 91% of rural dwellers reported listening to the radio in the past week (Montez, 2010). Radio coverage is also widespread with several radio stations broadcasting at the national level and local community stations covering each county (International Media Support, 2007). In such setting, the informational returns of cellphone access can be trivial.

Second, during the period of study the predominant cellphone technology available in Liberia was 2G (i.e., no data/Internet access through the phone). For instance, starting in 2014 only two towers in Monrovia and one tower near Roberts International Airport had 3G capabilities (i.e., Internet browsing ability) and none had 4G capabilities (broadband Internet, video streaming).[37] Thus, individuals at the time did not have access to video, social media, and other Internet content that can potentially be fertile ground for misinformation and EVD-related conspiracy theories. Therefore, it is not surprising that cellphone coverage within our context is not strongly associated with being misinformed.

Note that our data and results cannot be used to assess whether information was an important mechanism in containing EVD. It is plausible that information disseminated through radio, cellphones, and billboards played a crucial role in disease containment. Our results simply document that people with more coverage did not necessarily receive more information than people with less coverage.

7. Conclusion

Combining novel data on cellphone tower locations and Ebola cases in Liberia, we show that cellphone coverage contains the spread of the Ebola Virus Disease (EVD). Specifically, comparing villages at the margin of the signal strength threshold, we find that having access to cellphone coverage leads to a 10.8 percentage point reduction in the likelihood that a village has an EVD case. Using novel survey data collected after the epidemic, we assess the importance of several channels that may explain the observed relationship between cellphone coverage and epidemic containment. We provide evidence that the negative relationship between cellphone coverage and EVD is likely explained by facilitating access to treatment care. This result is plausible since, in the short run, the returns to timely ambulance service or care centers are likely higher and immediately realized. On the other hand, the effects of prevention may take longer to materialize as they essentially entail a change in health behavior. However, we cannot fully exclude that improving access to preventive care or information could have also contributed to containment.

Infectious disease outbreaks are still a major burden to low and middle-income countries (Holmes et al., 2017), and extreme events, such as health epidemics, are expected to remain a worldwide threat United Nations Office for Disaster Risk Reduction (2015). For instance, we are currently in the midst of a worldwide Coronavirus epidemic and a resurgence of Ebola in several African countries. Even though their onset might be hard to predict, the ultimate human and economic costs could be mitigated through appropriate governmental actions.

Our findings show how something as simple and ubiquitous as a cellphone can have positive externalities on health outcomes by allowing communities to better access health care treatment resources in times of crisis. From a policy perspective, it is fundamental for governmental stakeholders to know the relative effectiveness of potential tools, such as cellphones, in mitigating the effects of infectious diseases. Our results can help guide policymakers in choosing more efficient allocations of limited funds. For example, in normal times, longer term policies such as investments in the expansion of cellphone coverage to remote areas and information campaigns to prompt changes in preventive health behavior can be fruitful. During times of crisis, policymakers should take advantage of increased cellphone coverage to implement shorter-run, emergency policies (e.g., hotlines) and measures that enhance access to treatment care and thus improve the effectiveness of the response. However, further research should explore specific interventions that directly take advantage of this technology and the effects of these tools in the context of other health epidemics.

CRediT authorship contribution statement

Robert Gonzalez: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review & editing. **Elisa M. Maffioli:** Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Writing – original draft, Writing – review & editing.

Data availability

Data will be made available on request.

Appendix A. Estimation of the Irregular Terrain Model (ITM)

The estimation of cellphone coverage involved two major steps. First, we estimated the Irregular Terrain Model (ITM) to obtain signal strength on the ground for each individual tower using the CloudRF API (available at: https://api.cloudrf.com). Table A.1 lays out the specific inputs used in the estimation of the ITM. For each tower, the output from the cloudRF API is an ArcGIS-readable polygon shapefile containing as attribute the signal received on the ground for different bands of the polygon file. For each tower, the estimated signal strength ranged between −50 and −140 dBm. We clarify that the cloudRF API actually gives the absolute value of the signal strength, thus technically the output ranged between 50 and 140.

Second, the polygon shapefiles obtained from the ITM were processed in Arcmap 10.7 using the ArcPy package. The processing involved three steps. First, we projected the shapefiles using the *Project* tool to an appropriate coordinate system (UTM 29N) for analysis. Second, we converted the shapefiles to raster files with cell size of 50X50 meters using the *Polygon to Raster* tool. Third, we used the *Mosaic to Raster* tool to combine the raster files for each tower into one single raster file for all of Liberia (depicted in Fig. 1). For raster cells where there was overlap in coverage between several towers, we used the "Minimum" rule to assign the signal strength to that particular cell. This assigns the highest signal strength available to that cell since the lower the estimated absolute value dBm, the higher the signal strength.

[37] These observations are made by the authors using the GSMA Collins Bartholomew Coverage Explorer (GSMA, 2014) which provides the type of coverage (2G,3G,4G) for GSM networks across the world.

Table A.1
Main variables and parameters of Irregular Terrain Model (ITM).

Model variables	Description	Parameters
Transmitter (cell tower) characteristics		
Transmitter power (txw)	Transmission power (Watts)	5
Frequency (frq)	Radio wave frequency (MHz)	900
Latitude (lat)	Latitude of cell tower	e.g., 6.27069
Longitude (lon)	Longitude of cell tower	e.g., −10.73158
Transmitter height (txh)	Height of cell tower above ground (meters)	30
Radius (rad)	Maximum coverage radius (kilometers)	20
Antenna gain (txg)	Transmitter antenna gain (dBi)	2.14
Antenna (ant)	Antenna type	39 (omnidirectional)
Azimuth (azi)	Antenna azimuth angle (degrees)	0°
Antenna tilt (tlt)	Antenna vertical angle (degrees)	0°
Antenna polarization (pol)	Vertical/Horizontal	v (vertical)
Receiver (handheld device) characteristics		
Receiver sensitivity (rxs)	Minimum power received threshold (dBm)	−140
Receiver height (rxh)	Receiver height above ground (meters)	1.5
Antenna gain (rxg)	Receiver antenna gain (dBi)	2.14
Geographic characteristics		
Resolution (res)	Topographic model	30 (DSM30)
Clutter (clh)	Consider clutter (trees, buildings, etc.)	Yes
Climate (cli)	1: Equatorial	3
	2: Continental Subtropical	
	3: Maritime Subtropical	
	4: Desert	
	5: Continental Temperate	
	6: Maritime Temperate, over land	
	7: Maritime Temperate, over sea	
Terrain conductivity (ter)	Salt water : 80	15 (farmland)
	Fresh water : 80	
	Good ground : 25	
	Marshy land : 12	
	Farmland, forest : 15	
	Average ground : 15	
	Mountain, sand : 13	
	City : 5	
	Poor ground : 4	
Other inputs		
Model (pm)	Propagation model	1 (ITM)
Model subtype (pe)	Conservative, Average, Optimistic	2 (Average)
Measure (out)	Measurement units	2 (dBm)
Engine (eng)	API Engine	2 (SLEIPNIR)
Knife edge diffraction (ked)	Yes (= 1) or No (= 0)	0
Color scheme (col)	Color scheme	9 (Grayscale/GIS)

Notes: Refer to CloudRF API (available at: https://api.cloudrf.com for more information. CloudRF API code within parenthesis in "Model variables" column.

Appendix B. Additional figures and tables

See Figs. B.1–B.11 and Tables B.1–B.11

Table B.1
Determinants of coverage.

	Dep. variable: Signal strength (dBm)		
	Full sample (1)	Within 20 dBm (2)	Within 10 dBm (3)
Topographic controls			
Elevation (km)	0.010	0.008	0.024
	(0.010)	(0.009)	(0.015)
Slope (%)	2.380*	2.124*	−2.125
	(1.421)	(1.070)	(1.683)
Demographic controls			
Household size (cluster)	−0.656	0.240	−0.349
	(1.098)	(0.865)	(0.863)
Married (cluster)	−19.001*	−15.356	6.448
	(11.362)	(12.708)	(11.520)
Christian (cluster)	18.497	2.602	44.732
	(22.072)	(21.862)	(28.783)
Muslim (cluster)	31.975	5.327	42.642
	(22.235)	(21.543)	(26.031)
Other (cluster)	−97.375	41.542	7.111
	(91.214)	(60.378)	(56.018)
Economic controls			
Primary education (cluster)	−18.307*	−26.451**	5.856
	(10.371)	(12.283)	(9.634)
Secondary education (cluster)	−0.253	4.469	8.041
	(11.734)	(13.591)	(10.907)
Own house (cluster)	−10.402	−0.121	6.316
	(8.759)	(9.797)	(8.903)
House condition: good (cluster)	−17.310	3.191	−13.268
	(12.813)	(11.529)	(14.893)
Asset ownership (cluster)	33.152	13.276	−1.417
	(22.730)	(26.979)	(24.229)
Distance to Monrovia (km)	0.012	0.011	−0.033
	(0.012)	(0.011)	(0.022)
Distance to closest city (km)	−0.187***	−0.056	−0.008
	(0.065)	(0.070)	(0.064)
Additional controls			
Rural (cluster)	−6.403	2.760	−2.858
	(4.109)	(3.582)	(3.325)
Wealth index (DHS, cluster)	5.739	−2.566	5.121
	(5.084)	(4.845)	(5.933)
Occupation wage (cluster)	17.359**	5.700	−4.432
	(8.098)	(9.170)	(11.189)
Occupation self-employed (cluster)	−12.244*	−14.390**	6.317
	(6.908)	(5.961)	(8.680)
Occupation professional (cluster)	−46.456	−11.138	−3.997
	(39.527)	(51.551)	(34.298)
Observations	234.00	106.00	53.00
Topographic controls	0.11	0.04	0.27
Demographic controls	0.02	0.67	0.48
Economic controls	0.00	0.00	0.11

Notes: Robust standard errors. *, **, and *** indicate 10, 5, and 1 percent significance, respectively. Sources of data: ALOS Global Digital Surface Model; Demographic Health Survey (DHS, 2013). 322 DHS clusters are linked to the closest village by GPS coordinates. Geographical controls and distances are from the closest 244 linked villages. Other controls are collapsed at the DHS cluster level. Compared to Table 1, we excluded population and female population because the DHS clusters are artificially created for the purpose of the DHS surveys. In addition, Kpelle and Bassa do not have enough variation at the DHS cluster level. House condition good is defined as a mean between improved floor, wall and roof material based on WHO definitions. Asset ownership is defined as a mean of owning electricity, radio, phone, bank account, refrigerator, vehicle).

Table B.2
Summary statistics by coverage status (villages).

	Full sample			Within 20 dBm			Within 10 dBm		
	Inside (1)	Outside (2)	S.E. (3)	Inside (4)	Outside (5)	S.E. (6)	Inside (7)	Outside (8)	S.E. (9)
Topographic characteristics:									
Elevation (m)	122.8	165.3	(14.69)***	123.5	124.6	(12.26)	122.3	117.7	(11.69)
Slope (%)	0.68	0.89	(0.05)***	0.68	0.64	(0.04)	0.67	0.62	(0.04)
Demographic characteristics:									
Household size	4.68	5.02	(0.10)***	4.66	4.58	(0.07)	4.64	4.52	(0.08)
Population (log)	4.38	4.32	(0.07)	4.25	4.17	(0.06)	4.16	4.19	(0.08)
Female (%)	0.48	0.48	(0.00)	0.48	0.48	(0.00)	0.48	0.48	(0.01)
Married (%)	0.37	0.38	(0.01)	0.37	0.39	(0.01)**	0.38	0.39	(0.01)
Christian (%)	0.85	0.87	(0.03)	0.85	0.85	(0.02)	0.83	0.85	(0.02)
Muslim (%)	0.11	0.10	(0.03)	0.11	0.13	(0.02)	0.13	0.13	(0.02)
African religion (%)	0.01	0.01	(0.00)	0.01	0.00	(0.00)	0.00	0.00	(0.00)
Kpelle (%)	0.31	0.28	(0.04)	0.31	0.33	(0.03)	0.31	0.36	(0.03)*
Bassa (%)	0.25	0.24	(0.06)	0.26	0.24	(0.04)	0.26	0.21	(0.04)
Other ethnic group (%)	0.43	0.48	(0.05)	0.42	0.42	(0.03)	0.42	0.42	(0.03)
Economic characteristics:									
Primary education (%)	0.28	0.25	(0.01)**	0.28	0.26	(0.01)**	0.26	0.26	(0.01)
Secondary education (%)	0.38	0.33	(0.02)***	0.37	0.34	(0.01)**	0.35	0.35	(0.01)
Owns house (%)	0.80	0.87	(0.02)***	0.81	0.84	(0.01)***	0.82	0.82	(0.01)
House condition: Good (%)	0.26	0.14	(0.02)***	0.25	0.20	(0.01)***	0.22	0.21	(0.01)
Asset ownership (%)	0.13	0.11	(0.01)***	0.13	0.12	(0.01)*	0.12	0.12	(0.01)
Distance to Monrovia (km)	112.3	166.9	(12.51)***	111.1	126.1	(10.08)	111.3	119.9	(9.11)
Distance to closest city (km)	19.82	33.72	(2.15)***	19.67	22.52	(1.36)**	19.58	21.06	(1.24)
Observations	1856	7830		1698	2141		1112	801	

Notes: Columns (1), (2), (4), (5), (7) and (8) give the means of the corresponding variable. Columns (3), (6) and (9) give the clustered standard errors for the difference in means in parenthesis. Clustered standard errors at the district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.3
Summary statistics, RD sample villages *versus* other villages.

	(1) Within bandwidth [−8.9 dBm, 8.9 dBm]	(2) Below (< −8.9 dBm)	(3) Above (> 8.9 dBm)	(4) *P*-value ((1)–(2))	(5) *P*-value ((1)–(3))
Topographic characteristics:					
Elevation (m)	120.10	169.42	123.65	0.000	0.775
	(119.46)	(131.93)	(137.61)		
Slope (%)	0.66	0.91	0.69	0.000	0.413
	(0.62)	(0.94)	(0.67)		
Demographic characteristics:					
Household size	4.62	5.06	4.70	0.000	0.422
	(1.52)	(1.85)	(1.59)		
Population (log)	4.18	4.33	4.58	0.061	0.000
	(1.40)	(1.41)	(1.77)		
Female (%)	0.48	0.48	0.49	1.000	0.593
	(0.10)	(0.09)	(0.09)		
Married (%)	0.39	0.38	0.36	0.133	0.000
	(0.12)	(0.11)	(0.11)		
Christian (%)	0.83	0.88	0.88	0.258	0.067
	(0.30)	(0.25)	(0.24)		
Muslim (%)	0.13	0.10	0.09	0.302	0.061
	(0.29)	(0.24)	(0.22)		
African religion (%)	0.00	0.01	0.01	0.579	0.530
	(0.03)	(0.03)	(0.03)		
Kpelle (%)	0.34	0.27	0.31	0.170	0.475
	(0.40)	(0.39)	(0.36)		
Bassa (%)	0.24	0.24	0.24	0.885	0.819
	(0.38)	(0.40)	(0.37)		
Other ethnic group (%)	0.42	0.48	0.44	0.264	0.627
	(0.42)	(0.45)	(0.40)		
Primary education (%)	0.26	0.25	0.30	0.556	0.000
	(0.17)	(0.18)	(0.16)		
Secondary education (%)	0.35	0.33	0.42	0.330	0.000
	(0.21)	(0.21)	(0.22)		
Economic characteristics:					
Owns house (%)	0.82	0.87	0.78	0.000	0.056
	(0.27)	(0.22)	(0.29)		
House condition: Good (%)	0.21	0.14	0.31	0.000	0.000
	(0.21)	(0.17)	(0.27)		
Asset ownership (%)	0.12	0.11	0.15	0.213	0.000
	(0.09)	(0.09)	(0.11)		
Distance to Monrovia (km)	114.47	171.30	111.58	0.000	0.767
	(89.33)	(102.05)	(103.47)		
Distance to closest city (km)	20.41	34.84	19.47	0.000	0.540
	(15.61)	(20.67)	(16.42)		
Observations	1547	7207	932		

Notes: Columns (1), (2), (3) give the means and standard deviations of the corresponding variable. Columns (4) and (5) give the *p*-value for the difference in means between column 1 and 2 (column 4) and columns 1 and 3 (column 5). Clustered standard errors at the district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.4
Effect of coverage on number of reported EVD cases.

	Dep. Variable = Number of Reported EVD Cases				
	Baseline (1)	Controls (2)	Topography Poly (3)	Polynomial RD (4)	Kernel choice (5)
Coverage	−1.414**	−1.213**	−1.454***	0.191	−1.287**
	(0.561)	(0.514)	(0.552)	(5.141)	(0.507)
Mean outside coverage	0.395	0.395	0.395	0.158	0.395
Bandwidth (dBm)	3.60	3.70	3.63	50.00	3.12
Observations	640	640	640	7014	640
Districts	69	69	69	115	69

Notes: Columns (1) and (2) present estimates of β using a local linear regression specification of Eq. (1). All specification include controls for elevation and slope. Column (2) adds controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Optimal bandwidth chosen as in Calonico et al. (2014). Column (3) uses topography polynomial: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. It also includes distance to tower as additional control. Column (4) uses a wider bandwidth of 50 dBm and a third-degree polynomial specification for $f(\tilde{R}_i)$. Column (5) uses a rectangular kernel. Optimal bandwidths chosen as in Calonico et al. (2014) except for column (4) which uses a fixed, wider bandwidth. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.5
Effect of coverage on likelihood of EVD suspected death.

	Dep. Variable = $\mathbb{1}${Number of Suspected EVD Deaths > 0}					
	Baseline (1)	Controls (2)	Topography Poly (3)	Polynomial RD (4)	Kernel choice (5)	Probit model (6)
Coverage	−0.069*	−0.074**	−0.069*	−0.058	−0.067*	−0.243*
	(0.037)	(0.030)	(0.036)	(0.036)	(0.035)	(0.129)
Mean outside coverage	0.111	0.111	0.109	0.091	0.106	0.087
Bandwidth (dBm)	10.64	10.26	11.10	50.00	9.02	.
Observations	1913	1913	2139	7014	1741	1741
Districts	86	86	87	115	84	84
Marginal Effect						−0.048

Notes: Columns (1) and (2) present estimates of β using a local linear regression specification of Eq. (1). All specification include controls for elevation and slope. Column (2) adds controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Optimal bandwidth chosen as in Calonico et al. (2014). Column (3) uses topography polynomial: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. It also includes distance to tower as additional control. Column (4) uses a wider bandwidth of 50 dBm and a third-degree polynomial specification for $f(\tilde{R}_i)$. Column (5) uses a rectangular kernel. Column (6) estimates a probit specification of Eq. (1). Optimal bandwidths chosen as in Calonico et al. (2014) except for column (4) which uses a fixed, wider bandwidth. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.6
Effect of coverage on number of suspected EVD deaths.

	Dep. Variable = Number of Suspected EVD Deaths				
	Baseline	Controls	Topography Poly	Polynomial RD	Kernel choice
	(1)	(2)	(3)	(4)	(5)
Coverage	−0.906***	−0.603**	−0.889***	−0.001	−0.339
	(0.333)	(0.258)	(0.316)	(2.250)	(0.231)
Mean outside coverage	0.395	0.395	0.395	0.207	0.389
Bandwidth (dBm)	3.17	3.53	3.34	50.00	5.91
Observations	640	640	640	7014	1010
Districts	69	69	69	115	75

Notes: Columns (1) and (2) present estimates of β using a local linear regression specification of Eq. (1). All specification include controls for elevation and slope. Column (2) adds controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Optimal bandwidth chosen as in Calonico et al. (2014). Column (3) uses topography polynomial: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. It also includes distance to tower as additional control. Column (4) uses a wider bandwidth of 50 dBm and a third-degree polynomial specification for $f(\tilde{R}_i)$. Column (5) uses a rectangular kernel. Optimal bandwidths chosen as in Calonico et al. (2014) except for column (4) which uses a fixed, wider bandwidth. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.7
Effect of coverage on number of months affected by EVD.

	Number of months affected by EVD				
	Baseline (1)	Controls (2)	Topography Poly (3)	Polynomial RD (4)	Kernel choice (5)
Coverage	−0.209**	−0.188**	−0.220**	−0.158**	−0.199**
	(0.096)	(0.084)	(0.099)	(0.074)	(0.087)
Mean outside coverage	0.128	0.128	0.128	0.090	0.142
Bandwidth (dBm)	8.72	8.97	8.59	50.00	7.10
Observations	1547	1547	1547	7014	1369
Districts	83	83	83	115	81

Notes: Columns (1) and (2) present estimates of β using a local linear regression specification of Eq. (1). All specification include controls for elevation and slope. Column (2) adds controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Optimal bandwidth chosen as in Calonico et al. (2014). Column (3) uses topography polynomial: $h(\mathbf{G}_i) = \rho_1 elev_i + \rho_2 elev_i^2 + \rho_3 slope_i + \rho_4 elev_i \times slope_i + \rho_5 elev_i^2 \times slope_i + \rho_6 slope_i^2 + \rho_7 elev_i \times slope_i^2 + \rho_8 elev_i^2 \times slope_i^2$. t also includes distance to tower as additional control. Column (4) uses a wider bandwidth of 50 dBm and a third-degree polynomial specification for $f(\tilde{R}_i)$. Column (5) uses a rectangular kernel. Optimal bandwidths chosen as in Calonico et al. (2014) except for column (4) which uses a fixed, wider bandwidth. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.8
Effect of coverage on likelihood of EVD case, by distance to coverage boundary.

	Geographic distance to coverage boundary					
	No restriction (1)	Within 2 km (2)	Within 4 km (3)	No restriction (4)	Within 2 km (5)	Within 4 km (6)
Coverage	−0.108**	−0.110**	−0.107**	−0.101**	−0.105**	−0.100**
	(0.048)	(0.049)	(0.049)	(0.042)	(0.043)	(0.042)
Mean outside coverage	0.093	0.094	0.093	0.093	0.093	0.093
Bandwidth (dBm)	8.99	9.04	8.96	8.13	8.15	8.07
Observations	1547	1729	1542	1547	1535	1542
Districts	83	84	83	83	83	83
Controls	No	No	No	Yes	Yes	Yes

Notes: Dependent variable equal 1 if there is at least one EVD case reported within the village. Columns (1)–(3) present estimates of β using a local linear regression specification of Eq. (1). All specification include controls for elevation and slope. Columns (4)–(6) add controls for average household size, population, female, married, Christian, Muslim, African religion, Kpelle, Bassa, primary education, literacy, house ownership and quality, an index of asset ownership, distance to Monrovia, and distance to closest main city. Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.

Table B.9
Average travel distance to closest village with coverage.

	No dBm restrictions (1)	Within 20 dBm (2)	Within 10 dBm (3)
Mean (km)	25.94	6.73	5.69
Std. Dev.	45.69	10.45	10.26
Observations	4639	1060	367
Share of non-coverage sample	0.59	0.50	0.46

Notes: Stata program "georoute" (Weber and Péclat, 2017) used for travel distance calculations. Program uses Open Source Routing Machine (OSRM) engine for calculations and OpenStreetMap as base map. Villages with calculated travel distance shorter than Euclidean distance dropped from analysis. Calculated travel distance can be shorter than Euclidean distance if village is far from road network used in the calculation and program uses segment of road available. Share of non-coverage sample refers to the share of non-coverage villages for which a travel distance was calculated.

Table B.10
Effect of coverage on likelihood of EVD case, control villages close to coverage.

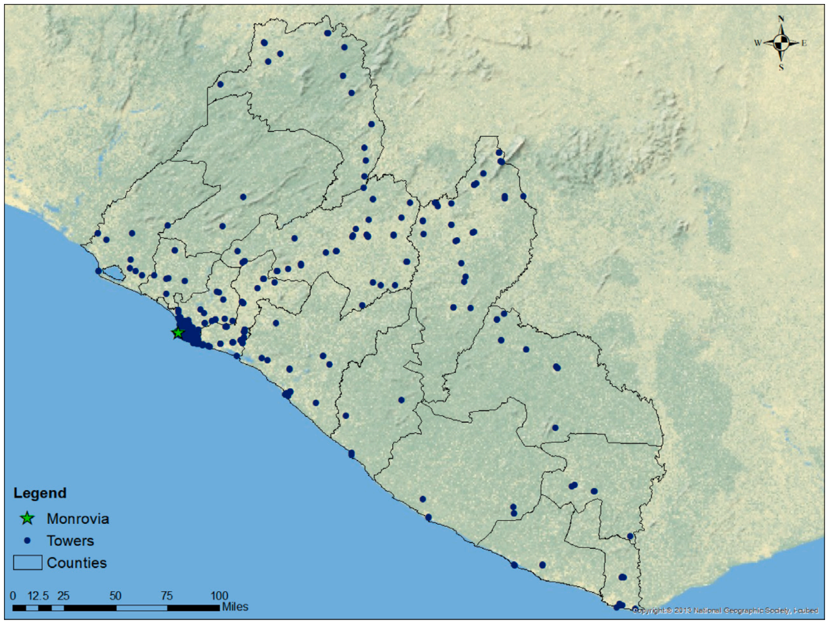
	Closest coverage village within:				
	2 km (1)	4 km (2)	6 km (3)	8 km (4)	10 km (5)
Coverage	−0.114*	−0.098*	−0.106**	−0.103**	−0.104**
	(0.060)	(0.055)	(0.052)	(0.049)	(0.048)
Obs	1515	1612	1651	1677	1695
Bandwidth	9.082	9.216	9.297	9.348	9.507

Notes: Estimates of β using a local linear regression specification of Eq. (1). All specifications include controls for elevation and slope. Optimal bandwidth chosen as in Calonico et al. (2014). Standard errors clustered at district level. *, **, and *** indicate 10, 5, and 1 percent significance, respectively. Stata program "georoute" (Weber and Péclat, 2017) used for travel distance calculations. Program uses Open Source Routing Machine (OSRM) engine for calculations and OpenStreetMap as base map. Villages with calculated travel distance shorter than Euclidean distance dropped from analysis. Calculated travel distance can be shorter than Euclidean distance if village is far from road network used in the calculation and program uses segment of road available.

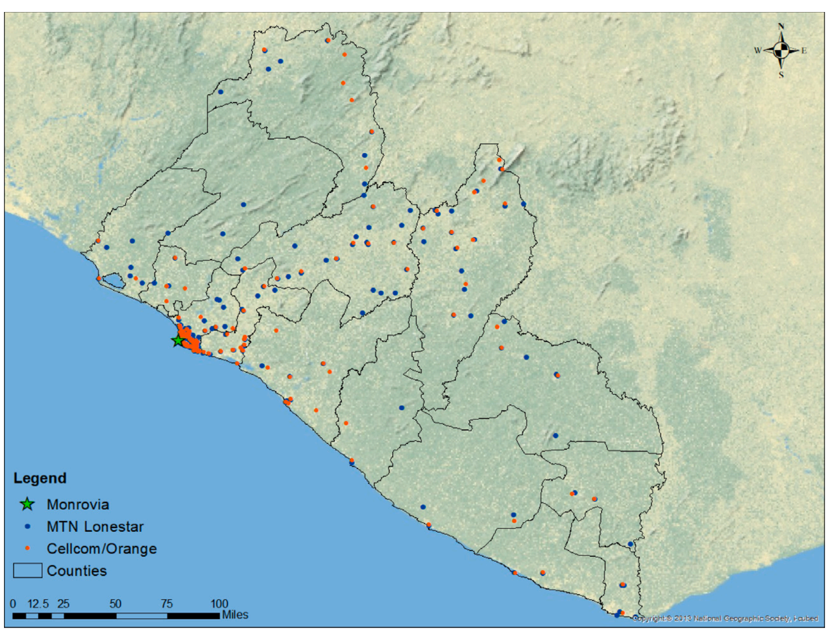
Table B.11
Main results using spatially-adjusted standard errors.

	Clustered S.E.	Conley SE (different distance cutoffs)		
	(1)	(2)	(3)	(4)
Coverage	−0.091*	−0.091**	−0.091**	−0.091**
	(0.046)	(0.039)	(0.041)	(0.046)
Mean outside coverage	0.093	0.093	0.093	0.093
Bandwidth (dBm)	8.99	8.99	8.99	8.99
Observations	1547	1547	1547	1547
Districts	83	83	83	83
Dist. cutoff	.	10	20	50

Notes: All columns present estimates of β using a local linear regression specification of Equation (1) that uses the optimal bandwidth of specification (1) in Table 2 (baseline specification). Note that the coefficient in column 1 is not bias-corrected. Columns 2–4 adjust standard errors for spatial correlation using a linear decay in the correlation between errors. Distance cutoff denotes the cutoff at which the correlation of errors beyond this cutoff is assumed to be zero. All specifications include controls for elevation and slope. *, **, and *** indicate 10, 5, and 1 percent significance, respectively.



(a) All Towers



(b) By Network Operator

Fig. B.1. Cellphone Towers, Liberia (2013). *Notes:* Cell towers' location obtained from the Liberia Telecommunications Authority (LTA).

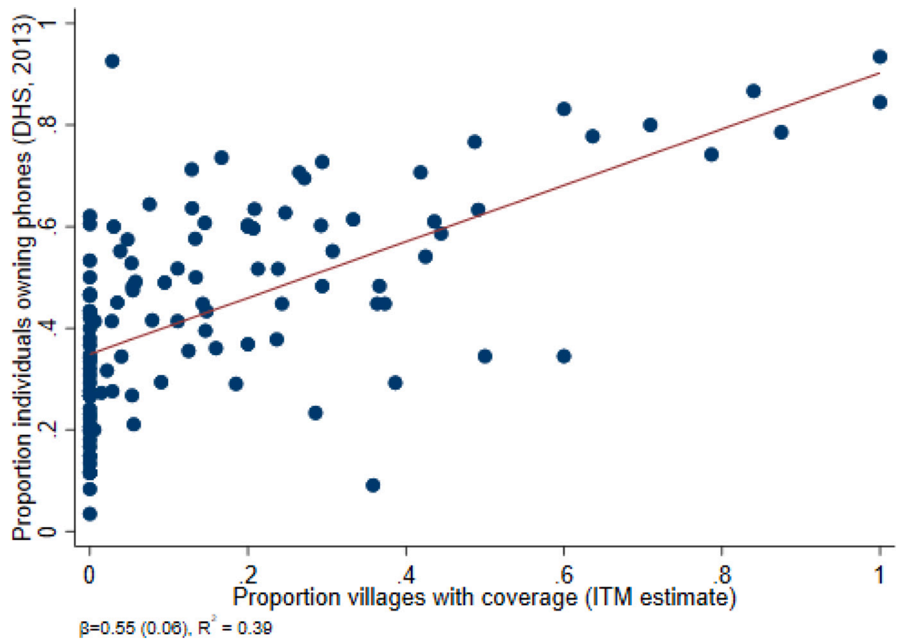
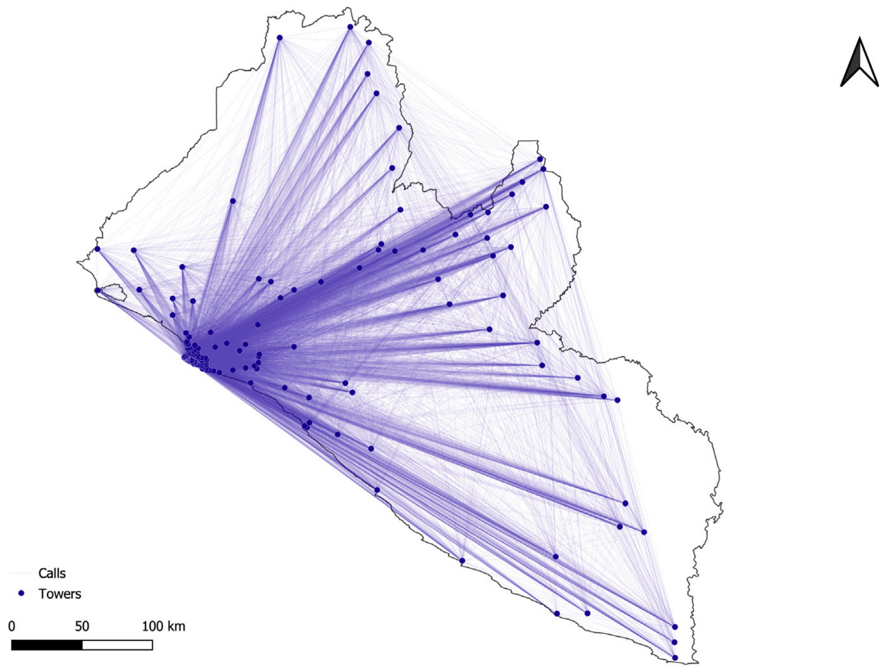
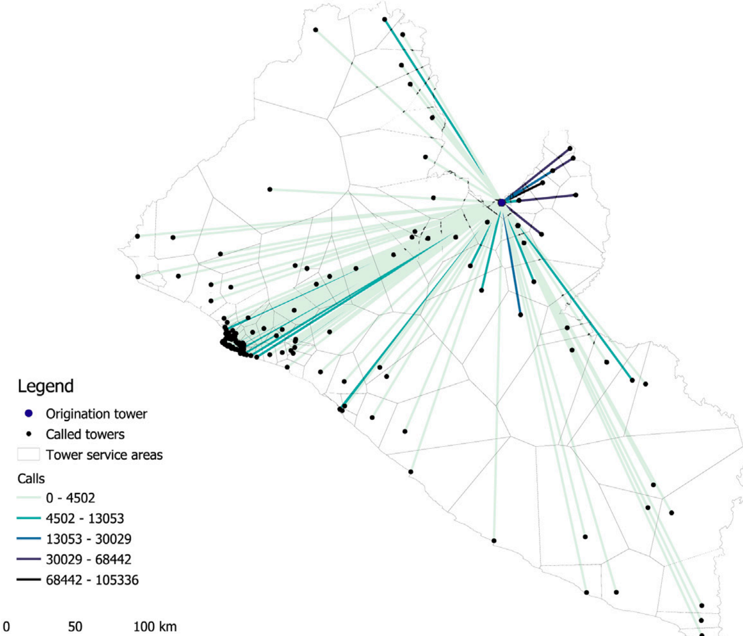


Fig. B.2. Predicted Coverage *versus* Reported Cellphone Ownership.
Notes: Correlation between the proportion of villages within a district that are estimated to have cellphone coverage by the ITM (X axis) and the proportion of individuals reporting owning a cellphone in the Demographic and Health Survey (DHS, 2013) (Y axis).



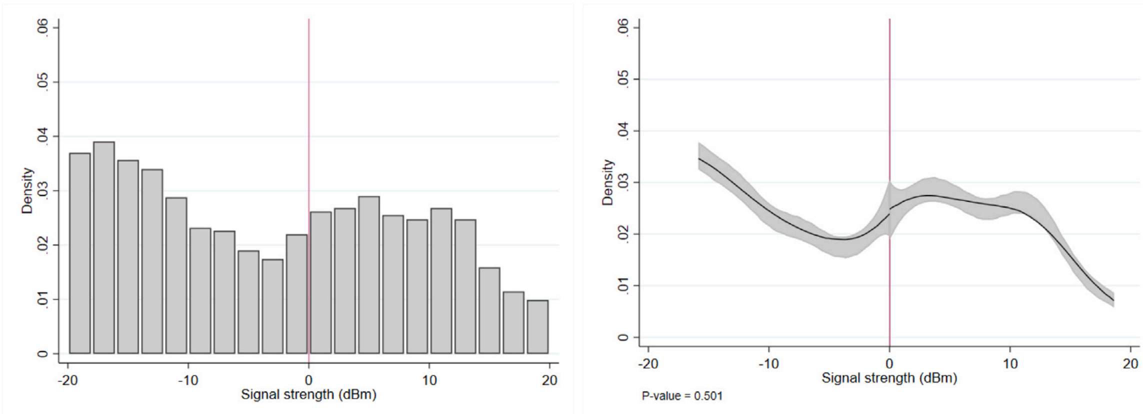
(a) All calls (June 15, 2015)



(b) Sample calls and called towers

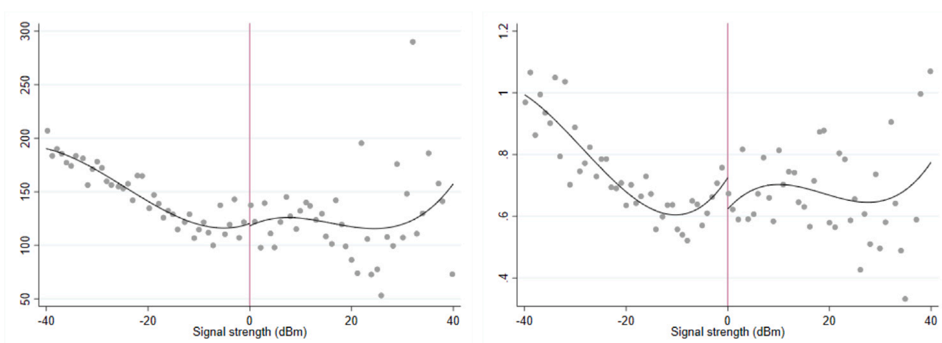
Fig. B.3. Example of Call Detail Records.

Notes: Panel a: All calls for the day of June 15, 2013. Panel b: Calls originating from cell tower in Ganta city (highlighted in blue) to all other towers in Liberia (black dots) in June and July of 2015. Color gradient represents the number of calls originating from Ganta city to the indicated tower. Darker shades mean a larger number of calls going to the indicated tower. Polygons indicate the service areas for each tower. Service areas calculated using a Voronoi/Thiessen polygon. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

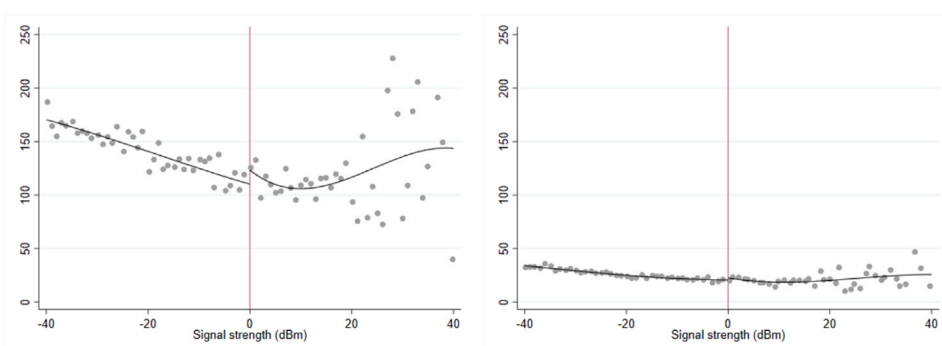


(a) Histogram (b) Test for manipulation

Fig. B.4. Histogram and Test for Manipulation of Forcing Variable.
Notes: "Signal strength (dBm)" refers to the received power or signal strength, measured in decibel milliwatts, at the village level. Signal strength is normalized so that negative values mean no coverage (i.e., not enough signal strength). Histogram bar width is 2 dBm. Panel (b) uses the test for breaks in the density of the forcing variable proposed in Cattaneo et al. (2019) and uses the code discussed in Cattaneo et al. (2018). *P*-value for test presented in figure caption.



Elevation (m) Slope (%)



Distance to Monrovia (km) Distance to closest city (km)

Fig. B.5. Regression Discontinuity Plots, Covariates.
Notes: Solid dots give the average of the specified variable for each bin of signal strength (dBm). "Signal strength (dBm)" refers to the received power or signal strength, measured in decibel milliwatts, at the village level. Signal strength is normalized so that negative values mean no coverage (i.e., not enough signal strength). The solid line trends give the predicted values from a regression of the outcome variable on a fourth degree polynomial in distance to the boundary that uses a triangular kernel.

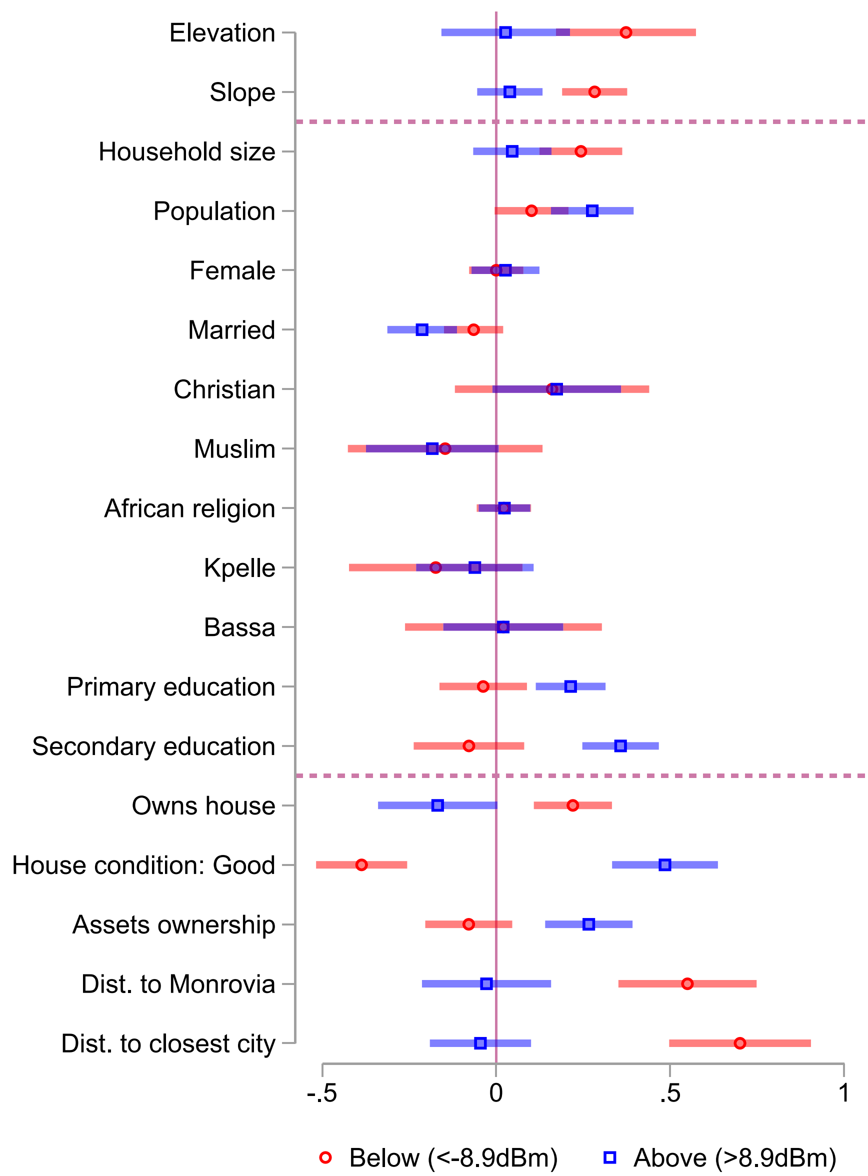


Fig. B.6. Difference in Means (RD Sample Villages vs Other Villages).
Notes: All variables are standardized for convenience. For each variable, dots give the difference in means between the specified group and the RD estimation sample. RD estimation sample is the sample within the Calonico et al. (2014) estimated bandwidth of analysis: [−8.9 dBm, 8.9 dBm]. Red dots refer to the sample below −8.9 dBm (less coverage than RD estimation sample). Blue dots refer to the sample above 8.9 dBm (more coverage than RD estimation sample). Lines around the dots give the 95% confidence interval for the difference in means. Dashed horizontal lines divide variables by topographic, demographic, and economic characteristics. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

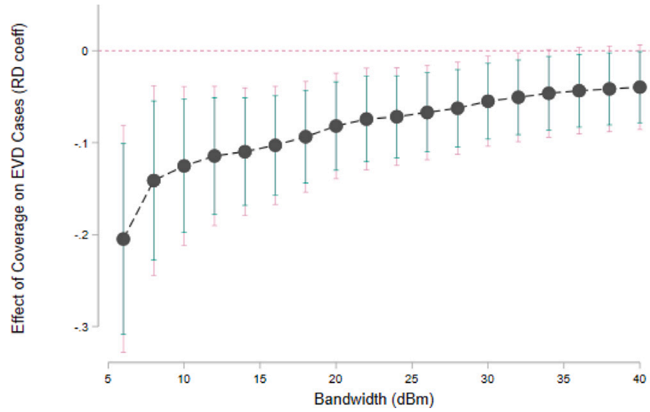


Fig. B.7. Bandwidth Sensitivity.
Notes: Solid dots indicate the RD estimate from Eq. (1) using the specified bandwidth. Range spikes indicate 95% and 90% confidence intervals of the estimates.

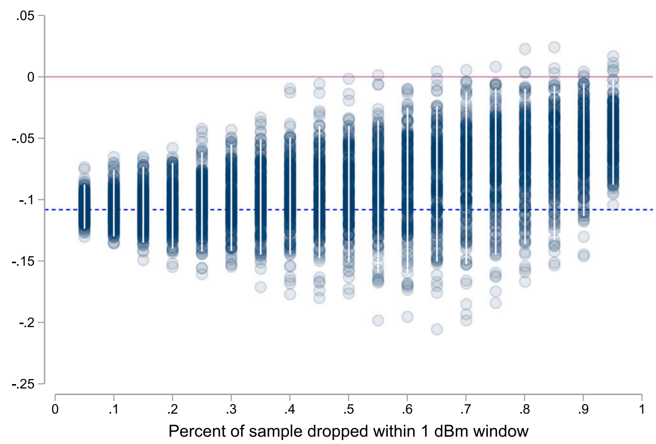


Fig. B.8. Sensitivity of Results to Observations near Treatment Cutoff.
Notes: Each dot gives the estimated RD coefficient from 250 replications that estimate Eq. (1) after randomly dropping the specified percent of observations within a 1 dBm window around the cutoff. White vertical lines indicate the 2.5th and 97.5th percentile of the estimated 250 coefficients within each percent dropped category. Blue dashed line gives the RD coefficient estimate from estimating Eq. (1) without any restrictions (i.e., dropping 0% of observations near cutoff).

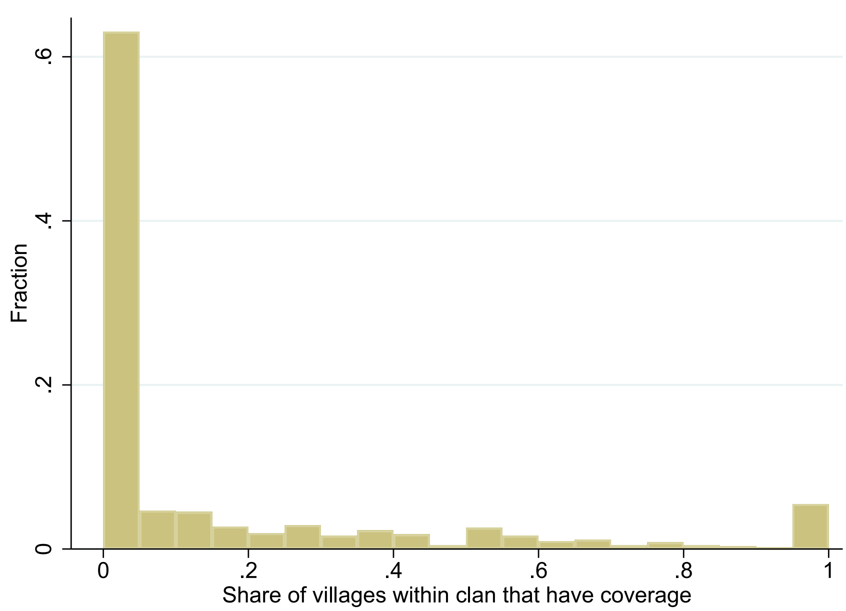
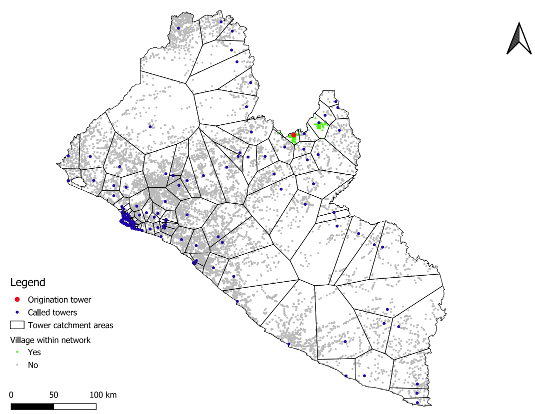
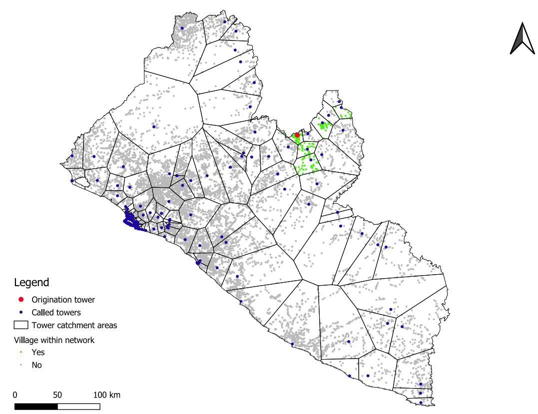


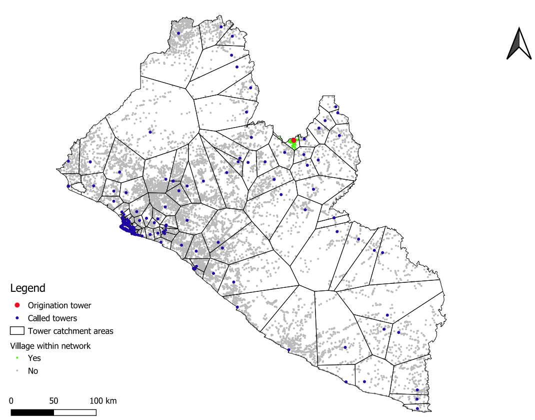
Fig. B.9. Histogram of the share of villages within a clan that have coverage.



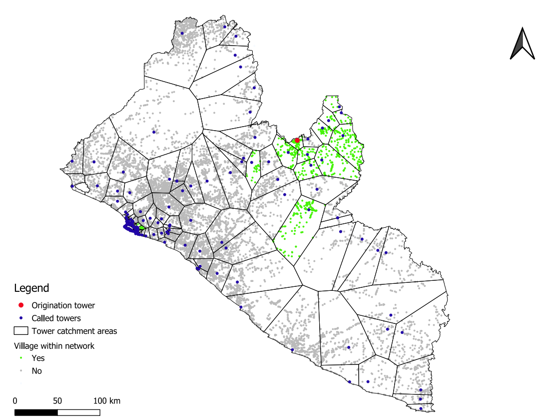
(a) Top tower



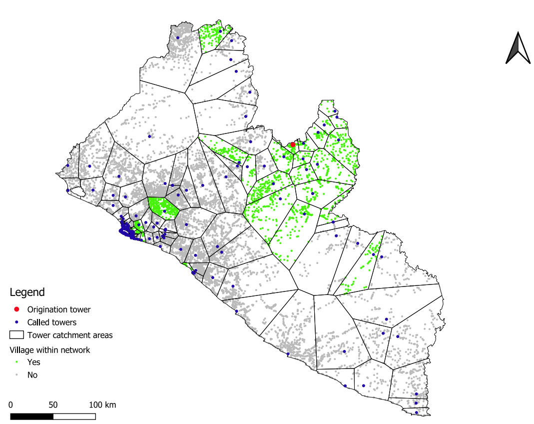
(b) Top 5 towers



(c) Top 50% calls

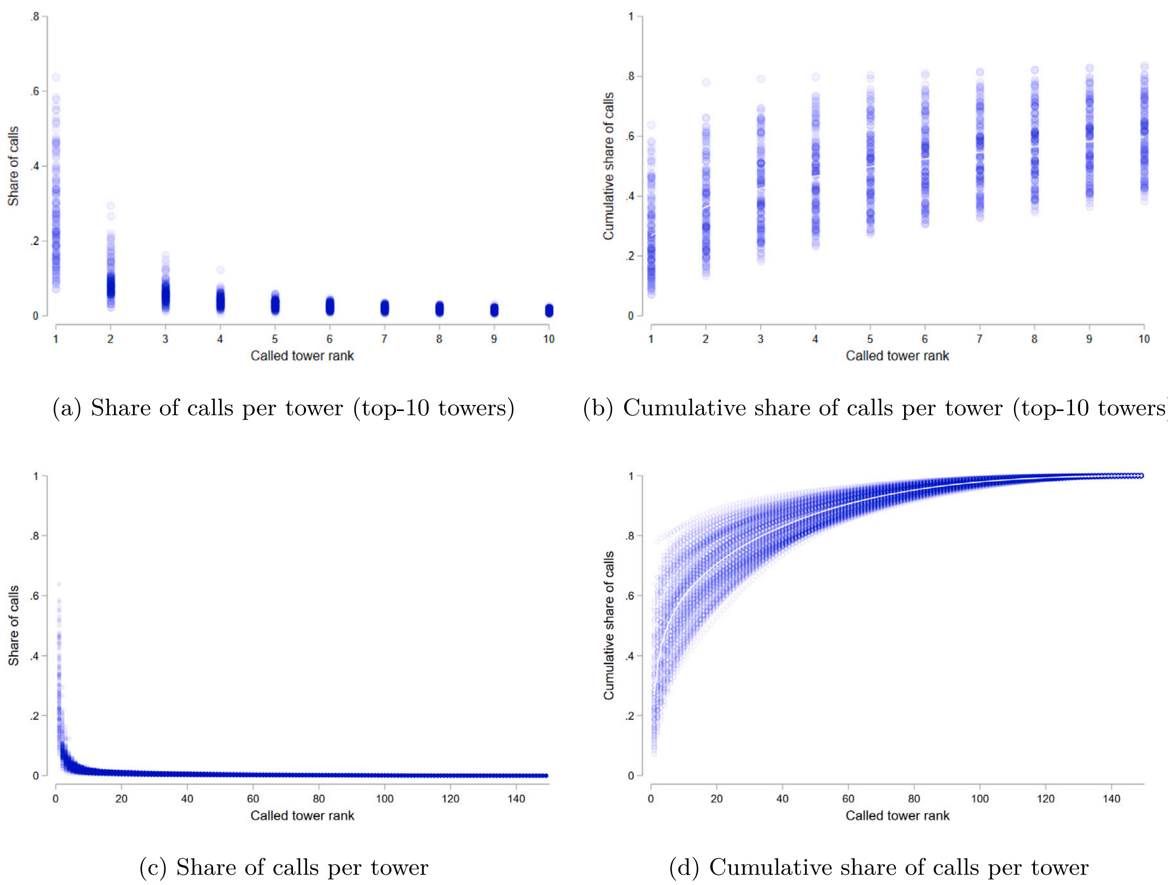


(d) Top 75% calls



(e) Top 90% calls

Fig. B.10. Example of Network for Ganta City, Different Definitions.
Notes: Networks for Ganta city tower, different definitions. Villages that are part of the network according to the specified definition are highlighted. Villages outside the network are gray. Given tower k, panels a and b define a network as all villages within the service area of tower k and the top and top-5 towers receiving calls originating from k. Panels c–e define a network as all villages within the service area of tower k and the towers receiving 50%, 75%, and 90%, respectively, of all calls originating from k. CDR data obtained from Cellcom Liberia.



(a) Share of calls per tower (top-10 towers)
(b) Cumulative share of calls per tower (top-10 towers)
(c) Share of calls per tower
(d) Cumulative share of calls per tower

Fig. B.11. Share and Cumulative Share of Calls per Tower.
Notes: Panels (a) and (c) depict for all towers in the sample, the share of all outgoing calls by the rank of the tower called. For example, the share of outgoing calls going to the top-ranked tower (most called tower) ranged between 10 to slightly above 60% (panel a). Panels (b) and (d) depict the cumulative share of outgoing calls.

Appendix C. Determining the coverage cutoff

In principle, although one may know the typical range for the minimum required signal strength or sensitivity cutoff, the exact cutoff c is generally unknown (Farahani, 2008). We employ three methods to determine potential cutoff(s).

First, we use the Difference in Kernels estimator described in Qiu (2011) and Porter and Yu (2015). We estimate Eq. (6) below for each potential signal strength cutoff r. In order to limit our search, we restrict values of r to be within 80 to 110 dBm.

$$M(r) = \frac{1}{nh}\sum_{i=1}^{n} Y_i K_1\left(\frac{R_i - r}{h}\right) - \frac{1}{nh}\sum_{i=1}^{n} Y_i K_0\left(\frac{R_i - r}{h}\right) \tag{6}$$

where R_i denotes the ITM-estimated received power (i.e., signal strength) in each village 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, respectively. We use a triangular kernel and a bandwidth h of 2 dBms. Intuitively, this procedure compares the kernel-weighted average of our outcome on the right and left sides of point r. $M(r)$ should be large if r is a cutoff point and small if r is not a cutoff point (or, if there is no treatment effect). Fig. C.1 presents the estimates of $M(r)$ for different values of r. There is a clear and distinct jump in the difference in the kernels estimator at $r = 95$.

Second, we employ one of the strategies described in Card et al. (2008). Specifically, we estimate a simple relationship between our outcome Y_i and the estimated signal strength R_i for several potential cutoffs r but within a fixed interval $[\underline{R}, \overline{R}]$:

$$Y_i = \alpha + \beta \mathbf{1}\left[R_i > r\right] + f(R_i) + \varepsilon_i \qquad \text{with} \qquad R_i \in [\underline{R}, \overline{R}] \tag{7}$$

where $f(.)$ is a function of the estimated signal strength R_i. We then choose potential cutoff r based on the r that yields the highest R^2. Intuitively, if there is a discontinuity at r, then any specification of (7)

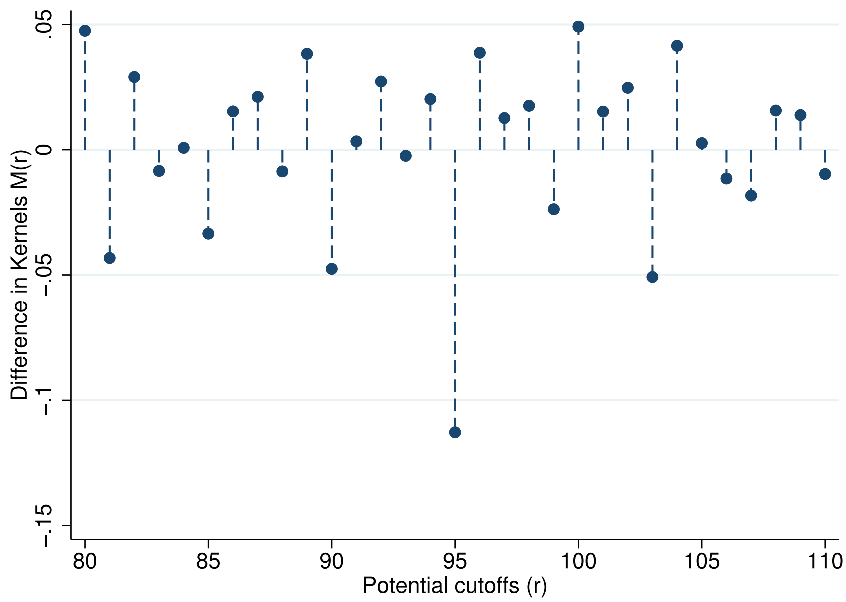


Fig. C.1. Difference in Kernels. *Notes:* Each dot represents the estimate of $M(r)$ in Eq. (6). Estimation.

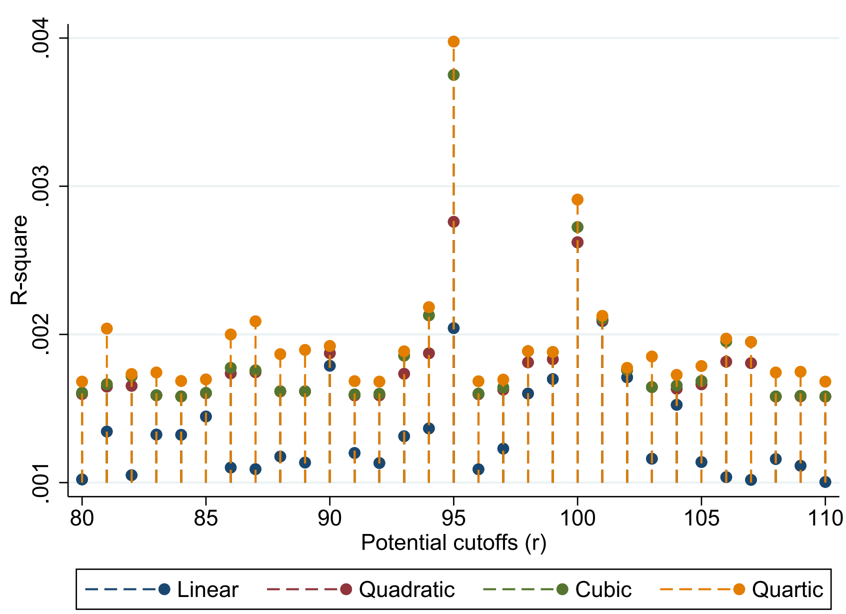
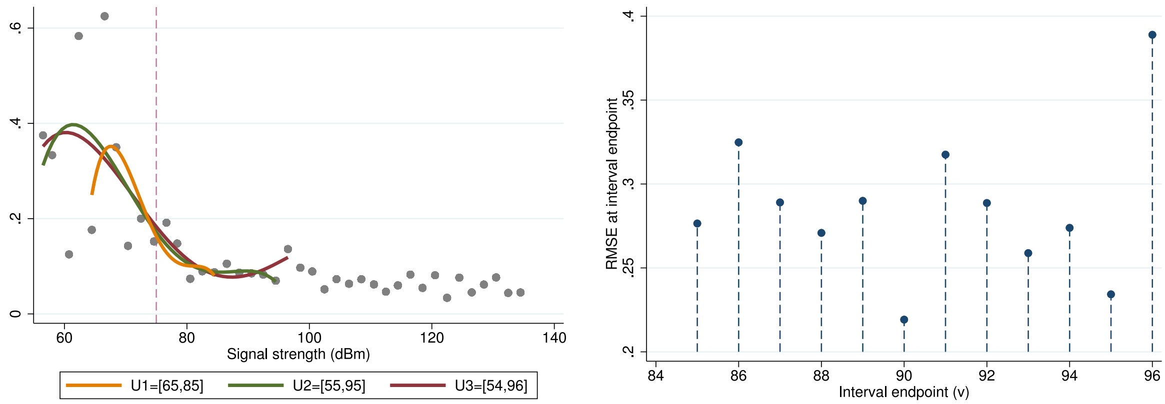


Fig. C.2. Model fit by Cutoff point r.
Notes: Each dot represents the R^2 obtained from estimating Eq. (7) using different values for r (x-axis), and using the specified polynomial for $f(R_i)$ in Eq. (7).



(a) Polynomial within interval U around $r_0 = 75$

(b) RMSE at interval endpoints

Fig. C.3. RMSE of Endpoint Residuals.
Notes: Solid dots in Panel A give the average of the outcome variable for each bin of signal strength (dBm). "Signal strength (dBm)" refers to the received power or signal strength, measured in decibel milliwatts, at the village level. Bin width is 2 dBm. The solid lines are the polynomials estimated within the specified intervals. Dashed vertical line is the initial $r_0 = 75$. Refer to text for more detail.

that uses a cutoff different than r is misspecified. As before, we restrict the search to values of r within 80 to 110 dBm and fix the interval $[\underline{R}, \overline{R}]$ to be within these values as well.

Fig. C.2 plots the R^2 for the specified cutoffs r and for a linear, quadratic, cubic, and quartic specifications of $f(R_i)$. Note that maximum fit is reached when Eq. (7) uses $r = 95$ suggesting that this is likely the cutoff at which our outcome jumps.

Lastly, we use a modification of the method proposed by Spokoiny (1998) adapted to our RD setting. Specifically, we proceed in three steps:

Step 1: We use a flexible polynomial $f(.)$ to estimate $Y_i = f(R_i - r_0) + \varepsilon_i$ within a neighborhood U of point r_0. Given our setting, we start with $r_0 = 75$ which is to the left of any plausible cutoff point and a fourth degree polynomial specification for $f(.)$.

Step 2: We then examine the RMSE of the residuals from Step 1 at point $v \in V$ where $V = U \cap U'$ is the set of boundary points of interval U. In our one-dimensional case, V are simply the left and right endpoints of the interval U. Furthermore, given that our potential cutoff is to the right of $r_0 = 75$, we focus only on the *right* endpoint residuals.

Step 3: We then gradually increase interval U around r_0 and repeat steps 1–2 until there is a clear jump in the RMSE of endpoint residuals. Cutoff point r is chosen as the endpoint of the maximal interval for which the endpoint residuals are "well-behaved" (i.e., the RMSE of the endpoint residuals does not jump significantly).

Intuitively, after we hit the maximal interval, the polynomial will have a hard time fitting the jump in the outcome occurring at the cutoff point. Panel A of Fig. C.3 illustrates the method by presenting polynomial $f(R_i - r_0)$ estimated within three different intervals around $r_0 = 75$: intervals U_1, U_2, and U_3, with right endpoints at 85, 95, and 96 dBm, respectively. Note that the fit of $f(.)$ is reasonably well at the endpoints of U_1 and U_2, but performs poorly at the endpoint of U_3. This likely suggests that we have reached the maximal interval at U_2. This is confirmed in panel B of Fig. C.3 that plots the RMSE at the endpoints of each interval used to estimate $f(R_i - r_0)$. Notice that once we get to a right endpoint of 96, the RMSE increases substantially. This suggests that the cutoff point is 95, the right endpoint of the maximal interval.

References

Adena, M., Enikolopov, R., Petrova, M., Santarosa, V., Zhuravskaya, E., 2015. Radio and the rise of the Nazis in Prewar Germany. Q. J. Econ. 130 (4), 1885–1939.

Agarwal, S., Perry, H.B., Long, L.-A., Labrique, A.B., 2015. Evidence on feasibility and effective use of mHealth strategies by frontline health workers in developing countries: Systematic review. Trop. Med. Int. Health 20 (8), 1003–1014.

Aker, J., 2010. Information from markets near and far: Mobile phones and agricultural markets in Niger. Am. Econ. J.: Appl. Econ. 46–59.

Aker, J., Fafchamps, M., 2014. Mobile phone coverage and producer markets: Evidence from west Africa. World Bank Econ. Rev. 29 (2), 262–292.

Aker, J., Ksoll, C., 2018. Can ABC lead to sustained 123? The medium-term effects of a technology-enhanced adult education program. Econom. Dev. Cult. Chang..

Aker, J., Ksoll, C., Lybbert, T.J., 2012. Can mobile phones improve learning? Evidence from a field experiment in Niger. Am. Econ. J.: Appl. Econ. 4 (4), 94–120.

Aker, J., Mbiti, I.M., 2010. Mobile phones and economic development in Africa. J. Econ. Perspect. 24 (3), 207–232.

Alexander, K., et al., 2015. What factors might have led to the emergence of Ebola in West Africa? PLoS Negl. Trop. Dis. 9 (6), e0003652.

Allgaier, J., Svalastog, A.L., 2015. The communication aspects of the Ebola virus disease outbreak in western Africa–do we need to counter one, two, or many epidemics? Croat. Med. J. 56 (5), 496.

Armand, A., Atwell, P., Gomes, J.F., 2020. The reach of radio: Ending civil conflict through rebel demobilization. Am. Econ. Rev. 110 (5), 1395–1429.

Bengtsson, L., Gaudart, J., Lu, X., Moore, S., Wetter, E., Sallah, K., Rebaudet, S., Piarroux, R., 2015. Using mobile phone data to predict the spatial spread of Cholera. Sci. Rep. 5, 8923.

Bengtsson, L., Lu, X., Thorson, A., Garfield, R., Von Schreeb, J., 2011. Improved response to disasters and outbreaks by tracking population movements with mobile phone network data: A post-earthquake geospatial study in Haiti. PLoS Med. 8 (8), e1001083.

Blumenstock, J.E., Eagle, N., Fafchamps, M., 2016. Airtime transfers and mobile communications: Evidence in the aftermath of natural disasters. J. Dev. Econ. 120, 157–181.

Braun, R., Catalani, C., Wimbush, J., Israelski, D., 2013. Community health workers and mobile technology: A systematic review of the literature. PLoS One 8 (6), e65772.

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

Card, D., Mas, A., Rothstein, J., 2008. Tipping and the dynamics of segregation. Q. J. Econ. 123 (1), 177–218.

Carey, J.M., Chi, V., Flynn, D., Nyhan, B., Zeitzoff, T., 2020. The effects of corrective information about disease epidemics and outbreaks: Evidence from Zika and yellow fever in Brazil. Sci. Adv. 6 (5), eaaw7449.

Cattaneo, M.D., Jansson, M., Ma, X., 2018. Manipulation testing based on density discontinuity. Stata J. 18 (1), 234–261.

Cattaneo, M.D., Jansson, M., Ma, X., 2019. Simple local polynomial density estimators. J. Amer. Statist. Assoc. 1–7.

Cole, S.A., Fernando, A.N., 2021. 'Mobile'izing agricultural advice technology adoption diffusion and sustainability. Econ. J. 131 (633), 192–219.

Conley, T., 1999. GMM estimation with cross sectional dependence. J. Econometrics 92, 1–45.

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

D'Ambrosio, M.V., Bakalar, M., Bennuru, S., Reber, C., Skandarajah, A., Nilsson, L., Switz, N., Kamgno, J., Pion, S., Boussinesq, M., Nutman, T., A. Fletcher, D., 2015. Point-of-care quantification of blood-borne filarial parasites with a mobile phone microscope. Sci. Translat. Med. 7, 286re4.

Durante, R., Pinotti, P., Tesei, A., 2019. The political legacy of entertainment TV. Amer. Econ. Rev. 109 (7), 2497–2530.

Enikolopov, R., Petrova, M., Zhuravskaya, E., 2011. Media and political persuasion: Evidence from Russia. Amer. Econ. Rev. 101 (7), 3253–3285.

Eppink, D., Kuebler, W., 1994. TIREM/SEM Handbook. Electromagnetic Compatibility Analysis Center Annapolis Md.

Fabregas, R., Kremer, M., Schilbach, F., 2019. Realizing the potential of digital development: The case of agricultural advice. Science 366 (6471), eaay3038.

Fallah, M.P., Skrip, L.A., Gertler, S., Yamin, D., Galvani, A.P., 2015. Quantifying poverty as a driver of Ebola transmission. PLoS Negl. Trop. Dis. 9 (12).

Farahani, S., 2008. Chapter 4: Transceivers requirements. ZigBee Wirel. Netw. Transceiv..

Feng, S., Grepin, K.A., Chunara, R., 2018. Tracking health seeking behavior during an Ebola outbreak via mobile phones and SMS. NPJ Digit. Med. 1 (51).

Freifeld, C.C., Chunara, R., Mekaru, S.R., Chan, E.H., Kass-Hout, T., Iacucci, A.A., Brownstein, J.S., 2010. Participatory epidemiology: Use of mobile phones for community-based health reporting. PLoS Med. 7 (12), e1000376.

Gagliarducci, S., Onorato, M.G., Sobbrio, F., Tabellini, G., 2020. War of the waves: Radio and resistance during world war II. Am. Econ. J.: Appl. Econ. 12 (4), 1–38.

Gonzalez, R.M., 2021. Cell phone access and election fraud: Evidence from a spatial regression discontinuity design in afghanistan. Am. Econ. J.: Appl. Econ. 13 (2), 1–51.

Grantz, K.H., Meredith, H.R., Cummings, D.A., Metcalf, C.J.E., Grenfell, B.T., Giles, J.R., Mehta, S., Solomon, S., Labrique, A., Kishore, N., et al., 2020. The use of mobile phone data to inform analysis of COVID-19 pandemic epidemiology. Nat. Commun. 11 (1), 1–8.

GSMA, 2014. Mobile coverage explorer. In: Collins Bartholomew. Available at: https://www.collinsbartholomew.com/mobile-coverage-maps/mobile-coverage-explorer.

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

Guriev, S., Melnikov, N., Zhuravskaya, E., 2020. 3G internet and confidence in government. Q. J. Econ..

Hampton, K.N., Sessions, L.F., Her, E.J., 2011. Core networks, social isolation, and new mediA: How internet and mobile phone use is related to network size and diversity. Inform. Commun. Soc. 14 (1), 130–155.

Holmes, K.K., Bertozzi, S., Bloom, B.R., Jha, P., Gelband, H., DeMaria, L.M., Horton, S., 2017. Major Infectious Diseases, Third ed. The International Bank for Reconstruction and Development and The World Bank, Washington (DC).

Hymowitz, D., 2017. State of Emergency: lessons from How Government Fought Ebola. Technical report, Tony Blair Institute for Global Change.

IFRC, 2014. Emergency Appeal Operation Update Ebola Virus Disease Emergency Appeals (Liberia, Sierra Leone, Guinea, Nigeria, Senegal and Africa Coordination). International Federation of Red Cross and Red Crescent.

International Media Support, 2007. A review of media support in the post-conflict transitional period and recommendations for future actions: Strengthening liberia's media. In: The partnership for media and conflict prevention in West Africa.

Jack, W., Suri, T., 2011. Mobile money: The economics of M-PESA.

Jack, W., Suri, T., 2014. Risk sharing and transactions costs: evidence from Kenya's mobile money revolution. Amer. Econ. Rev. 104 (1), 183–223.

Jack, W., Suri, T., Townsend, R., 2010. Monetary Theory and Electronic Money: Reflections on the Kenyan Experience. FRB Richmond Economic Quarterly.

JAXA, 2016. ALOS global digital surface model ALOS world 3D 30 m (AW3D30), Japan aerospace exploration agency.

Jensen, R., 2007. The digital provide: Information (technology), market performance, and welfare in the south Indian Fisheries sector. Q. J. Econ. 122, 879–924.

Kirsch, T.D., Moseson, H., Massaquoi, M., Nyenswah, T.G., Goodermote, R., Rodriguez-Barraquer, I., Lessler, J., Cumings, D.A.T., Peters, D.H., 2017. Impact of interventions and the incidence of Ebola virus disease in Liberia - implications for future epidemics. Health Policy Plan. 32, 205–214.

Kling, J.R., Liebman, J.B., Katz, L.F., 2007. Experimental analysis of neighborhood effects. Econometrica 75 (1), 83–119.

Krishna, A., Thompson, T.L., 2019. Misinformation about health: A review of health communication and misinformation scholarship. Am. Behav. Sci. 0002764219878223.

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

Liberian Ministry of Health, 2017. 2014 Annual report.

LISGIS, 2008. 2008 National Population and Housing Census. Liberia Institute of Statistics and Geo-Information Services (LISGIS), Available at: https://www.lisgis.net/page_info.php?7d5f44532cbfc489b8db9e12e44eb820=MzQy.

Longley, A.G., Rice, P.L., 1968. Prediction of Tropospheric Radio Transmission Loss Over Irregular Terrain. a Computer Method-1968. Institute for telecommunication sciences boulder co.

LTA, 2014. Annual Report 2013–2014. Liberia Telecommunications Authority.

Lu, X., Bengtssona, L., Holmea, P., 2012. Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. USA 109 (29), 11576–11581.

Maffioli, E.M., 2020. Collecting data during an epidemic: A novel mobile phone research method. J. Int. Dev..

Maffioli, E.M., 2021. The political economy of health epidemics: Evidence from the Ebola outbreak. J. Dev. Econ. 151.

Maffioli, E.M., Gonzalez, R., 2022. Are socio-demographic and economic characteristics good predictors of misinformation during an epidemic? PLOS Glob. Public Health 2 (3), e0000279.

Manacorda, M., Tesei, A., 2020. Liberation technology: Mobile phones and political mobilization in africa. Econometrica 88 (2), 533–567.

Mensah, J.T., Tafere, K., Abay, K.A., 2022. Saving Lives Through Technology. World Bank, Washington, DC.

Milusheva, S., 2020. Using Mobile Phone Data to Reduce Spread of Disease, no. 9198. World Bank Policy Research Working Paper.

Montez, D., 2010. Community Radio a Vital Resource for Liberians. AudienceScapes: Intermedia Knowledge Center.

Nyei, I., 2014. Decentralizing the state in Liberia: The issues, progress and challenges. Stabil.: Int. J. Secur. Dev. 3 (1).

Nyenswah, T.G., Kateh, F., Bawo, L., Massaquoi, M., Gbanyan, M., Fallah, M., et al., 2016. Ebola and its control in Liberia, 2014–2015. Emerg. Infect. Diseases 22 (2).

Obasola, O.I., Mabawonku, I., Lagunju, I., 2015. A review of e-health interventions for maternal and child health in Sub-Sahara Africa. Matern Child Health J. 19, 1813–1824.

Oliver, N., Lepri, B., Sterly, H., Lambiotte, R., Deletaille, S., De Nadai, M., Letouzé, E., Salah, A.A., Benjamins, R., Cattuto, C., et al., 2020. Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle. 6, (23), p. eabc0764,

Olken, B.A., 2009. Do television and radio destroy social capital? Evidence from Indonesian villages. Am. Econ. J.: Appl. Econ. 1 (4), 1–33.

Onyeonoro, U.U., Ekpemiro, U.C., Abali, C., Nwokeukwu, H.I., 2015. Ebola epidemic-the Nigerian experience. Pan Afr. Med. J. 22 (Suppl 1).

Ortiz-Martínez, Y., Jiménez-Arcia, L.F., 2017. Yellow fever outbreaks and Twitter: Rumors and misinformation. Am. J. Infect. Control 45 (7), 816–817.

Oyeyemi, S.O., Gabarron, E., Wynn, R., 2014. Ebola, Twitter, and misinformation: A dangerous combination? Bmj 349, g6178.

Pathak, R., Poudel, D.R., Karmacharya, P., Pathak, A., Aryal, M.R., Mahmood, M., Donato, A.A., 2015. Youtube as a source of information on Ebola virus disease. North Am. J. Med. Sci. 7 (7), 306.

Pew Research Center, 2011. Social networking sites and our lives: How people's trust, personal relationships, and civic and political involvement are connected to their use of social networking sites and other technologies. Technical report.

Pew Research Center, 2019. In Emerging Economies, Smartphone and Social Media Users Have Broader Social Networks. Technical report.

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

Qiu, P., 2011. Jump regression analysis. In: International Encyclopedia of Statistical Science. Springer, Berlin, Heidelberg.

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

Roberts, H., Seymour, B., Fish, S.A., Robinson, E., Zuckerman, E., 2017. Digital health communication and global public influence: A study of the Ebola epidemic. J. Health Commun. 22 (sup1), 51–58.

Ruble, K., 2015. The Village That Beat Ebola: How One Liberian Community Avoided the Outbreak. Vice News.

Sacks, J.A., Zehe, E., Redick, C., Bah, A., Cowger, K., Camara, M., Diallo, A., Nasser, A., Gigo, I., Dhillon, R.S., Liua, A., 2015. Introduction of mobile health tools to support Ebola surveillance and contact tracing in Guinea. Glob. Health: Sci. Prac. 2 (4), 646–659.

Seybold, J.S., 2005. Introduction to RF Propagation. John Wiley & Sons.

Shapiro, J.N., Weidmann, N.B., 2015. Is the phone mightier than the sword? Cellphones and insurgent violence in Iraq. Int. Organ. 247–274.

Spokoiny, V.G., 1998. Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice. Ann. Stat. 26 (4), 1356–1378.

United Nations Office for Disaster Risk Reduction, 2015. The Human Cost of Weather-Related Disasters: 1995–2015. Technical report.

Venkatraman, A., Mukhija, D., Kumar, N., Nagpal, S., 2016. Zika virus misinformation on the internet. Travel Med. Infect. Dis. 14 (4), 421–422.

Walsh, C., 2022. Social impacts of new radio markets in Ghana: A dynamic structural analysis. J. Econ. Lit. 4 (3), 43–60.

Weber, S., Péclat, M., 2017. A simple command to calculate travel distance and travel time. Stata J. 17 (4), 962–971.

Wesolowski, A., Qureshi, T., Boni, M.F., Sundsøy, P.R., Johansson, M.A., Rasheed, S.B., Engø-Monsen, K., Buckee, C.O., 2015. Impact of human mobility on the emergence of dengue epidemics in Pakistan. Proc. Natl. Acad. Sci. 112 (38), 11887–11892.

White House, 2014. White House, Office of the Press Secretary. 2014b. White House Update on U.S. Response to Ebola, 02 Dec. 2014. Technical report, Available from: https://www.whitehouse.gov/the-press-office/2014/12/02/fact-sheet-update-ebola-response.

World Bank, 2014. Press release. World Bank Group to Nearly Double Funding in Ebola Crisis to $400 Million, September 25 2014. Technical report, Available from: http://www.worldbank.org/en/news/press-release/2014/09/25/world-bank-group-nearly-double-funding-ebola-crisis-400-million.

World Health Organization, 2014. Ebola: Experimental therapies and rumoured remedies. World Health Organization.

World Health Organization, 2016. Situation Report, Liberia: June 2016. Technical report.

Yang, C., Yang, J., Luo, X., Gong, P., 2009. Use of mobile phones in an emergency reporting system for infectious disease surveillance after the Sichuan earthquake in China. Bull. World Health Organ. 87, 619–632.