Social Impacts of New Radio Markets in Ghana: A Dynamic Structural Analysis

Christoph Walsh*

May 3, 2023

Abstract

In the 1990s, Ghana opened up the radio broadcasting market to commercial stations. Exploiting quasirandom variation in radio coverage caused by coverage spilling through gaps in mountainous areas, I estimate the effect of radio coverage on social outcomes, such as malaria among children, and find positive benefits. I then estimate a dynamic structural model of radio station entry to simulate counterfactual regulatory policies that aim to increase access to radio coverage and deliver these benefits to new communities. I find that allowing stronger transmitters and targeted entry subsidies in unserved markets are effective ways to increase access to coverage.

Key words: Dynamic oligopoly, Radio regulation, Effects of media
JEL codes: D43, L13, L82

* Tilburg University, Department of Econometrics and Operations Research. Correspondence email: cbtwalsh@uvt.nl. I am indebted to my PhD advisor Marc Rysman for his invaluable guidance and encouragement throughout this project. I would also like to thank Francesco Decarolis, Jihye Jeon, and Andrew Newman for their valuable comments and suggestions. This paper also benefited from helpful comments from presentations at Boston University, Boston College, IIOC, The Federal Communications Commission, Tilburg University, CREST, Mannheim, QuantCo Berlin, Sciences Po, and EARIE. Tiantian Yu provided excellent research assistance.

1 Introduction

Researchers have documented positive effects of radio coverage on a wide range of social and economic outcomes, such as education (Keefer and Khemani, 2014), health (Vaughan and Swalehe, 2000), political participation (Wang, 2021), peace (Paluck and Green, 2009) and women's empowerment (Cheung, 2012; Okuyama, 2021). Access to mass media is also important to inform individuals about certain pressing issues, such as elections or the spread of infectious diseases. In rural parts of many developing countries, radio is the most popular form of mass media, as lower literacy rates and electricity penetration rule out newspapers, television, and internet for many individuals. Radio coverage is therefore important to ensure rural communities have access to information.

The radio broadcasting sector, however, is highly regulated because radio makes use of the limited frequency spectrum. These regulations can affect the entry decisions of radio stations, which in turn affect which areas receive radio coverage. This can then have an impact on the social and economic outcomes of communities. In Ghana, the media sector was under state control until a new constitution was promulgated in 1992 which allowed the entry of privately-owned radio stations. There has since been a large rise in the number of commercial radio stations, with over 300 commercial stations entering in the first twenty years. The entry of commercial radio stations is regulated in several dimensions which can affect how many stations enter and whether coverage will spread to rural areas. Regulations such as licensing fees and restrictions on transmitter strengths may result in certain areas without access to radio coverage.

In this article, I use data from Ghana's broadcasting regulator and topographical data to compute the coverage area of each radio station. Using these data, I first complement the existing literature that document the effects of radio coverage by estimating the effect of coverage on health and economic outcomes. I then estimate a dynamic structural model of radio station entry where I use the stations' overlapping coverage areas and the distribution of local languages to quantify the intensity of competition. Using this estimated structural model, I simulate entry patterns under counterfactual regulation schemes that aim to increase access to radio coverage. Based on which areas receive coverage under these alternative regulations, I evaluate their effectiveness using the estimated effects of radio coverage on health and economic outcomes.

To explore the effects of radio coverage, I exploit quasirandom variation in coverage caused by coverage spilling through gaps in mountainous areas. A radio station located

in a particular community broadcasts its content to the population living nearby. This station's coverage will eventually be blocked by hilly terrain. However, its coverage could spill through gaps in mountainous areas in the form of streaks. If these coverage streaks are sufficiently far away from the source, I argue that the radio station is unlikely to have strategically placed its radio mast within its primary broadcasting area to capture a particular far-away location that is near the border of a coverage streak. Therefore the locations near the borders of these streaks receive coverage in an as-if random fashion. I estimate the effect of radio coverage on different outcomes by comparing the locations on either side of these streak borders. Previous work has also exploited irregular terrain when estimating the effects of coverage (for example, Olken (2009), Yanagizawa-Drott (2014), and Gonzalez (2021)). In this article, the outcome variables and coverage data vary at a very fine geographic resolution, which enables using this coverage streaks approach to estimate the effects of radio coverage.

One outcome variable I use is the malaria parasite rate among children, as measured by the Malaria Atlas Project (Bhatt et al., 2015). Malaria is prevalent throughout all regions in Ghana and in 2017 accounted for 11.8% of all deaths in children under five years of age and 40.6% of health center admissions (Ghana Health Service, 2017). Malaria can also have long-lasting effects. For example, Bleakley (2010) finds that persistent childhood malaria infection can reduce adult income by as much as 50%. Radio has the potential to warn listeners of the risks of malaria and give information on how to prevent it. One way in which radio informs about malaria and other health-related issues is through "edutainment" programs. These are entertaining shows such as soap operas that also include informative messages such as how to prevent contracting diseases. According to those surveyed in the Demographic Health Surveys in 2003-2014, 80% stated they heard messages about malaria on the radio, whereas for television and newspapers this number was 52% and 15% respectively. I find that individuals in areas that receive coverage experience an additional 1.6 percentage-point drop in the malaria prevalence rate. I also document using the Demographic Health Survey data that households in areas with radio coverage are also more likely to have their children sleep under mosquito bed nets.

Areas receiving radio coverage may learn of employment opportunities and more productive practices. A second variable I use is nighttime luminosity as seen from space, which has been used in the literature as a proxy for local GDP (for example, Henderson et al. (2012)). I find that areas receiving coverage experience an increase in nighttime luminosity of 1.7%. Finally, I also explore the effect of coverage on the Normalized Dif-

ference Vegetation Index (NDVI), which is a satellite image-based measure of agricultural productivity used by Asher and Novosad (2020), among others. I find areas receiving coverage experience an increase of 1.1% in summer NDVI, indicating that these areas see an increase in agricultural productivity.

Based on the evidence of the positive impacts of radio coverage, I proceed to estimate a dynamic structural model of radio station entry in order to evaluate the effectiveness of counterfactual regulation schemes in delivering the benefits of radio to new communities. Stations typically compete in oligopoly markets involving many dynamic and strategic interactions, such as preemptive entry. Furthermore, because of increased radio ownership and the growth of the market over time, much of the gains from entry are only realized in later time periods. These considerations would be difficult to capture in a static model.

In the model, the country is split into a grid of a large number of locations. Radio listeners in each location choose to tune into one of the different stations available to them. Listeners obtain different payoffs from public stations, local commercial stations broadcasting in their local language, and other commercial stations. The optimal choice of station for each listener determines the listener share for each radio station in each location. The profits for an active radio station from a location is determined by the product of the listener share, the population living in the location, and the monetization rate of listeners. The total profits for a station is then the sum of these profits over the entire area where the station has coverage. The coverage areas of each potential station in turn depend on the ruggedness of the surrounding terrain.

A potential radio station in a location chooses whether or not to obtain a broadcasting license based on a privately-observed entry cost and the expected present discounted value of future profits. Due to the nonstationary nature of the data, I assume the dynamic game has a finite horizon. Because there are a large number of stations, I make a simplifying behavioral assumption on potential stations' strategies. Similar to the moment-based Markov equilibrium approach of Ifrach and Weintraub (2017), I assume that stations' equilibrium policy functions are a function of a summary measure of the market state instead of the market state itself. For this summary measure I use the counterfactual profits the station would receive if they were currently active. This is a measure of how the current market state maps into payoffs. In the model, stations believe rivals play according to this policy function and enter according to the expected present discounted value of future profits based on these beliefs. In equilibrium, these beliefs are consistent with rivals' actual strategies.

Due to the computational burden of solving for an equilibrium, I do not solve for the equilibrium when estimating the model parameters. For each trial value of the model parameters in estimation, I first approximate the equilibrium policy function with a logistic regression using the observed entry decisions and the stations' counterfactual profits at those parameter values. I use this fitted equilibrium policy function to simulate rivals' behavior when simulating the expected present discounted value of profits from entering for each station. The parameters are then estimated by maximizing the likelihood of the observed entry decisions using these value functions. This estimation approach is similar to the two-step approaches developed by, for example, Bajari et al. (2007). However, because the profit function is nonlinear in the parameters the value function needs to be simulated for each trial value of the parameters.

The model estimates show that consumers prefer to listen to commercial stations from their local language area more than public stations, and prefer public stations to commercial stations from outside their language area. Profits are also increasing over time, as radio ownership and the monetization of listeners increase.

Radio stations are subject to significant policy oversight as they make use of a limited frequency spectrum. Yet these regulations can impact which locations receive radio coverage. Because access to radio coverage can bring important benefits, it is important to understand how the regulatory environment can impact access. I then use the structural model, together with the estimates of the positive effects of radio coverage, to evaluate counterfactual regulation schemes.

I first consider a counterfactual policy where stations have transmitters that are 50% stronger, which would result in larger coverage areas. The effect of such a policy change on entry is not obvious ex-ante. On the one hand, radio stations could have a larger base of listeners, which would make entry more attractive. On the other hand, they may experience greater competition, which would make entry less attractive. The total cost of transmission split over a station's life makes up only a very small fraction of its total costs, and therefore I assume stations' costs are unchanged for this counterfactual exercise. Yordy (2008) collected data on the startup costs and annual costs of many radio stations in Africa, including one commercial station in Ghana. For this station, the cost of the transmitter represented only 1.6% of its total costs. Its annual energy costs were also less than 2% of its operating costs. Because the cost of transmission is only a small fraction of total costs, the cost of obtaining a stronger transmitter would also have little

impact on total costs.[1]

I find that the larger listener base effect dominates the competition effect and there is entry of 12% more commercial stations by 2015. The proportion of the population receiving coverage also increases by 5% to 87.9%. Using the estimated effects of radio coverage, in 2015 there would be 2,740 fewer malaria infections among children, an increase in GDP of $33.65m, and an increase of 0.042% in agricultural productivity. There would also be additional benefits, such as additional political participation and women's empowerment as found by previous researchers. Therefore, a policy allowing stronger transmitters would be effective in delivering the benefits of radio to more communities, and would not be costly to implement.

I then consider the cost of a targeted entry subsidy that would achieve the same level of access by 2015 as the transmitter strength counterfactual. The subsidy I consider is proportional to the number of individuals newly served by radio coverage after a station's entry. For example, if a newly entered radio station reaches 30,000 people, but 20,000 of these were already served by a different station broadcasting from another location, then the station receives a subsidy only for the 10,000 newly served individuals. This subsidy can also be interpreted as a reduction in the authorization fees for qualifying stations. I find that the subsidy that results in the same level of access would have a total cost of 4.2% of total industry profits over the sample period. The subsidy is equivalent to receiving a once-off payment of 7.8% of the average annual profits of a station for each additional 10,000 individuals newly served by radio. This is more costly than the higher transmitter strength policy, but it results in more stations and also a higher level of access earlier in the sample. For example, 72.4% are served with the targeted subsidy by the year 2000, whereas only 54.7% are served under the transmitter strength policy (compared to 46.7% in the baseline case).

Related Literature: This article contributes to four main strands of literature. First, it adds to the wide literature studying the effects of mass media on social outcomes. Gentzkow (2006), Oberholzer-Gee and Waldfogel (2009), Gentzkow et al. (2011) and Cagé (2019) study the effect of media on voter turnout whereas Gerber et al. (2009), Enikolopov et al. (2011), Larreguy et al. (2020), Garcia-Arenas (2016) and Durante et al. (2019) study its

[1] I collected a sample of prices of 93 transmitters from 17 brands from a retailer in Ghana's capital Accra. For brands where products differ only by their transmission power, a 50% more powerful transmitter costs on average 45.3% more. However, because transmission costs represents only a small fraction of total costs, the total costs of opening and running a station would not change significantly under the policy.

effect on political leaning. Strömberg (2004), Besley et al. (2002) and Snyder and Strömberg (2010) find that government responsiveness and political accountability are greater in areas with better media coverage. Farré and Fasani (2013) find that media informs individuals about the net gains from migration and Keefer and Khemani (2014) find positive effects on parental investment in children's education. Jensen and Oster (2009) find an increase in women's status in the household and lower domestic violence, La Ferrara et al. (2012) find that soap operas lower fertility and Kearney and Levine (2015) find that reality TV leads to fewer teenage pregnancies. Not all of the effects of mass media are positive, however. Olken (2009) finds that radio and television lower social capital and trust, DellaVigna et al. (2014) find increased nationalism and ethnically offensive graffiti from cross-border coverage and Yanagizawa-Drott (2014) finds increased violence in Rwanda from propaganda on the radio. Paluck (2009) and Paluck and Green (2009), however, find that educational radio soap operas reduced violence in Rwanda. Furthermore, Armand et al. (2020) find that messages on the radio reduced conflict during the Lord's Resistance Army insurgency. From these results, we see that the majority of the effects of mass media are positive, particularly for radio. This motivates studying how regulation and competition affect the entry decisions of radio stations, as they can deliver important benefits to rural communities. This article contributes to this literature by estimating the effect of radio coverage on malaria, growth and agricultural productivity in Ghana. My identification strategy is also able to exploit quasirandom variation in coverage arising from coverage streaks, which is possible in my context because the outcome variables and coverage data vary at a very fine geographic level.

Second, it adds to the literature estimating dynamic oligopoly models of entry (such as Ryan (2012) and Igami (2017)). The model in this article differs from previous work in that it allows the coverage areas of the stations, which are determined by geographical features, to define how stations compete. This setup allows for unique counterfactual experiments, such as the effect of stronger transmitters on entry decisions.

Third, it also contributes to the literature using structural estimation tools from industrial organization to study policy issues in developing countries. Hidalgo and Sovinsky (2022) estimate a structural demand model for internet in Columbia; Rysman et al. (2023) estimate a dynamic oligopolistic model of bank branching in Thailand; Ryan (2021) estimates an optimal bidding model in the Indian electricity market; and Vitali (2022) estimates a location choice model with consumer search frictions using data on Ugandan garment firms.

Finally, it also adds to market studies on the radio broadcasting industry. (Berry and Waldfogel, 1999) and Berry et al. (2016) find that there is more entry than is socially optimal in US radio markets. As this article is in a developing context, I instead argue that radio access should be provided to rural areas without access. Sweeting (2009, 2010, 2013) studies advertising times, mergers and musical performance rights and Jeziorski (2014a,b, forthcoming) studies ownership caps and mergers in the US radio industry. In my context, there are no mergers and very few firms operating multiple stations, so I focus solely on the entry decisions of stations.

2 Industry Background and Data

FM Broadcasting in Ghana

After Ghana's independence from the United Kingdom in 1957, the country went through various military governments where the media was under state control. This ended in 1992 when a new constitution allowed the entry of private media in Ghana. However, the government at the time delayed the provision and allocation of broadcasting licenses to private owners. In 1994, a pirate radio station called Radio EYE was set up in the nation's capital Accra as a form of protest which pressured the government to begin issuing licenses.[2] The National Communications Authority (NCA) was then established as the regulator overseeing the issuing of broadcasting licenses and licenses were first issued in 1995. Since then, there has been a large growth in the number of commercial stations, with 352 commercial stations entering in the first twenty years since liberalization.[3]

Radio is a very important form of mass media in Ghana, particularly during the sample period. For many individuals, radio may be the only form of mass media available to them. According to the 2010 Housing and Population Census, adult literacy was 74.1% which rules out newspapers for a quarter of the population. Only 64.2% of households reported using electricity. Even those with electricity, power outages (known locally as *Dumsor*) are very frequent. This rules out televisions for a large portion of the population. Furthermore, only 7.9% of households owned a desktop or laptop computer and only 7.8%

[2]There were also pirate radio stations on university campuses before campus licenses were issued. Since then, however, pirate radio stations are uncommon. According to Kafewo (2006) there were none in 2006. The regulator also carries out routine spectrum monitoring exercises and shuts down pirate stations when found.

[3]This is approximately one-third of the number of radio stations that the US had in per capita terms.

of the population 12 years and older had access to the internet. In rural areas, literacy, electricity penetration and internet penetration are even lower. Radios are inexpensive and do not always require electricity to operate. They can be run on batteries and hand crank radios are also available. According to the Demographic Health Surveys in 2003, 74.7% of households had a radio, and 28.6% of households had radio but did not own a television or read the newspaper.[4] Radio stations can operate at a very local level due to the lower cost of creating content compared to television, allowing them to broadcast in the local language of the community. The country has a population of about 30 million, yet there are over 60 actively-spoken languages.

For licensing purposes, radio stations in Ghana are classified as one of five types: Public, Public Foreign, Commercial, Community and Campus.[5] The license type is important because it determines how stations can generate revenue, the permissible size of its coverage radius and the application fees, initial fees and annual fees the stations must pay. Commercial stations are required to pay initial authorization fees and annual spectrum fees. Commercial stations and Public Foreign stations are restricted to have a 45km coverage radius. According to the NCA, this restriction was put in place to avoid frequency congestion and to allow for frequency re-use in different parts of the country in order to enhance media pluralism. However, during the sample period the only area with a number of stations approaching the limit is the nation's capital Accra. Community and Campus stations are restricted to a 5km radius. Public stations, on the other hand, do not have any coverage limitations.

Radio Station Coverage Data

I use data on the entry and exit of radio stations from Ghana's National Communications Authority (NCA). The NCA lists of all the authorization dates of license holders since the first authorizations in 1995. Of the 352 authorized commercial radio stations, 296 are independently owned, whereas 19 companies held two broadcasting licenses and 6 companies held three broadcasting licenses. As the opening of multiple stations by the same firm rarely occurs in the data, I abstract away from network expansion in the structural model. Compared to commercial station entry, there was very little entry of campus, community

[4] I show statistics from the Demographic Health Surveys across years in Table A.1 in the Online Appendix.

[5] Section A.2 in the Online Appendix states how the NCA defines each license type.

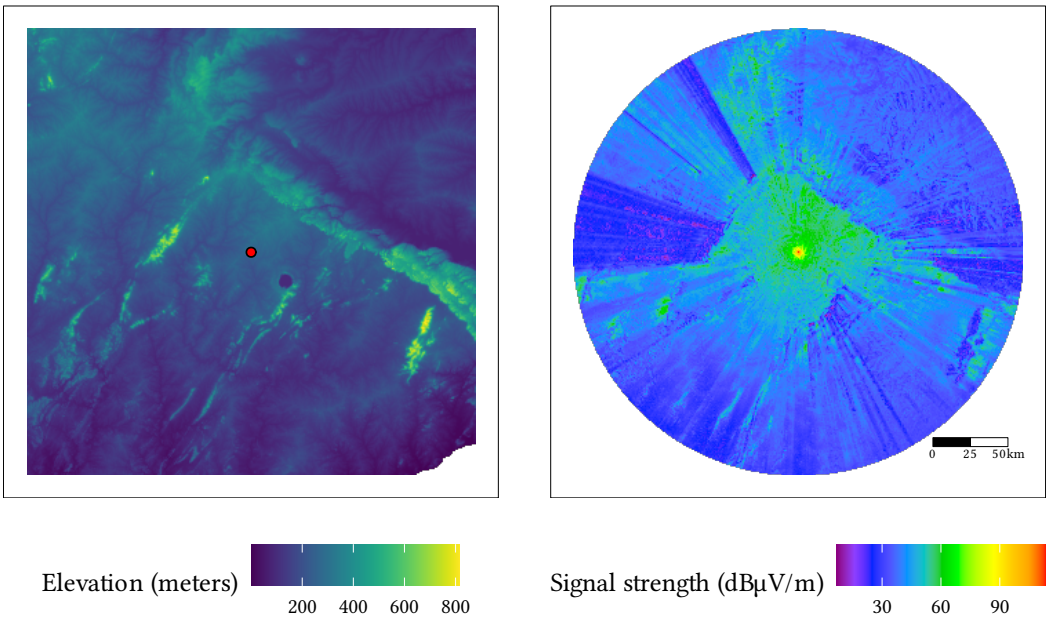
and public stations.[6] Because there are relatively few campus and community stations, and their coverage areas are much smaller than commercial and public stations, I omit them from the structural analysis.

I merge the NCA data with data from FMLIST, a worldwide radio station database. These data contain various other information about the stations, such as their height above ground, wattage and GPS coordinates. I use the Longley-Rice Irregular Terrain Model to compute the field strength of each radio station's coverage. This model was initially used by the Federal Communications Commission to predict the extent of stations' coverage areas to ensure that the coverage from stations with the same frequency in different locations did not overlap. The model computes how the field strength of radio transmission degrades as the signal reaches obstructions such as hills and mountains. For these obstructions, I use data from the Shuttle Radar Topography Mission, which is a worldwide digital elevation model at three arc-second resolution (roughly 90m×90m). I use the command-line software SPLAT (radio frequency Signal Propagation, Loss, And Terrain analysis tool) to perform these calculations.

Figure 1 shows an example of this calculation for one radio station. The left panel in Figure 1 shows the elevation near the city of Kumasi in Ghana. Brighter areas represent areas with higher elevation. The central point represents the location of one radio station's tower. The right panel in Figure 1 shows the signal strength of that radio station's coverage in the same area. The signal strength is measured in decibel-microvolts per meter (dBμV/m). Bright areas represent where the signal quality is high and dark areas represent where the signal quality is very poor. As the signal reaches mountainous obstructions, the color darkens rapidly, indicating a sharp drop in signal quality. The Federal Communications Commission, as well as other regulatory bodies around the world, consider 60 dBμV/m to be the threshold of signal strength for FM radio broadcasting. In this article, I will also use the 60 dBμV/m cutoff. As can be seen from the figure, coverage with a signal strength of above 60 dBμV/m can go past the regulatory 45km radius. A small amount of coverage goes past 45km for the majority of stations in my data. However, the majority of a station's coverage area lies within a 45km radius for all commercial stations my data, and very little of any stations' coverage areas go past 45km at a 72 dBμV/m strength, another cutoff used by regulatory agencies.

The resulting coverage data from the software is at a 3 arc-second resolution (approx-

[6] I show the cumulative number of authorized stations by license type over time in Figure A.2 in the Online Appendix.



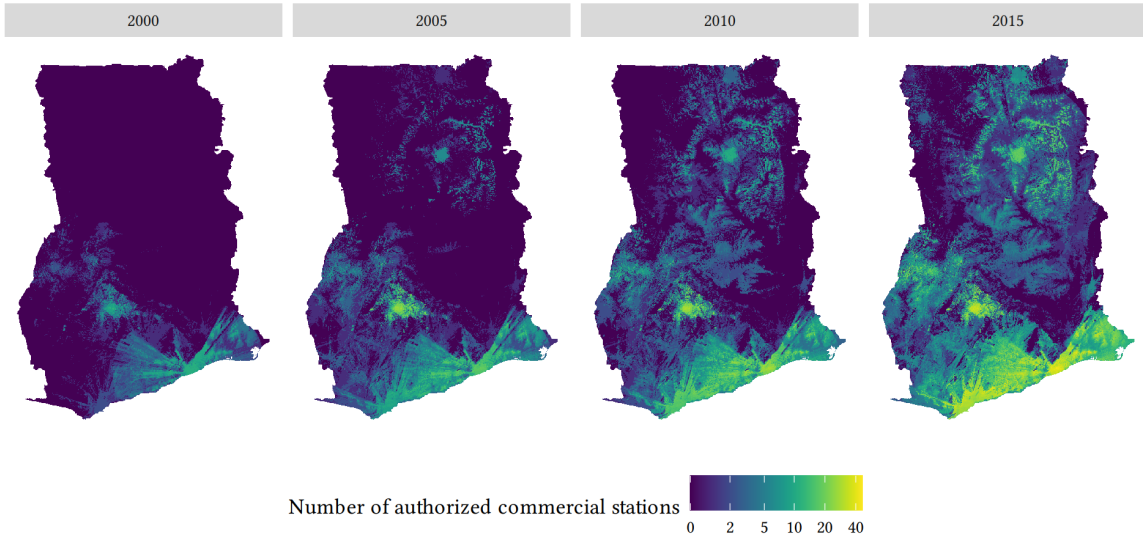
The left figure shows the elevation in a particular area where the central point indicates the location of a radio station's tower. The right figure shows the computed signal strength of that radio station in the same area, measured in decibel-microvolts per meter.

FIGURE 1: Example radio station coverage output.

imately 90m×90m) which results in several million data points for each station. In order to merge these data with other spatial data, I aggregate the coverage data to 30 arc seconds which makes the resolution approximately 900m×900m. For the structural model, I aggregate the data further to a 5 arc-minute resolution to maintain tractability.

It should be noted that these coverage predictions are computed from a model and as a result are not perfect. The elevation data does not take into account other obstructions, such as forests and large buildings. Kasampalis et al. (2013) perform a validation exercise comparing predictions of the model to actual field readings of field strength. They find that the predictions from the SPLAT software has an average error of -0.5 dBμV/m and a standard deviation of 5.5 dBμV/m. As 60 dBμV/m is considered the minimum strength required for a station to be picked up by a radio, the average error is very small.[7] I also aggregate the data from 3 arc seconds to 30 arc seconds which would smooth out this error. Furthermore, I also tested the SPLAT software by computing the coverage areas

[7]In Figure A.1 in the Online Appendix I show a histogram of the probabilities of station-cells being pivotal (i.e. crossing to the other side of the 60 dBμV/m threshold) according to this error distribution. 90.2% of station-cells effectively have a zero probability, 95% have a probability less than 5%, and the average probability is 1.2%.



The number of authorized commercial stations available with signal strength above 60 dBμV/m at each point in Ghana at approximately 900m×900m resolution, 2000, 2005, 2010 and 2015.

FIGURE 2: Number of authorized commercial stations.

of different US radio stations and found consistent coverage predictions with a portable radio.

Figure 2 shows the output from generating the coverage maps for every radio station. The figure shows the total number of authorized commercial stations available at each point in 1995, 2005 and 2015 at approximately 900m×900m resolution using the 60 dBμV/m coverage threshold. There is a large variation in the number of stations across the country, as well as large variation over time. The country's capital Accra, the coastal city in the southeast, has the most stations. The northern regions have very few stations, with many areas not having any coverage even by the end of the sample period.

The NCA also have reports which document whether the stations are currently "On Air" or "Off Air". I use reports from 2009Q3-2016Q3 which enables me to observe entry and exit in the data at a quarterly frequency during this period. Exit is very rare in the data, so the structural model focuses only on the entry decisions of radio stations.

Historical data on content and advertising times are unavailable for the stations in the sample period. Therefore I abstract away from horizontal differentiation in format choice and advertising in the structural model. In section A.3 in the Online Appendix, I show

summary information on languages and genres for a subset of the stations in the data where data were available. From this we see that most stations broadcast a mix of news, talk and music, but there are a small number of stations specializing in certain genres, such a religious stations. We also observe that stations often broadcast in a mixture of English (the lingua franca in Ghana) and a local language. In the structural model, I allow listeners to obtain a different flow utility from listening to commercial station broadcasting from in their language area compared to stations from a different language area. I use data from Glottolog to obtain the distribution of local languages in the country. This contains the locations for 64 distinct local languages for Ghana. I provide additional details about these data in Section A.4 in the Online Appendix.

Additional Data

Malaria Incidence: To study the effect of radio coverage on malaria, I use data from the Malaria Atlas Project. News programs and "edutainment" programs on the radio can inform individuals about how to avoid contracting malaria. According to those surveyed in the Demographic Health Surveys in 2003, 2008 and 2014, 80% stated they heard messages about malaria on the radio. For television and newspapers, this number was 52% and 15% respectively. Bhatt et al. (2015) constructed detailed maps of the malaria parasite rate among 2-10 year olds in Sub-Saharan Africa for the years 2000-2015 at a 15 arc-second resolution. The maps are estimated using a geostatistical model that uses various survey datasets and climate data. Figure A.3 in the Online Appendix shows these data from 2000 to 2015. The rate is high throughout most of the country, particularly in the earlier periods. Cities have a much lower rate of malaria, but my analysis focuses on the effect of radio coverage in rural areas.

Nighttime Luminosity: Individuals may learn about improved farming practices or employment opportunities, which can increase growth in the area. To study effect of radio coverage on development, I use a measure based on nighttime luminosity. Nighttime luminosity data have been used as proxies for local GDP in an ever-increasing number of applications, for example, Henderson et al. (2012) and Michalopoulos and Papaioannou (2013). Henderson et al. (2012) also find that the relationship between the growth rate of nighttime luminosity and GDP is approximately linear, particularly for longer time horizons. I obtain these data from the National Oceanic and Atmospheric Administration. The data are based on satellite images captured by the US Air Force at night between 8:30 PM

and 10:00 PM local time around the world. These images are then processed and cleaned to represent the average amount of light emanating from a geographic location during a year. Observations obstructed by clouds are excluded, as well as observations with light coming from forest fires, gas flares, sunlight (during summer months) and moonlight. Nighttime luminosity data are available from 1992-2013 at a 30 arc-second resolution around the world (approximately 900m×900m in Ghana). Values in the data are represented on a scale that ranges from 0 to 63 which measures the amount of light captured by the camera's sensor. This scale is top-coded at 63, although top-coding is rare in Ghana. Only 0.04% and 0.14% of observations are top-coded in 1992 and 2013 respectively. For certain years, data from two different satellites are available. In these years, I take the average of the values.

Finally, because I use the temporal variation in nightlight luminosity, it is necessary to intercalibrate the digital number values across years. The reason for this is the data come from different satellites, each with different camera settings (this is documented, for example, by Wu et al. (2013)). We use a similar approach to Rysman et al. (2023) and scale all luminosity values in a year such that the sum of values within the country's borders equals the national GDP from that year (constant $PPP GDP from the World Bank). I show maps of the intercalibrated nighttime luminosity for different years in Figure A.4 in the Online Appendix.

Normalized Difference Vegetation Index: To study the effect of radio coverage on agricultural productivity, I use a measure based on satellite imagery called the Normalized Difference Vegetation Index (NDVI). The index measures the amount of light absorbed by plants for photosynthesis. When the vegetation is lush, a higher proportion of light will be absorbed by the vegetation. I obtain NDVI images at an 8 arc-second resolution from the Global Inventory Monitoring and Modeling System for 2000-2016. Like Asher and Novosad (2020), I take the average NDVI values during the growing season in the summer months when NDVI peaks. I show maps of the measure for different years in Figure A.5 in the Online Appendix. Here we can see that southern Ghana has a higher vegetation index, apart from the larger metropolitan areas.

Population: The structural entry model requires a measure of potential listenership. For this I use rasterized population maps from NASA's Socioeconomic Data and Applications Center (SEDAC). These are available at five-year intervals from 1990-2020 at a 30

arc-second resolution. These maps were constructed by combining administrative data at the district and town level with satellite data to measure the urban extent of towns. I show an example population map in Figure A.6 in the Online Appendix.

3 Positive Outcomes of Radio Coverage

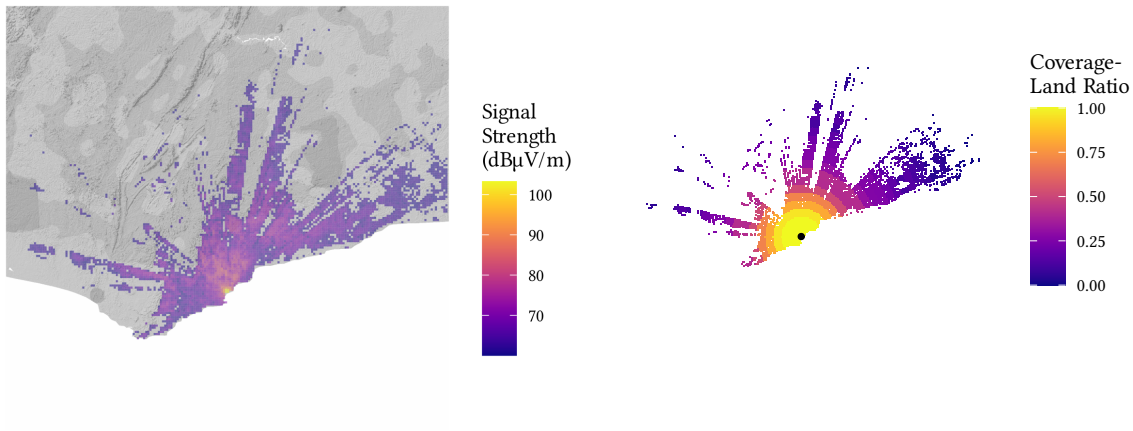
In the literature review, we learned that radio and other forms of mass media affect a host of different social and economic outcomes. In this section, I show that these positive effects continue to hold in this context for outcome variables where we have fine-grained geographic data.

Identification Strategy

Radio coverage is not randomly assigned throughout the country because stations may prefer to locate in areas with higher population and less competition. However, irregular terrain creates some randomness in radio coverage which can be exploited to estimate its effects on different outcomes. Figure 3a shows the coverage area for a radio station in Sekondi-Takoradi. This city has a population of approximately half a million and the coverage area of the station encompasses the city and some surrounding areas. As a result of hills away from the station, the coverage stops abruptly in certain areas. However, there are some gaps in the hills which allow some streaks of coverage to spill through into some rural areas. Given the population living in these streaks is a small fraction of the station's main coverage area in the city, I argue that the station is not positioning its radio mast strategically within the city to include certain areas beyond the hills. Therefore locations close to the border of one of these coverage streaks receive coverage in an as-if random fashion. We can then compare the outcomes of areas on either side of the borders of these streaks to estimate the effect of radio coverage on these outcomes.[8]

The coverage data is a grid with a resolution of approximately 900m×900m per cell. Each observation is then a cell-year combination. I identify cells that are near the border of a coverage streak in a particular year in the following way. First, for every station, I find the ratio of coverage to land at distance bins away from each transmitter. This is shown

[8]In the following section I discuss the structural model of radio station competition that I estimate. We can compute the share of profits coming from these coverage streaks using this estimated model. For the median station, this is only 0.2%, giving further evidence that stations are unlikely to strategically place their masts in order to capture particular areas in these streaks.



(A) Example coverage streaks caused by coverage spilling through gaps in mountainous areas.

(B) The ratio of coverage to land at distance bins away from the transmitter.

FIGURE 3: Graphical example of identification strategy.

in Figure 3b. Cells very close to the station have a ratio of one because all cells with land at those distances have coverage. Further away, however, the ratio falls as some coverage gets blocked by hills. When the coverage-land ratio is sufficiently small, which mostly occurs in the outer range of its coverage and in more rural areas, I argue that the station did not strategically place its tower to capture particular points on the edge of its coverage area. I identify streaks of coverage as areas where the coverage-land ratio is below 0.2.

Due to the presence of hills, it is also possible that the coverage-land ratio is small near the source of coverage. In these cases, the station may have intended for coverage to reach those locations. Therefore I only include areas that are in the upper quintile of distance away from the station and at least 10km from the nearest transmitter. The elevation change causing the coverage streak should also be far away from a location, as local elevation changes may affect outcomes directly. Therefore I exclude locations with elevation changes exceeding fifty meters in a 9km×9km grid around the observation. Finally, it is possible a location is near a coverage streak of one station but in the non-random coverage area of a different station. I only include locations near coverage streaks that are not in the normal coverage area of other stations in the regressions.

In the regressions I include all cell-year observations that are within one kilometer of the border of a coverage streak. Thus, a particular cell-year is in coverage streak if (i) it

has radio coverage, (ii) the coverage-land ratio is less than 0.2, (ii) it is in the outer quintile of the station's coverage, (iii) it is at least 10km from the nearest station, and (iv) it has a maximum elevation change of 50m (at a 90m×90m resolution) within a 9km×9km grid surrounding it. In Section A.5 in the Online Appendix I show an example of a coverage streak according to these definitions and describe exactly which cells are included in the regressions.

The estimating equation using this subset of data is:

$$y_{\ell t} = \beta \times coverage_{\ell t} + \alpha_i + \delta_t + \epsilon_{\ell t} \tag{1}$$

where ℓ indexes cells (a 900m×900m plot of land), t indexes years, $y_{\ell t}$ is an outcome variable of interest, $coverage_{\ell t}$ is a dummy variable for if cell ℓ has coverage from any station at time t, α_i are cell fixed effects, δ_t are year fixed effects, and $\epsilon_{\ell t}$ is the error term. This approach requires the outcome variable to vary at a very fine geographic level. I use malaria prevalence among children measured by Bhatt et al. (2015), local GDP measured by the amount of light in nighttime satellite images, and agricultural productivity measured by how much light is absorbed by vegetation in satellite imagery (NDVI).

Estimating equation (1) via the standard two-way fixed effects (TWFE) estimator has been shown to be biased when observations are treated at different time periods and the treatment effects are heterogeneous (see, for example, De Chaisemartin and d'Haultfœuille (2020) and Callaway and Sant'Anna (2021)). This is because the TWFE approach estimates a weighted sum of all possible individual 2×2 difference-in-differences estimates where the weights can often be negative. Moreover, in a staggered difference-in-differences design, the TWFE approach uses already-treated units as part of the control group for later-treated units, which is inappropriate in the presence of dynamic effects. Therefore I instead use the method from Callaway and Sant'Anna (2021) which does not suffer from negative weighting and does not use already-treated units as control groups for later-treated units. Their approach involves estimating the group-time average treatment effects on the treated (ATT) for every treated group (cells receiving coverage for the first time in a particular year) and every post-treatment time period. These estimates are then aggregated using group sizes as weights.

Estimating a group-time ATT involves estimating a TWFE regression on a subset of the data with two time periods. I will explain how this works through an illustrative example. Consider estimating the group-time ATT for 2010 for cells receiving coverage

	Dependent variable:		
	Malaria prevalence	Nighttime luminosity	Vegetation index
	(1)	(2)	(3)
Control groups: Never and not yet received coverage			
Coverage	−0.016	0.017	0.007
	(0.003)	(0.007)	(0.003)
Control groups: Never received coverage only			
Coverage	−0.014	0.018	0.007
	(0.003)	(0.006)	(0.003)
Sample size	25328	34176	12560
Mean dependent variable	0.618	0.033	0.663

Estimates of the overall average treatment effect on the treated following Callaway and Sant'Anna (2021) using the outcome regression approach with cell fixed effects, year fixed effects, and weights based on group size. Robust standard errors clustered at the cell level are in parentheses. Nighttime luminosity is transformed using the inverse hyperbolic sign transformation.

Table 1: Outcomes of radio coverage results.

for the first time in 2005. The two time periods included in the regression would be 2004 (pre-treatment) and 2010 (post-treatment). The cells included are all observations in the treated group (those receiving coverage for the first time in 2005) and the control group. There are two possible choices of control groups under their approach. The first is the group of "never-treated" cells, which in this case are all cells that never received coverage by the end of the sample period. The second control group additionally includes the "not-yet-treated" cells, which are all cells that received coverage for the first time after 2010. I present results using both control groups. The group-time ATT is then estimated for all combinations of groups (years where locations receive coverage for the first time) and post-treatment time periods. The overall ATT is then the weighted average of all group-time ATTs, with weights corresponding to the size of the groups.

Results

Table 1 shows the results. The upper panel shows the results for all three variables using never- and not-yet-treated units as the control group, whereas the bottom panel uses only the never-treated units. Column (1) shows that areas receiving coverage on average experience a drop in the malaria prevalence rate among children of either 1.4 or 1.6 percentage

points, depending on the control group. This is over a baseline rate of 61.8%, so it represents approximately a 2% drop over the baseline. In Section A.6 in the Online Appendix I merge the radio coverage data with Demographic Health Survey data and find descriptive evidence that children are more likely to sleep under mosquito bed nets in areas receiving radio coverage. This shows a plausible mechanism behind these findings.

Column (2) shows the results for nighttime luminosity. Nighttime luminosity, scaled to be measured in units local GDP of $100,000 per cell, is transformed by the inverse hyperbolic sine transformation. The results shows that areas receiving coverage experience an increase in local GDP of 1.7-1.8%. Finally, column (3) shows the results for the vegetation index, a measure of agricultural productivity. This is measured on a scale from 0 to 1. We observe that areas receiving coverage experience a growth in NDVI of 0.7 percentage points over a baseline of 66.3%, thus a 1.1% increase. Areas receiving radio coverage may hear of employment opportunities and more productive farming practices, increasing both GDP and NDVI. The results for NDVI thus also provide a plausible mechanism for the nighttime luminosity results.

I present a number of robustness checks for these results in the Online Appendix. In Table A.2 I show the same results using only not-yet treated cells as the control group, omitting the never-treated cells entirely from the analysis. The results are qualitatively the same, with malaria prevalence having a smaller but still statistically significant magnitude and NDVI having a higher magnitude. In Table A.3 I use the same control groups as Table 1 but I only include areas where the coverage-land ratio is below 0.1 (instead of 0.2), and where coverage is in the upper decile of distance from the station (instead of upper quintile). All results have the same sign and are similar in magnitude, but with less power due to the smaller number of observations. In Table A.4 I show results using a relaxed maximum elevation change criterion of 100m and find very similar results. In Table A.6 I present results using only commercial stations, as they are the focus of the counterfactual analysis. These regressions exclude all cells that have coverage from any public station and is based on a smaller sample. Thus they are based only on the coverage streaks of commercial stations. All results have the same sign as the baseline results. The results for malaria prevalence and nighttime luminosity are larger in magnitude and remain statistically significant, while the results for NDVI are smaller in magnitude and lose statistical significance. Finally, in Table A.5 in the I perform a placebo test where I randomly draw a different year of first receiving coverage for each location. As expected, in each case the effect is small and not statistically significant.

4 Structural Model of Radio Station Entry

With estimates of the positive effects of radio coverage on health and economic outcomes, we are now interested in how regulation and competition affect the entry decisions of commercial radio stations. I now introduce and estimate a dynamic structural model of radio station entry which I will then use to simulate counterfactual regulation schemes to evaluate their impacts on these health and economic outcomes.

Model

Setup: The country is made up of L equal-sized locations indexed by ℓ. Location ℓ has a population $M_{\ell t}$ in time period t and has a dominant language $\kappa_\ell \in \mathcal{K} = \{1, \ldots, K\}$ which is fixed over time. Throughout the country there is a set $\mathcal{F} = \{1, \ldots, F\}$ of potential stations. Each potential station f is associated with a particular location $\widetilde{\ell}_f$. If station f is actively broadcasting, its coverage will reach location $\widetilde{\ell}_f$ and surrounding locations, depending on the hilliness of the surrounding terrain. We use the indicator $\gamma_{f\ell} \in \{0, 1\}$ to denote if station f could be heard in location ℓ if it were on air.[9]

Stations are either commercial stations or public stations and we use the indicator $b_f \in \{0, 1\}$ to indicate if station $f \in \mathcal{F}$ is a public station. Commercial stations broadcast in a mix of the lingua franca (English) and a local language denoted by $\widetilde{\kappa}_f \in \mathcal{K}$. Because public stations have a wider coverage area, they broadcast in the lingua franca and several different local languages, and do not broadcast in a single dominant local language. We denote this by $\widetilde{\kappa}_f = 0$.

Station Life Cycle: Each potential station f in period t is in one of four different phases, denoted by $\varsigma_{ft} \in \{1, 2, 3, 4\}$, where $\boldsymbol{\varsigma}_t = (\varsigma_{1t}, \ldots, \varsigma_{Ft})$ denotes the phases of all potential stations in period t. These phases are: (1) unauthorized and off air, (2) authorized but off air, (3) on air, and (4) exited. I illustrate these phases in Figure 4. To go on air, a station must obtain a license and undergo setup. We denote by $a_{ft} \in \{0, 1\}$ the action of station f in phase 1 in period t, where $a_{ft} = 1$ denotes obtaining a license and $a_{ft} = 0$ denotes choosing the outside option. If a station chooses the outside option, it perishes and another potential station takes its place in the following period. If a station obtains a license, they begin setup.

[9] The total number of potential entrants, F, includes all 393 commercial and public stations that actually entered plus an additional 256 potential stations that did not.

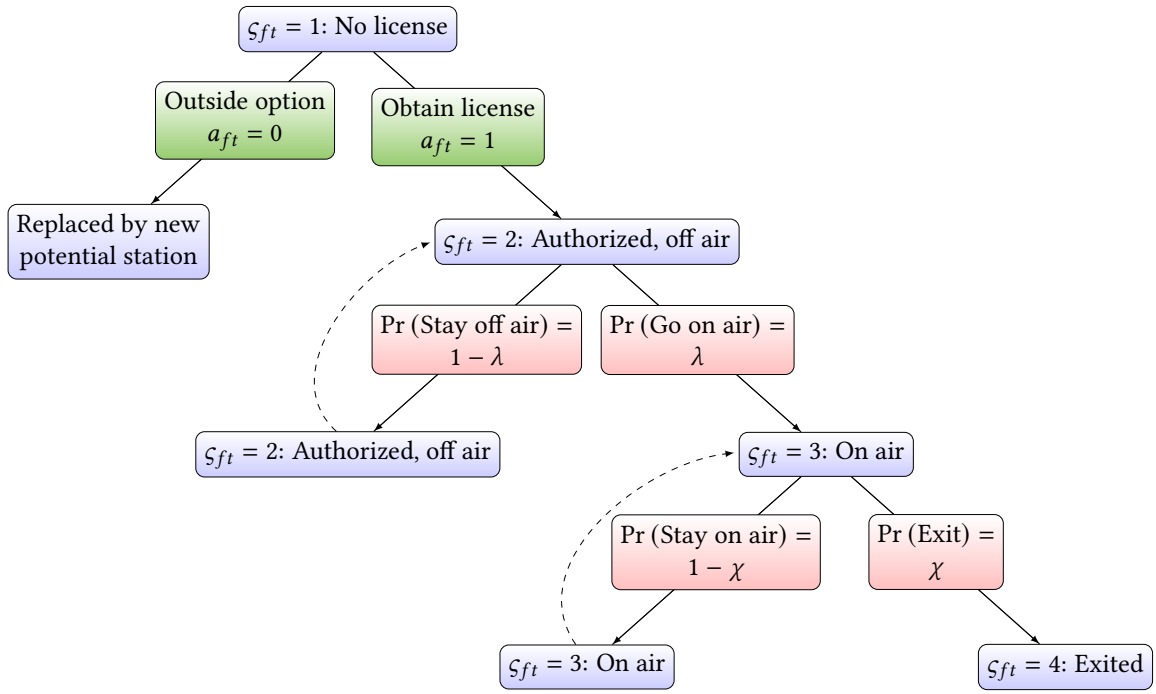


FIGURE 4: Life cycle of a commercial radio station in the model.

For commercial stations, there is some uncertainty for the station about how long this setup will take before it can go on air. Each period an authorized but not yet active commercial station goes on air with probability λ. With probability $1 - \lambda$, the station has not completed setup and remains in phase $\varsigma_{ft} = 2$. This parameter captures different market frictions that cause delays in the setup process. Thus the probability that the station goes on air j years after authorization is $(1 - \lambda)^j \lambda$. With probability λ, a commercial station will begin broadcasting immediately after authorization.

Authorized and on-air commercial stations exogenously exit and go off air with probability χ each period. When this occurs they enter phase $\varsigma_{ft} = 4$. Exited stations cannot re-enter. Stations can only proceed to a higher phase and cannot return to a lower phase. Public stations do not have a setup phase and do not exit, and thus are only ever in either phase 1 (unauthorized and off air) or 3 (on air). Thus for public stations it is as if $\lambda = 1$ and $\chi = 0$.

Listenership: Each time period t is divided into a number of subintervals τ where consumers make radio-listening decisions. At time τ in period t, consumer i in location ℓ decides to listen to the radio with probability μ_t. With probability $1 - \mu_t$ they

do another activity. The consumer can only listen to a station if the station is both on air ($\varsigma_{ft} = 3$) and its coverage reaches location ℓ ($\gamma_{f\ell} = 1$). We denote by $\mathcal{J}_{\ell t}\left(\boldsymbol{\varsigma}_t\right) = \left\{f \in \mathcal{F} : \varsigma_{ft} = 3 \text{ and } \gamma_{f\ell} = 1\right\}$ as the set of stations that are available to be listened to in location ℓ at time t. They obtain the following utility from listening to station $j \in \mathcal{J}_{\ell t}\left(\boldsymbol{\varsigma}_t\right)$ at that moment:

$$u_{if\ell t\tau} = \delta_{f\ell} + v_{if\ell t\tau} \tag{2}$$

where

$$\delta_{f\ell} = \theta^{Out} \mathbb{1}\left\{\widetilde{\kappa}_f \neq \kappa_\ell\right\}\left(1 - b_f\right) + \theta^{Pub} b_f \tag{3}$$

and where $v_{if\ell t\tau}$ is a Gumbel-distributed taste shock for consumer i at time τ. The mean utility for commercial stations broadcasting in the local language of the area (when $b_f = 0$ and $\widetilde{\kappa}_f = \kappa_\ell$) is normalized to zero. Consumers can have different mean utility for commercial stations broadcasting in a different language ($b_f = 0$ and $\widetilde{\kappa}_f \neq \kappa_\ell$) and for public stations ($b_f = 1$). Due to the Gumbel-distributed taste shocks, the listening share of station f in location ℓ throughout the period t is then:

$$\mathfrak{s}_{f\ell t}\left(\boldsymbol{s}_t\right) = \mu_t \frac{\mathbb{1}\left\{\varsigma_{ft} = 3\right\} \gamma_{f\ell} \exp\left(\delta_{f\ell}\right)}{\sum_{f' \in \mathcal{J}_{\ell t}(\varsigma_t)} \exp\left(\delta_{f'\ell}\right)} \tag{4}$$

where $\boldsymbol{s}_t$ is market state at time t, which includes the phase of each station, the population at each point in the country, and the time period: $\boldsymbol{s}_t = \left(\boldsymbol{\varsigma}_t, \{M_{\ell t}\}_{\ell=1}^{L}, t\right)$. The listening share for a station in locations where it is off air ($\varsigma_{ft} \neq 3$) or where its coverage does not reach ($\gamma_{f\ell} = 0$) is zero.[10]

Station Profits: If a potential entrant obtains a license (chooses $a_{ft} = 1$), they pay an entry cost equal to ε_{ft}, which includes the cost of the license and a setup fee. All stations share a common discount factor equal to $\beta \in (0, 1)$. The entry cost has a common distribution for all stations with cdf $F_\varepsilon\left(\varepsilon_{ft}\right)$. The value of ε_{ft}, however, is privately observed by the potential station. After authorization the station begins the following period as an authorized but off-air station with probability $1-\lambda$, or as an active station with probability λ. An authorized station that is not on air ($\varsigma_{ft} = 2$) earns zero profits. Once a station is on

[10] A location with a single on-air commercial station broadcasting in the local language would have a listening share of μ_t. In a location with two such stations, each would have a listening share of $\frac{\mu_t}{2}$. In a location with one local-language commercial station and one public station, the commercial station would have a listening share of $\frac{\mu_t}{1+e^{\theta_2}}$ and the public station $\frac{\mu_t e^{\theta_2}}{1+e^{\theta_2}}$.

air ($\varsigma_{ft} = 3$), it earns profits equal to:

$$\pi_f(\boldsymbol{s}_t) = \mathbb{1}\left\{\varsigma_{ft} = 3\right\} \alpha_t \sum_{\ell=1}^{L} M_{\ell t} \mathfrak{s}_{f\ell t}(\boldsymbol{s}_t) - FC \tag{5}$$

where α_t is the monetary value of an annual listener in period t (which I call the monetization rate) and FC are the station's fixed costs.[11] Active stations exogenously exit and go off air with probability χ each period. Exited stations ($\varsigma_{ft} = 4$) earn zero profits.

The timing of the model within a period is as follows: (1) potential stations observe their private entry costs and make entry decisions, (2) authorized but not-yet-active stations become active with probability λ and already-active stations exit with probability χ, (3) consumers realize their taste shocks for active stations, $\nu_{if\ell t\tau}$, and make listening choices, and (4) active stations earn profits.

Equilibrium Entry Strategies: I assume that stations know the process governing the evolution of population in each location over time, as well as the entry moments of public stations. A station's strategy function is denoted by $\sigma_f(\boldsymbol{s}_t, \varepsilon_{ft}) \in \{0, 1\}$ which maps the market state and the private entry cost into an action a_{ft}. Furthermore, $\boldsymbol{\sigma}_{-f}(\boldsymbol{s}_t, \boldsymbol{\varepsilon}_{-f,t}) = \left\{\sigma_j(\boldsymbol{s}_t, \varepsilon_{jt})\right\}_{j \in \mathcal{F} \setminus f}$ denotes the strategy function of all other potential stations $j \neq f$. To ease notation, this will be written as $\boldsymbol{\sigma}_{-f,t}$.

To define optimal strategies, we first need to define the value function. There is a terminal period T where the market state no longer changes. In the terminal period, the value function for an active station ($\varsigma_{ft} = 3$) is:

$$v_f^3(\boldsymbol{s}_T) = \frac{\pi_f(\boldsymbol{s}_T)}{1 - \beta} \tag{6}$$

whereas for other types of stations (unauthorized, inactive or exited), the terminal period value function is $v_f^1(\boldsymbol{s}_T) = v_f^2(\boldsymbol{s}_T) = v_f^4(\boldsymbol{s}_T) = 0$.

For periods $t < T$, we first define the following functions. The value of being an active

[11] This specification also includes the profits from within coverage streaks used in the reduced-form analysis. Although I argue there that stations are unlikely to strategically place their masts to capture particular areas in their coverage streaks, I include those areas in their profit function here for simplicity and completeness. Because after estimation the profits from coverage streaks is very small, this assumption is likely to have little impact on the counterfactual analysis.

station is given by:

$$v_f^3\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}\right) = (1-\chi)\Big(\pi_f\left(\boldsymbol{s}_t\right) + \beta\mathbb{E}\left[v_f^3\left(\boldsymbol{s}_{t+1}, \boldsymbol{\sigma}_{-f,t}\right)\middle|\boldsymbol{s}_t, \varsigma_{ft+1} = 3, \boldsymbol{\sigma}_{-f,t}\right]\Big) \tag{7}$$

With probability $1-\chi$, the station remains active and receives profits $\pi_f\left(\boldsymbol{s}_t\right)$ and begins the following period as an active station. With probability χ, the station exits and receives a payoff of $v_f^4\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}\right) = 0$ forever. The value of being an authorized but inactive station ($\varsigma_{ft} = 2$) is given by:

$$\begin{aligned} v_f^2\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}\right) =& \lambda\left(\pi_f\left(\boldsymbol{s}_t\right) + \beta\mathbb{E}\left[v_f^3\left(\boldsymbol{s}_{t+1}, \boldsymbol{\sigma}_{-f,t+1}\right)\middle|\boldsymbol{s}_t, \varsigma_{ft+1} = 3, \boldsymbol{\sigma}_{-f,t}\right]\right) + \\ & (1-\lambda)\,\beta\mathbb{E}\left[v_{ft+1}^2\left(\boldsymbol{s}_{t+1}, \boldsymbol{\sigma}_{-f,t+1}\right)\middle|\boldsymbol{s}_t, \varsigma_{ft+1} = 2, \boldsymbol{\sigma}_{-f,t}\right] \end{aligned} \tag{8}$$

With probability λ, the station becomes active and earns profits but with probability $1-\lambda$ it remains inactive. Finally, the value of being a potential entrant ($\varsigma_{ft} = 1$) is given by:

$$v_{ft}^1\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}, \varepsilon_{ft}\right) = \max\left\{0, v_{ft}^2\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}\right) - \varepsilon_{ft}\right\} \tag{9}$$

The potential entrant chooses the maximum of the outside option (0) and entering. The probability of a station entering is then $F_\varepsilon\left(v_{ft}^2\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}\right)\right)$.

Because the number of different possible states $\boldsymbol{s}_t \in \mathcal{S}$, is very large, I assume that potential stations form their strategies based on a summary statistic of the current market state, similar to the moment-based Markov equilibrium approach of Ifrach and Weintraub (2017). I assume that potential stations form entry strategies based on the counterfactual profits the station would earn if it were currently on air in the current market state. I denote these counterfactual profits by the function $\widetilde{\pi}_f\left(\boldsymbol{s}_t\right)$. This assumption restricts stations to have the same strategy at all market states that would give the station the same profits if they were active. For example, suppose a potential entrant could have one rival with a large overlapping coverage area or two rivals with smaller overlaps but both scenarios leading to the same flow profits if active. In both scenarios the potential entrant would experience the same intensity of competition but with a different number of rivals. This assumption restricts the station's entry strategy to be the same in both scenarios.

In addition to this, I further assume that the stations' strategy functions are symmetric within a time period. The simplified strategy function is then denoted by $\widetilde{\sigma}_t\left(\widetilde{\pi}_f\left(\boldsymbol{s}_t\right), \varepsilon_{ft}\right)$

with the collection of other firms' strategies denoted by $\widetilde{\boldsymbol{\sigma}}_{-f,t}$. In a symmetric equilibrium, for each potential station f and for each time period t, each station chooses the optimal action given their beliefs of the rivals' strategies, and each station's beliefs are consistent with the rivals' actual strategies:

$$\widetilde{\sigma}_t\left(\widetilde{\pi}\left(\boldsymbol{s}_t\right), \varepsilon_{ft}\right) = \underset{a_{ft}\in\{0,1\}}{\arg\max}\ a_{ft}\left(v_{ft}^2\left(\boldsymbol{s}_t, \widetilde{\boldsymbol{\sigma}}_{-f,t}\right) - \varepsilon_{ft}\right) \tag{10}$$

Structural Model Estimation

Parameterization: For the set of locations $1, \ldots, L$ I take the entire country and divide it into a grid with a 5 arc-minute resolution. This means each location is approximately 9000m×9000m in size. I use a coarser resolution compared to Section 3 to ease the computation burden in estimation. For each station, I take the average signal strength in each location and denote the station as having coverage in a location ($\gamma_{f\ell}$) if the average signal strength in the location exceeds the 60 dBμV/m threshold. For the language of a location, κ_ℓ, I take the dominant language of the centroid of location ℓ using the Glottolog data.

For the public stations in the model ($b_f = 1$), I use stations licensed as either "Public" or "Public (Foreign)". Because there is relatively less entry of public stations I assume that these enter deterministically and the entry dates are known by the potential commercial stations. For the set of potential entrants, I use all the commercial stations that ever entered in my data, as well as additional potential entrants in all populated areas in the country. To do this, I assume there is an additional potential entrant in every area with positive nighttime luminosity values. I cluster contiguous areas with positive nighttime luminosity values and for each cluster I set the potential entrant's coordinates to be the centroid of the cluster.[12] This approach leads to 256 additional potential entrants, in addition to the 352 commercial stations in the data. In the Online Appendix, I illustrate this approach in Figure A.8 and show the full set of station locations in Figure A.9.

I use the NASA SEDAC rasterized population data for $M_{\ell t}$, the population of location ℓ in time period t. Because the population data are only available at five-year intervals, I use exponential interpolation to obtain the population between the intervals. I assume the population of each location evolves deterministically in the model.

[12]The potential listenership (population within the coverage area) from broadcasting from the cluster's centroid is very similar compared to broadcasting from other parts of the cluster. I computed the potential listenership at all points in a fine grid in two clusters and found very little variation. The coefficient of variation in each case was 1.5% and 1.6%. I therefore choose to set the potential entrant locations at the centroid for simplicity.

I estimate the probability of transitioning from being authorized to active, λ, and the exogenous probability of exit, χ, using the observed transitions in the data. I estimate these before estimating the structural parameters. To estimate λ, I take the average number of quarters it takes for newly-authorized stations to become active using the NCA reports. After transforming to an annual rate I find $\lambda = 0.578$. In Figure A.7 in the Online Appendix I show a histogram of the time to become active from the date of authorization. In the data, there are no significant differences in the average setup time across regions, so I treat this probability as exogenous and assume it is equal for all potential entrants. For χ, I take the average quarterly exit rate observed in the data and transform it to an annual rate. I find $\chi = 0.011$.

Station variable profits are proportional to the product of the monetization rate, α_t, and the listening probability, μ_t. These are not separately identified in my model. Furthermore, I also need to normalize the product $\alpha_t \mu_t = 1$ relative to one time period. Therefore I parameterize their product as $\alpha_t \mu_t = \exp\left(\theta^{\alpha\mu} t\right)$ where the time trend can account for both increases in radio ownership and the monetization rate over time.

I set the annual discount factor to $\beta = 0.9$, following Ryan (2012) and Igami and Yang (2016). I assume the entry cost distribution $F_\varepsilon\left(\varepsilon_{ft}\right)$ is log-normal, where $\log\left(\varepsilon_{ft}\right) \sim \mathcal{N}\left(\theta^{EC}, \sigma_\varepsilon^2\right)$. As is common in discrete choice models, the variance of the entry cost shocks is not separately identified and I set it to $\sigma_\varepsilon^2 = 1$. Furthermore, in models without endogenous exit, the fixed cost, FC, is not separately identified from the entry cost. I therefore assume that the entry cost parameter, θ^{EC}, also contains the expected present discounted value of future fixed costs.[13] The full vector of parameters to be estimated in the structural model is then $\boldsymbol{\theta} = \left(\theta^{Out}, \theta^{Pub}, \theta^{\alpha\mu}, \theta^{EC}\right)$. With these parameterizations, the flow profits for a station can be expressed as:

$$\pi_f\left(\boldsymbol{s}_t\right) = \mathbb{1}\left\{\varsigma_{ft} = 3\right\} e^{\theta^{\alpha,\mu} t} \sum_{\ell=1}^{L} M_{\ell t} \frac{\gamma_{f\ell} e^{\theta^{Out}\mathbb{1}\left\{\widetilde{\kappa}_f \neq \kappa_\ell\right\}\left(1-b_f\right)+\theta^{Pub} b_f}}{\sum_{f' \in \mathcal{J}_{\ell t}\left(\varsigma_t\right)} e^{\theta^{Out}\mathbb{1}\left\{\widetilde{\kappa}_{f'} \neq \kappa_\ell\right\}\left(1-b_f\right)+\theta^{Pub} b_{f'}}} \tag{11}$$

Estimation: I do not solve for the equilibrium for each trial parameter vector in estimation as solving for an equilibrium is very time intensive. I therefore use an approach that side-steps the equilibrium computation in estimation, similar to the two-step approach in (Bajari et al., 2007).

[13] Because of the random setup time and exogenous exit rate, the expected present discounted value of future fixed costs is given by $\sum_{\tau=0}^{\infty}(1-\lambda)^\tau \lambda \beta^\tau \sum_{s=0}^{\infty} \beta^s (1-\chi)^s FC = \lambda FC / [(1-(1-\lambda)\beta)(1-(1-\chi))\beta]$.

Given a guess of the parameter vector $\boldsymbol{\theta}$, I first compute the counterfactual profits, $\widetilde{\pi}_f(\boldsymbol{s}_t)$, each potential entrant would have if active, given the observed phases, ς_{ft}, of all other stations in the data. I then fit a reduced-form equilibrium policy function by with a logistic regression using the observed entry decisions and the counterfactual profits at the trial parameter, $\boldsymbol{\theta}$. I also include a trend term in this logistic regression to account for the fact that the equilibrium policy function can differ in each time period. I show the estimates from this policy function at the estimated structural parameters in Table A.7 in the Online Appendix, which shows that rivals are more likely to enter when when counterfactual profits are higher and in later time periods.

For each potential entrant in each time period, I then forward simulate the profits the station would earn each period if it entered, where I use the reduced-form equilibrium policy function to draw entry decisions for rival entrants. I simulate forward 20 periods for each station and average over 2,500 paths. For each path, I draw new phases for each station using the estimated λ and χ. This simulation approach gives an estimate of $v_f^2\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}\right)$ in equation (8) given parameters $\boldsymbol{\theta}$. I denote this by $\widehat{v}_f^2\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}, \boldsymbol{\theta}\right)$. Using this together with the trial entry cost θ^{EC} gives the probability of observing the potential entrant's action, $a_{ft} \in \{0, 1\}$, at the parameter vector $\boldsymbol{\theta}$:

$$\Pr\left(a_{ft} \middle| \boldsymbol{\theta}\right) = \Phi\left(\left(2a_{ft} - 1\right)\left(\log\left(\widehat{v}_f^2\left(\boldsymbol{s}_t, \boldsymbol{\sigma}_{-f,t}, \boldsymbol{\theta}\right)\right) - \theta^{EC}\right)\right) \tag{12}$$

I estimate the structural parameters via maximum likelihood using these conditional choice probabilities:

$$\widehat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \sum_{f \in \left\{f' \in \mathcal{F} : b_{f'} = 0\right\}} \sum_{t \in \mathcal{T}_f} \log\left(\Pr\left(a_{ft} \middle| \boldsymbol{\theta}\right)\right) \tag{13}$$

where $\mathcal{T}_f$ denotes the set of time periods where potential station f was in phase 1 in the data.

The listener utility parameters, θ^{Out} and θ^{Pub}, are identified through differential entry patterns in the cross section where, holding the population within their coverage area fixed, entrants are more or less likely to enter when faced with different degrees of competition from commercial stations from inside and outside their language area and public stations. The parameter over the product of listenership and monetization, $\theta^{\alpha\mu}$, is identified through the increased entry over time given the amount of competition, beyond that explained by increasing population. Finally the mean entry cost parameter, θ^{EC}, is identified through the share of potential entrants entering in the data.

Parameter	Estimate	Standard Error
Commercial station not in local language, θ^{Out}	−2.806	(0.150)
Public station, θ^{Pub}	−0.404	(0.150)
Time trend, $\theta^{\alpha\mu}$	0.114	(0.000)
Entry cost mean, θ^{EC}	6.336	(0.033)

Table 2: Structural Parameter Estimates.

Parameter Estimates

Table 2 shows the structural parameter estimates. The first two parameters are the estimates of the listeners' mean utilities for different types of stations, as in equation (3). Commercial stations broadcasting from the local language have a normalized mean utility of 0. Because both parameters are negative, consumers obtain the highest mean utility from local commercial stations. Commercial stations from outside the local language have a large negative value, whereas public stations have a value closer to zero, indicating consumers prefer public stations to these. The time trend is positive, indicating that the product of the listening share and monetization of listeners is increasing over time.[14] Because the entry cost parameter changes with the number of additional potential entrants, it cannot be given a monetary interpretation.

In the Online Appendix I show plots that demonstrate the model's fit. To do this, I solve the model via backward induction at the estimated model parameters. Using the resulting equilibrium strategy functions I simulate 1,000 different paths of the industry over time.[15] In the left panel of Figure A.10 I show the average number of newly authorized commercial stations by year from these simulations compared to the data. The model is unable to match the year-by-year noise in entry but follows the overall trend reasonably well, although slightly underpredicting entry on average. In the right panel I show the average proportion of the country's population served by commercial stations each year from the simulations compared to the data. Again, this matches the data reasonably well except slightly underpredicts coverage in the early years. After 2000 until 2015 the model matches the data well. In Figure A.11 I show maps of the total number of stations across

[14] At the firm states, ς_{ft}, in the data, the population within firms' coverage areas explain 34.64% of the variation in profits, while the time trend only explains 10.46%. Thus population explains relatively more of the variation in profits.

[15] The model I estimate is not guaranteed to have a unique equilibrium. I have, however, tested the robustness of the model's predictions by solving the equilibrium with different initial guesses and in each case the predictions are virtually identical.

time. The model predicts the most entry occurring in the south of the country and near Kumasi, as we observe in the data. However, in the data more stations enter in these larger cities earlier in the sample period than the model predicts. It should be noted that the estimation approach avoids computing the equilibrium of the game and estimates the parameters by simulating the value functions forward given a reduced-form estimate of the equilibrium strategies, whereas the equilibrium computation involves solving the game via backward induction. The fact that the model predictions match the data reasonably well is evidence of the validity of the estimation strategy.

5 Improving Access to Radio

I now combine the reduced-form results from Section 3 with the estimated structural model from Section 4 to investigate which alternative regulation schemes could improve access to radio coverage, and as a result improve the health and economic outcomes of areas receiving coverage.

Increasing Transmitter Strengths

Commercial radio stations in Ghana are restricted to a 45km broadcasting radius. In rural areas, this may not make it worthwhile to enter if the number of listeners is not large enough. The coverage from stations operating in larger towns will also not spill out further into rural areas if transmitter strengths are restricted. One way to increase access to coverage in rural areas would be to allow commercial stations to have a larger broadcasting radius. The effect of such a policy on entry is not obvious ex-ante. On the one hand, a larger radius increases the listenership of a station which increases profits. However, this could also increase the amount of competition a station faces, which would decrease the station's listener share and decrease profits. The overall effect on a station's number of listeners is therefore ambiguous.

To run this counterfactual experiment, I first I calculate what each entrant's coverage area would have been if they were allowed to use transmitters that were 50% stronger. This results in new coverage indicators $\gamma_{f\ell}$ for each commercial station, which can affect a station's number of potential listeners. The median station's coverage area increases by 50%, but this change can differ across stations depending on the surrounding terrain. Using these new values, I solve for the new equilibrium strategies of the players. As argued

in the introduction, the total cost increase from higher transmission strengths is likely to be negligible and therefore I assume stations' costs are unchanged in the counterfactual.

I show the results from this counterfactual in Figure 5.[16] The left panel shows that with stronger transmitters, more stations enter compared to the baseline. This implies that the effect of greater potential listenership dominates the effect of more competition for most stations.[17] The right panel shows that a greater share of the population is served under the stronger transmitter strength policy. By 2015, the percentage of the population served by radio increases by 5% to 87.9%.[18] In Figure A.14 in the Online Appendix, I also show the same results but split by population density terciles. These show that the policy had little impact on entry in urban areas, but encouraged entry in the rural areas.[19]

We can now combine these results with the reduced form results from Section 3 to evaluate the effect of the transmitter strength policy on malaria among children, night-time luminosity growth and agricultural productivity. According to the 2010 census in Ghana (Ghana Statistical Service, 2012), approximately 23% of the population is in the 2-10 age range. This is the age range at which the reduced-form malaria estimates are based on. Therefore, assuming the proportion of children in the newly-covered areas is the same as the rest of the country, there are an estimated 171,273 children in the new areas that receive coverage from this policy in the year 2015. Based on this number and the estimated effect of coverage on the malaria parasite rate found in Section 3, there are a predicted 2,740 fewer infections among children in 2015. Although this number is small for a country with over 5 million children, there are also the large economic costs of malaria to consider.[20]

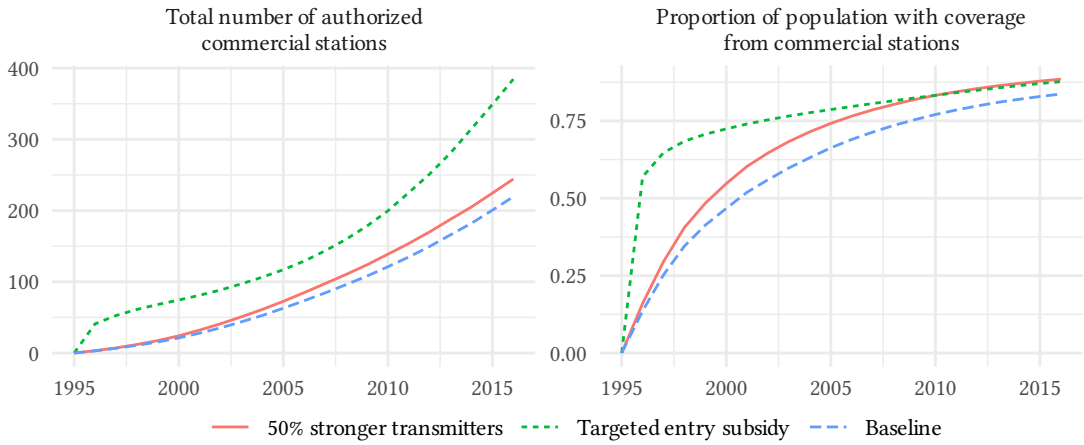
[16]For comparison purposes, I also repeat this exercise for a 100% increase and a 50% decrease. These are shown in Figure A.12 in the Online Appendix. The 50% decrease in signal strength has a negative impact on entry and coverage which is larger in magnitude than the 50% strength increase. 22.3% fewer stations enter and the percentage of population with coverage in 2015 falls to 71.2%. Compared to the 50% increase, the 100% increase leads to more entry of stations but has a similar impact on the percentage of population with coverage.

[17]In Section A.7 in the Online Appendix I discuss how the population in stations' coverage areas and competition change under the policy. I also discuss the impact of a station unilaterally increasing its signal strength on profits.

[18]An increase in the number of stations entering together with wider coverage areas could imply frequency congestion. However, this is unlikely for the following reasons. In Accra, the nation's capital and most congested area, there are only 2 more stations active compared to the baseline scenario which still had room to allocate additional frequencies during the sample period. The total number of stations across the county increases by only 9.5%, and the increase in percentage area served is only 4 percentage points.

[19]In the Online Appendix, Figure A.15 shows the results of the counterfactuals by geography and Figure A.16 shows the results in histograms of the number of stations available in each location.

[20]The WHO estimate that the case fatality rate for malaria is between 0.01%-0.4%. Thus the estimated reduction in fatalities is between 0.27-10.96 in 2015. Typical estimates of the value of a statistical life are



Average across 1,000 simulations.

FIGURE 5: Counterfactual simulation results.

By 2015, an additional 4 percentage points of the country's land area receives coverage from the policy. In Section 3 we found that coverage increases local GDP measured by nighttime luminosity by 1.7%. Thus the policy increases GDP in 2015 by 0.07%, or $33.65m. We also found that coverage increases local agricultural productivity measured by NDVI by 1.1%. Thus the policy contributes to a modest 0.042% increase in agricultural productivity in 2015.

The reduced-form analysis estimates the effects of coverage using coverage streaks which are typically rural areas. Thus these effects are likely local to rural areas. I use these effects here for areas newly served by radio which are also rural. The average population density of areas included in the reduced-form analysis is 47.3% of the national average. The population density of newly served areas in 2015 under the policy, which we use to compute the impacts, is 58% of the national average. Thus the areas are comparable in terms of how rural they are.

Radio coverage may also bring additional benefits to these communities such as increased political accountability, parental investment in education, women's status in the household, and other effects found in previous literature. This policy also does not involve the policymaker using public funds to subsidize the stations. Instead, the larger broadcasting radius incentivizes firms to enter by itself.

several million dollars in magnitude. Therefore the policy has a large impact in monetary terms.

Targeted Entry Cost Subsidies

An alternative approach to achieve the same level of access to coverage would be to provide entry subsidies to stations. Because more populated areas would become served regardless of the subsidy, a more effective approach than a universal subsidy would be to provide a subsidy only to stations entering into unserved markets. For this counterfactual experiment I consider a once-off subsidy that is proportional to the number of newly-served individuals from the station entering. I consider subsidies of different magnitudes and find the subsidy that delivers the same share of population served in 2015 as the transmitter strength policy. Because stations are required to pay licensing fees, this subsidy can also be considered as a reduction in the licensing fees for stations entering in unserved areas. Because the policy is designed to have the same proportion of the population served in 2015 as the 50% stronger transmitter policy, by construction it has the same impacts on malaria, growth and agricultural productivity in that year.

The effect of the resulting subsidy on entry and access is shown in Figure 5. Compared to the transmitter strength counterfactual, the subsidy encourages entry much earlier in the sample period. This is because stations want to enter ahead of rivals to avail of the subsidy: once a location is served later entrants cannot benefit from the subsidy. This also increases the proportion of the population served substantially earlier in the sample. This gap eventually decreases over time as stations in the baseline scenario find it profitable to enter as the product of population, listenership and monetization increase over time. The increased number of stations continues to the end of the sample period, but with the same share of population served by 2015 by construction.

Without detailed data on station profits it is difficult to provide a monetary value on the total cost of the subsidy. However, we can calculate the total cost of the subsidy as a proportion of total industry profits in the model. Over the entire sample period, the subsidy has a total cost of 4.2% of total industry profits. The subsidy is equivalent to receiving a once-off payment of 7.8% of the average annual profits of a station for each additional 10,000 individuals newly served by radio.

Figure A.13 in the Online Appendix shows the impact of subsidies of different magnitudes. From this we can see that there are diminishing returns to the size of the subsidy. In fact, a subsidy costing only 1.2% of total industry profits would result in 86.7% served by 2015, instead of 87.9% under the transmitter strength policy.

6 Conclusion

Radio has been found to have many beneficial effects on many parts of society, including political participation, education and health. This is especially important in developing economies where radio is more common and accessible compared to other forms of mass media. Therefore it is important for the broadcasting regulator to ensure licensing and broadcasting restrictions do not preclude certain rural areas from obtaining access to radio coverage, as it can bring a host of additional benefits to these communities.

Complementing the previous research on the effects of media, I find that radio coverage reduces malaria among children, increases growth (measured by nighttime luminosity), and increases agricultural productivity (measured by the proportion of green light in summer-month satellite images). I estimate these effects by exploiting coverage streaks spilling through gaps in hilly areas, where locations near the borders of these coverage streaks receive radio coverage in an as-if random fashion.

These positive effects of radio motivate the understanding of how commercial radio stations make their entry decisions, and how they respond to various forms of regulation. Regulations on transmitter strengths and entry costs could deter entry in rural areas, resulting in less information provision for these communities. Using a dynamic structural model of radio station entry, I simulate counterfactual policies with different regulation schemes that aim to increase access to coverage in rural areas. I find that allowing stations to increase their transmission power by 50% results in an increase of 5% of the population served by coverage by the end of the sample period. To achieve the same level of access with a targeted entry subsidy that subsidizes stations for the number of new individuals served by coverage would cost 4.2% of total industry profits over the sample period.

References

Armand, A., P. Atwell, and J. F. Gomes (2020): "The Reach of Radio: Ending Civil Conflict through Rebel Demobilization," *American Economic Review*, 110, 1395–1429.

Asher, S. and P. Novosad (2020): "Rural Roads and Local Economic Development," *American Economic Review*, 110, 797–823.

Bajari, P., C. L. Benkard, and J. Levin (2007): "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75, 1331–1370.

Berry, S., A. Eizenberg, and J. Waldfogel (2016): "Optimal Product Variety in Radio Markets," *The RAND Journal of Economics*, 47, 463–497.

Berry, S. T. and J. Waldfogel (1999): "Free Entry and Social Inefficiency in Radio Broadcasting," *The RAND Journal of Economics*, 30, 397–420.

Besley, T., R. Burgess, et al. (2002): "The Political Economy of Government Responsiveness: Theory and Evidence from India," *The Quarterly Journal of Economics*, 117, 1415–1451.

Bhatt, S., D. Weiss, E. Cameron, D. Bisanzio, B. Mappin, U. Dalrymple, K. Battle, C. Moyes, A. Henry, P. Eckhoff, et al. (2015): "The Effect of Malaria Control on Plasmodium Falciparum in Africa Between 2000 and 2015," *Nature*, 526, 207–211.

Bleakley, H. (2010): "Malaria eradication in the Americas: A retrospective analysis of childhood exposure," *American Economic Journal: Applied Economics*, 2, 1–45.

Cagé, J. (2019): "Media Competition, Information Provision and Political Participation: Evidence from French Local Newspapers and Elections, 1944–2014," *Journal of Public Economics*, 104077.

Callaway, B. and P. H. Sant'Anna (2021): "Difference-in-Differences with Multiple Time Periods," *Journal of Econometrics*, 225, 200–230.

Cheung, M. (2012): "Edutainment Radio, Women's Status and Primary School Participation: Evidence from Cambodia," Tech. rep., Stockholm University, Department of Economics.

De Chaisemartin, C. and X. d'Haultfœuille (2020): "Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, 110, 2964–96.

DellaVigna, S., R. Enikolopov, V. Mironova, M. Petrova, and E. Zhuravskaya (2014): "Cross-Border Media and Nationalism: Evidence from Serbian Radio in Croatia," *American Economic Journal: Applied Economics*, 6, 103–132.

Durante, R., P. Pinotti, and A. Tesei (2019): "The Political Legacy of Entertainment TV," *American Economic Review*, 109, 2497–2530.

Enikolopov, R., M. Petrova, and E. Zhuravskaya (2011): "Media and Political Persuasion: Evidence from Russia," *American Economic Review*, 101, 3253–3285.

Farré, L. and F. Fasani (2013): "Media Exposure and Internal Migration—Evidence from Indonesia," *Journal of Development Economics*, 102, 48–61.

Garcia-Arenas, J. (2016): "The Impact of Free Media on Regime Change: Evidence from Russia," *The Quarterly Journal of Economics*, 128, 105–164.

Gentzkow, M. (2006): "Television and Voter Turnout," *The Quarterly Journal of Economics*, 121, 931–972.

Gentzkow, M., J. M. Shapiro, and M. Sinkinson (2011): "The Effect of Newspaper Entry and Exit on Electoral Politics," *American Economic Review*, 101, 2980–3018.

Gerber, A. S., D. Karlan, and D. Bergan (2009): "Does the Media Matter? A Field Experiment Measuring the Effect of Newspapers on Voting Behavior and Political Opinions," *American Economic Journal: Applied Economics*, 1, 35–52.

Ghana Health Service (2017): "The Health Sector in Ghana: Facts and Figures 2017," Ghana Ministry of Health.

Ghana Statistical Service (2012): "2010 Population & Housing Census: Summary Report of Final Results," May 2012.

Gonzalez, R. M. (2021): "Cell Phone Access and Election Fraud: Evidence from a Spatial Regression Discontinuity Design in Afghanistan," *American Economic Journal: Applied Economics*, 13, 1–51.

Henderson, V., A. Storeygard, and D. N. Weil (2012): "Measuring Economic Growth from Outer Space," *American Economic Review*, 102, 994–1028.

Hidalgo, J. and M. Sovinsky (2022): "Subsidies, Speed and Switching? Impacts of an Internet Subsidy in Colombia," CEPR Discussion Paper No. DP17663.

Ifrach, B. and G. Y. Weintraub (2017): "A Framework for Dynamic Oligopoly in Concentrated Industries," *The Review of Economic Studies*, 84, 1106–1150.

Igami, M. (2017): "Estimating the Innovator's Dilemma: Structural Analysis of Creative Destruction in the Hard Disk Drive Industry, 1981–1998," *Journal of Political Economy*, 125, 798–847.

Igami, M. and N. Yang (2016): "Unobserved Heterogeneity in Dynamic Games: Cannibalization and Preemptive Entry of Hamburger Chains in Canada," *Quantitative Economics*, 7, 483–521.

Jensen, R. and E. Oster (2009): "The Power of TV: Cable Television and Women's Status in India," *The Quarterly Journal of Economics*, 124, 1057–1094.

Jeziorski, P. (2014a): "Effects of Mergers in Two-Sided Markets: The US Radio Industry," *American Economic Journal: Microeconomics*, 6, 35–73.

——— (2014b): "Estimation of Cost Efficiencies from Mergers: Application to US Radio," *The RAND Journal of Economics*, 45, 816–846.

——— (forthcoming): "Empirical Model of Dynamic Merger Enforcement–Choosing Ownership Caps in US Radio," *Management Science*.

Kafewo, S. (2006): "African Media Development Initiative: Ghana Context," https://ghana.mom-gmr.org/uploads/tx_lfrogmom/documents/19-476_import.pdf.

Kasampalis, S., P. I. Lazaridis, Z. D. Zaharis, A. Bizopoulos, S. Zettas, and J. Cosmas (2013): "Comparison of Longley-Rice, ITM and ITWOM propagation models for DTV and FM broadcasting," in *Wireless Personal Multimedia Communications (WPMC), 2013 16th International Symposium on*, IEEE, 1–6.

Kearney, M. S. and P. B. Levine (2015): "Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing," *American Economic Review*, 105, 3597–3632.

Keefer, P. and S. Khemani (2014): "Mass Media and Public Education: The Effects of Access to Community Radio in Benin," *Journal of Development Economics*, 109, 57–72.

La Ferrara, E., A. Chong, and S. Duryea (2012): "Soap Operas and Fertility: Evidence from Brazil," *American Economic Journal: Applied Economics*, 4, 1–31.

Larreguy, H., J. Marshall, and J. M. Snyder Jr (2020): "Publicising Malfeasance: When the Local Media Structure Facilitates Electoral Accountability in Mexico," *The Economic Journal*, 130, 2291–2327.

Michalopoulos, S. and E. Papaioannou (2013): "Pre-Colonial Ethnic Institutions and Contemporary African Development," *Econometrica*, 81, 113–152.

Oberholzer-Gee, F. and J. Waldfogel (2009): "Media Markets and Localism: Does Local News en Español Boost Hispanic Voter Turnout?" *American Economic Review*, 99, 2120–28.

Okuyama, Y. (2021): "Empowering Women Through Radio: Evidence from Occupied Japan," .

Olken, B. A. (2009): "Do Television and Radio Destroy Social Capital? Evidence from Indonesian Villages," *American Economic Journal: Applied Economics*, 1, 1–33.

Paluck, E. L. (2009): "Reducing Intergroup Prejudice and Conflict using the Media: a Field Experiment in Rwanda," *Journal of Personality and Social Psychology*, 96, 574.

Paluck, E. L. and D. P. Green (2009): "Deference, Dissent, and Dispute Resolution: An Experimental Intervention using Mass Media to Change Norms and Behavior in Rwanda," *American Political Science Review*, 103, 622–644.

Ryan, N. (2021): "The Competitive Effects of Transmission Infrastructure in the Indian Electricity Market," *American Economic Journal: Microeconomics*, 13, 202–42.

Ryan, S. P. (2012): "The Costs of Environmental Regulation in a Concentrated Industry," *Econometrica*, 80, 1019–1061.

Rysman, M., R. M. Townsend, and C. Walsh (2023): "Bank Branching Strategies in the 1997 Thai Financial Crisis and Local Access to Credit," CEPR Discussion Paper No. DP17869.

Snyder, J. M. and D. Strömberg (2010): "Press Coverage and Political Accountability," *Journal of Political Economy*, 118, 355–408.

Strömberg, D. (2004): "Radio's Impact on Public Spending," *The Quarterly Journal of Economics*, 189–221.

Sweeting, A. (2009): "The Strategic Timing Incentives of Commercial Radio Stations: An Empirical Analysis Using Multiple Equilibria," *The RAND Journal of Economics*, 40, 710–742.

——— (2010): "The Effects of Mergers on Product Positioning: Evidence from the Music Radio Industry," *The RAND Journal of Economics*, 41, 372–397.

——— (2013): "Dynamic Product Positioning in Differentiated Product Markets: The Effect of Fees for Musical Performance Rights on the Commercial Radio Industry," *Econometrica*, 81, 1763–1803.

Vaughan, Peter W., E. M. R. A. S. and R. M. Swalehe (2000): "Entertainment-education and HIV/AIDS prevention: A Field Experiment in Tanzania," *Journal of Health Communication*, 5, 81–100.

Vitali, A. (2022): "Consumer Search and Firm Location: Theory and Evidence from the Garment Sector in Uganda," Working Paper.

Wang, T. (2021): "Waves of Empowerment: Black Radio and the Civil Rights Movement," Working paper.

Wu, J., S. He, J. Peng, W. Li, and X. Zhong (2013): "Intercalibration of DMSP-OLS Nighttime Light Data by the Invariant Region Method," *International Journal of Remote Sensing*, 34, 7356–7368.

Yanagizawa-Drott, D. (2014): "Propaganda and Conflict: Evidence from the Rwandan Genocide," *The Quarterly Journal of Economics*, 129, 1947–1994.

Yordy, C. (2008): "The Economics of Rural Radio in Africa: An Introductory Study into the Costs and Revenues," African Farm Radio Research Initiative (AFRRI), Farm Radio International, Ottawa, Canada.

Online Appendix to:

The Social Impacts of New Radio Markets in Ghana: A Dynamic Structural Analysis

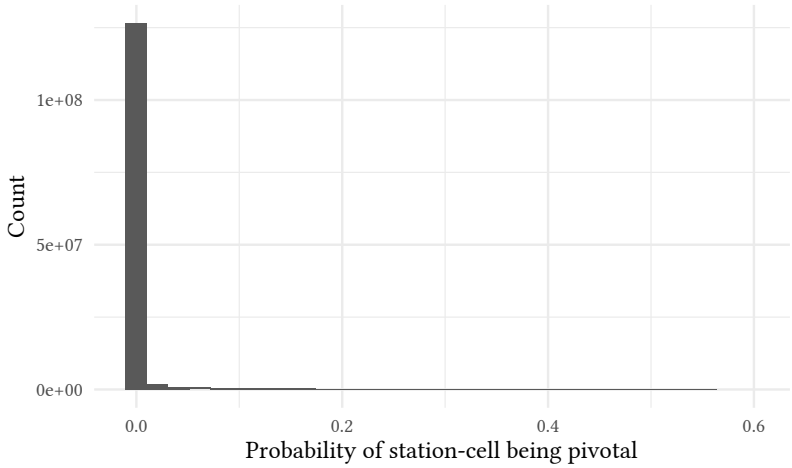
by Christoph Walsh

A.1 Additional Figures and Tables

	1993	1998	2003	2008	2014
% with radio	44.5	54.5	74.7	78.6	69.4
% with TV	16.8	25.4	31.0	47.2	64.2
% where respondent reads the newspaper	19.9	24.4	31.8	31.5	24.1
% with only radio	22.8	23.0	28.6	28.9	14.1
% with no radio, TV or newspaper	44.8	35.5	15.2	14.8	13.7

"% only radio" is the percentage of households in the wave that own a radio but do not own a TV nor read the newspaper.

Table A.1: Media access summary statistics from the Demographic Health Survey waves.



Let the estimated signal strength for station f in location (cell) ℓ be $\widehat{x}_{f\ell}$. For each station-cell combination I draw 1,000 new signal strength values according to $\widehat{x}_{f\ell} + \nu_{f\ell}$, where $\nu_{f\ell} \sim \mathcal{N}(-0.5, 30.25)$. These are the mean and variance of the errors found by Kasampalis et al. (2013). Using these simulated values I compute the probability that the signal strength crosses to the other side of the 60 dBμV/m threshold. The histogram displays the distribution of these probabilities.

Figure A.1: Histogram of probabilities that station-cells are pivotal relative to the 60 dBμV/m threshold.

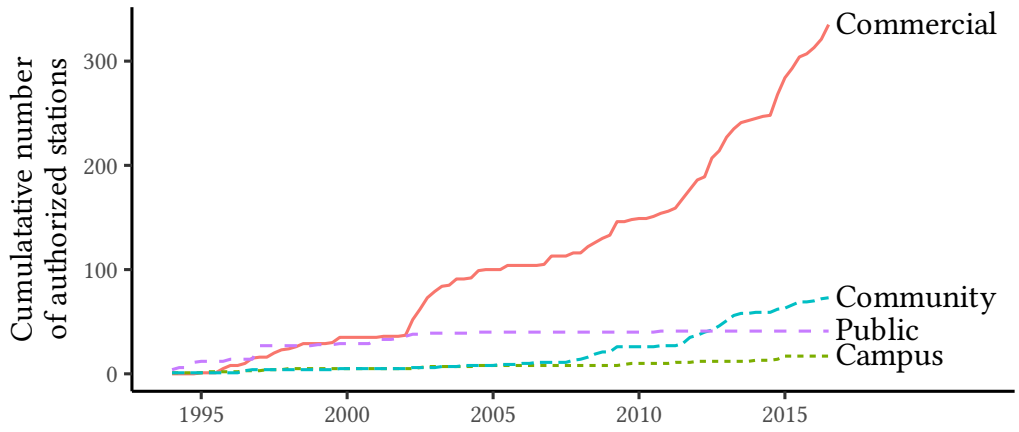


FIGURE A.2: Cumulative number of issued licenses by station type.

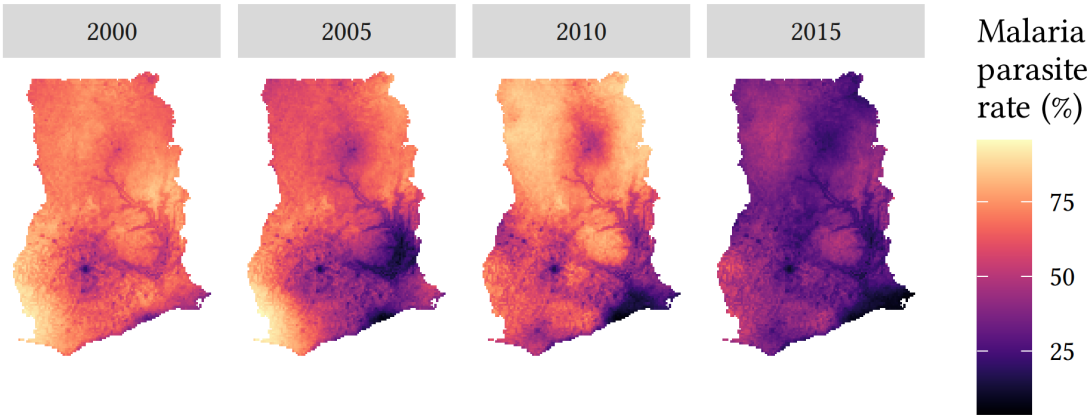
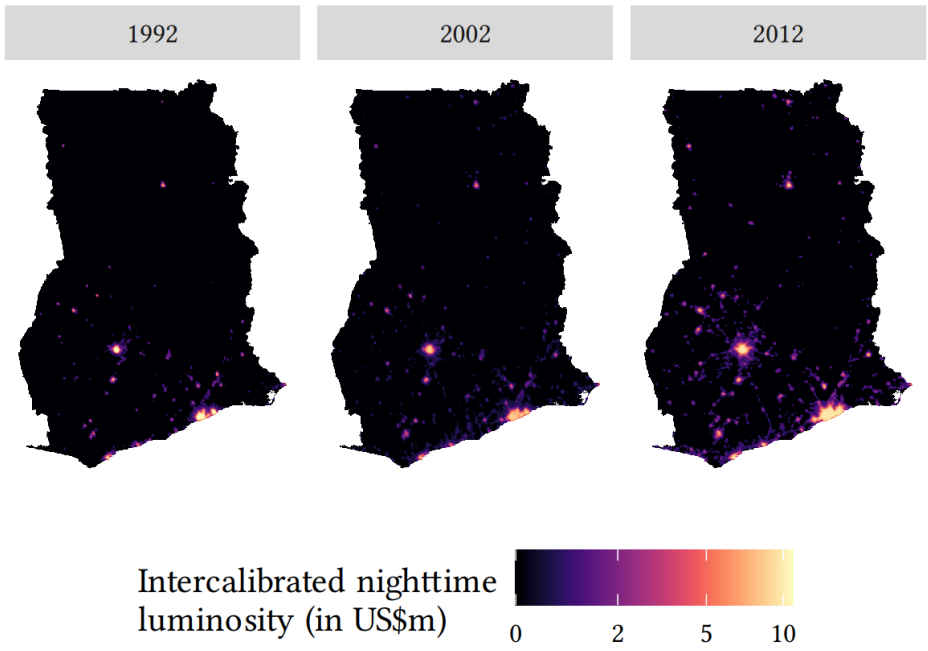


FIGURE A.3: Malaria parasite rate in 2-10 year olds over time.



The nighttime luminosity values are intercalibrated as follows. For each year, I find a scaling by dividing that year's national GDP by sum of values within the country's borders in that year. I then scale all values by this number for that year. Values represent GDP in a 900m×900m area.

Figure A.4: Intercalibrated nighttime luminosity in Ghana, 1992, 2002 and 2012.

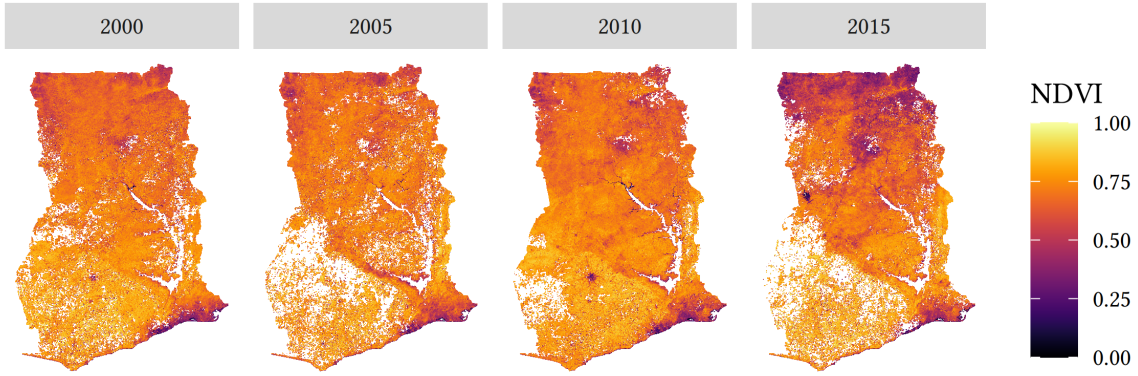


Figure A.5: Normalized Difference Vegetation Index (NDVI) in Ghana, 2000-2015. The maps show the average values during July-August of each year.

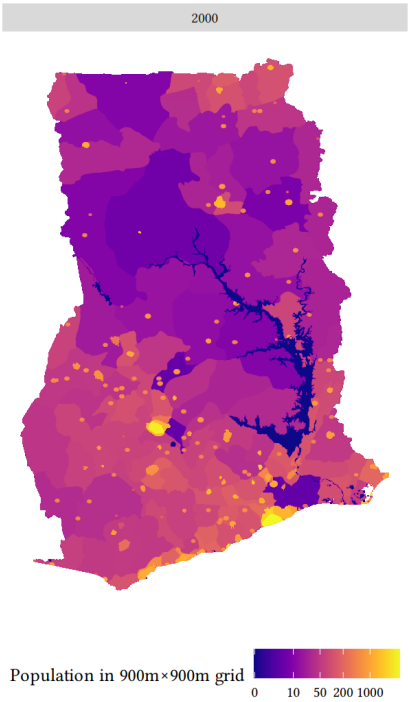


FIGURE A.6: Example NASA Socioeconomic Data and Applications Center population map.

	Dependent variable:		
	Malaria prevalence	Nighttime luminosity	Vegetation index
	(1)	(2)	(3)
Control group: Not yet received coverage only			
Coverage	−0.009	0.015	0.018
	(0.004)	(0.008)	(0.005)
Sample size	12000	20128	5472
Mean dependent variable	0.641	0.051	0.659

Estimates of the overall average treatment effect on the treated following Callaway and Sant'Anna (2021) using the outcome regression approach with cell fixed effects, year fixed effects, and weights based on group size. Robust standard errors clustered at the cell level are in parentheses. Nighttime luminosity is transformed using the inverse hyperbolic sign transformation.

TABLE A.2: Outcomes of radio coverage results using only not-yet-treated cells with never-treated cells omitted.

	Dependent variable:		
	Malaria prevalence	Nighttime luminosity	Vegetation index
	(1)	(2)	(3)
Control groups: Never and not yet received coverage			
Coverage	−0.010	0.013	0.013
	(0.006)	(0.008)	(0.005)
Control groups: Never received coverage only			
Coverage	−0.007	0.014	0.013
	(0.005)	(0.008)	(0.005)
Sample size	9936	13872	4080
Mean dependent variable	0.592	0.023	0.650

Estimates of the overall average treatment effect on the treated following Callaway and Sant'Anna (2021) using the outcome regression approach with cell fixed effects, year fixed effects, and weights based on group size. Robust standard errors clustered at the cell level are in parentheses. Nighttime luminosity is transformed using the inverse hyperbolic sign transformation.

Table A.3: Outcomes of radio coverage results using a coverage-land ratio below 0.1 and coverage in the upper decile of distance away from the station.

	Dependent variable:		
	Malaria prevalence	Nighttime luminosity	Vegetation index
	(1)	(2)	(3)
Control groups: Never and not yet received coverage			
Coverage	−0.014	0.010	0.009
	(0.002)	(0.007)	(0.002)
Control groups: Never received coverage only			
Coverage	−0.012	0.008	0.008
	(0.002)	(0.007)	(0.002)
Sample size	40656	51440	17024
Mean dependent variable	0.603	0.038	0.672

Estimates of the overall average treatment effect on the treated following Callaway and Sant'Anna (2021) using the outcome regression approach with cell fixed effects, year fixed effects, and weights based on group size. Robust standard errors clustered at the cell level are in parentheses. Nighttime luminosity is transformed using the inverse hyperbolic sign transformation.

Table A.4: Outcomes of radio coverage results using cells with a maximum elevation range within a 9km×9km grid around the cell of less than 100m.

	Dependent variable:		
	Malaria prevalence	Nighttime luminosity	Vegetation index
	(1)	(2)	(3)
Control groups: Never and not yet received coverage			
Coverage (placebo)	−0.001	−0.003	0.000
	(0.002)	(0.005)	(0.002)
Control groups: Never received coverage only			
Coverage (placebo)	−0.004	−0.002	0.000
	(0.002)	(0.005)	(0.002)
Sample size	32608	42704	17040
Mean dependent variable	0.633	0.043	0.660

Estimates of the overall average treatment effect on the treated following Callaway and Sant'Anna (2021) using the outcome regression approach with cell fixed effects, year fixed effects, and weights based on group size. Robust standard errors clustered at the cell level are in parentheses. Nighttime luminosity is transformed using the inverse hyperbolic sign transformation.

Table A.5: Outcomes of radio coverage results with placebo first year of coverage.

	Dependent variable:		
	Malaria prevalence	Nighttime luminosity	Vegetation index
	(1)	(2)	(3)
Control groups: Never and not yet received coverage			
Coverage	−0.064	0.062	0.001
	(0.005)	(0.021)	(0.009)
Control groups: Never received coverage only			
Coverage	−0.066	0.064	0.000
	(0.005)	(0.020)	(0.008)
Sample size	21264	28208	10016
Mean dependent variable	0.614	0.038	0.665

Estimates of the overall average treatment effect on the treated following Callaway and Sant'Anna (2021) using the outcome regression approach with cell fixed effects, year fixed effects, and weights based on group size. Robust standard errors clustered at the cell level are in parentheses. Nighttime luminosity is transformed using the inverse hyperbolic sign transformation. Cells with any coverage from public stations are excluded from the sample.

Table A.6: Outcomes of radio coverage results using coverage from commercial stations only.

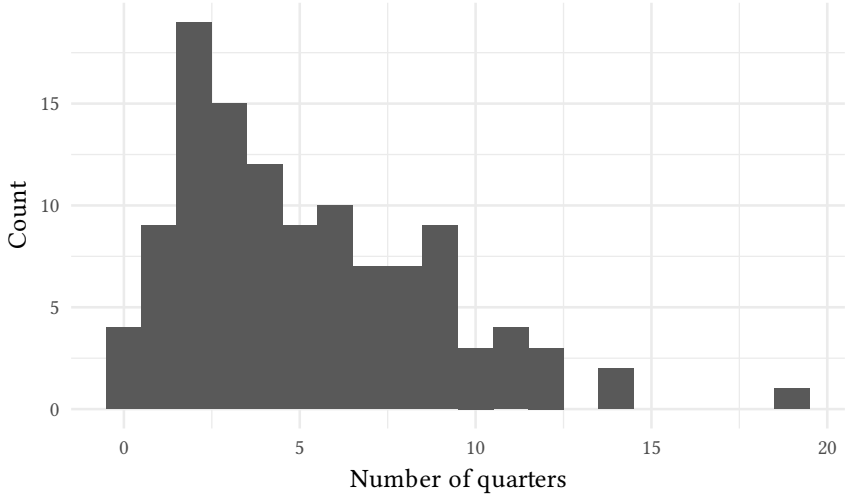
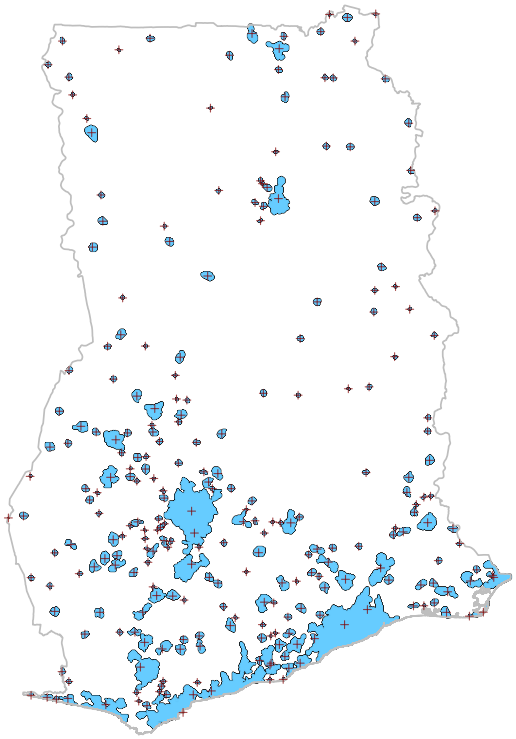
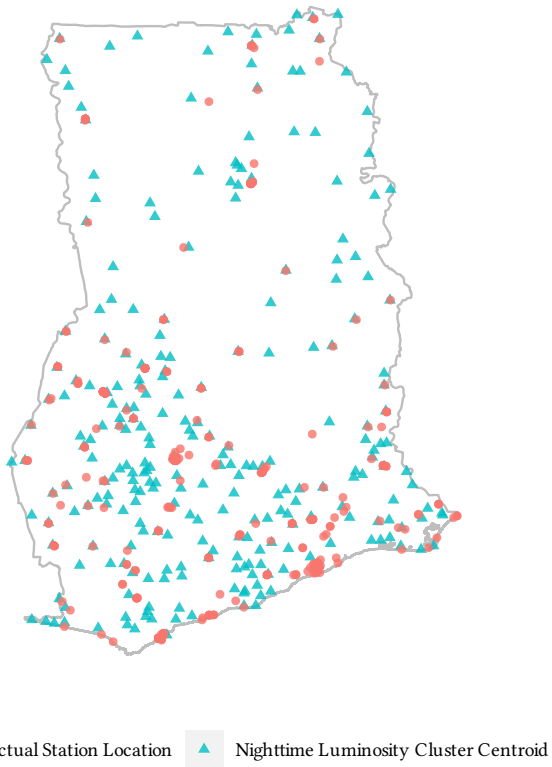


FIGURE A.7: Number of quarters between date of authorization and date first active.



Additional potential entrants are placed at the centroid of all nighttime luminosity clusters, representing populated areas in the country. The blue ares are the nighttime luminosity clusters and the red points are their centroids.

FIGURE A.8: Using nighttime luminosity to construct additional potential entrants.



Locations of all potential entrants, both firms from data and centroids of nighttime luminosity clusters.

FIGURE A.9: Potential entrant locations.

Variable	Estimate	Standard Error
Intercept	−5.083	(0.155)
Counterfactual profits	0.024	(0.007)
Time trend	0.124	(0.010)

TABLE A.7: Fitted policy function for rivals' actions during estimation at the estimated structural parameters.

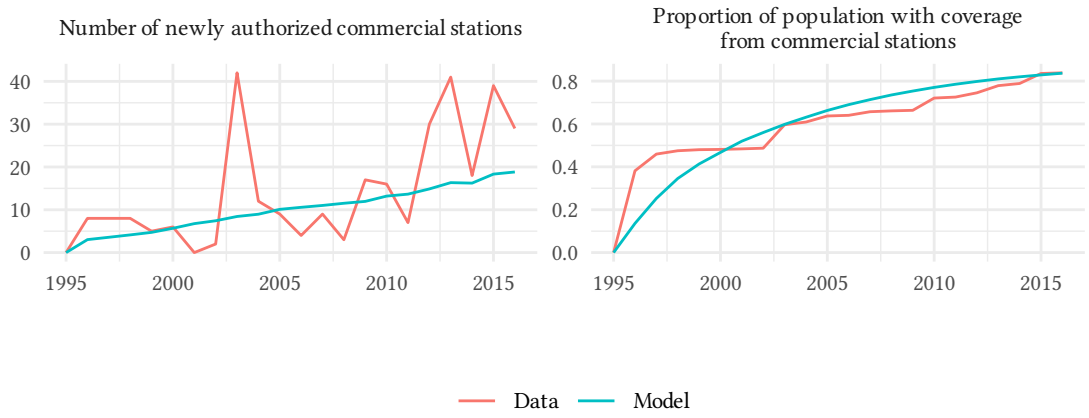


FIGURE A.10: Model fit in total number of commercial stations and percentage of population with coverage.

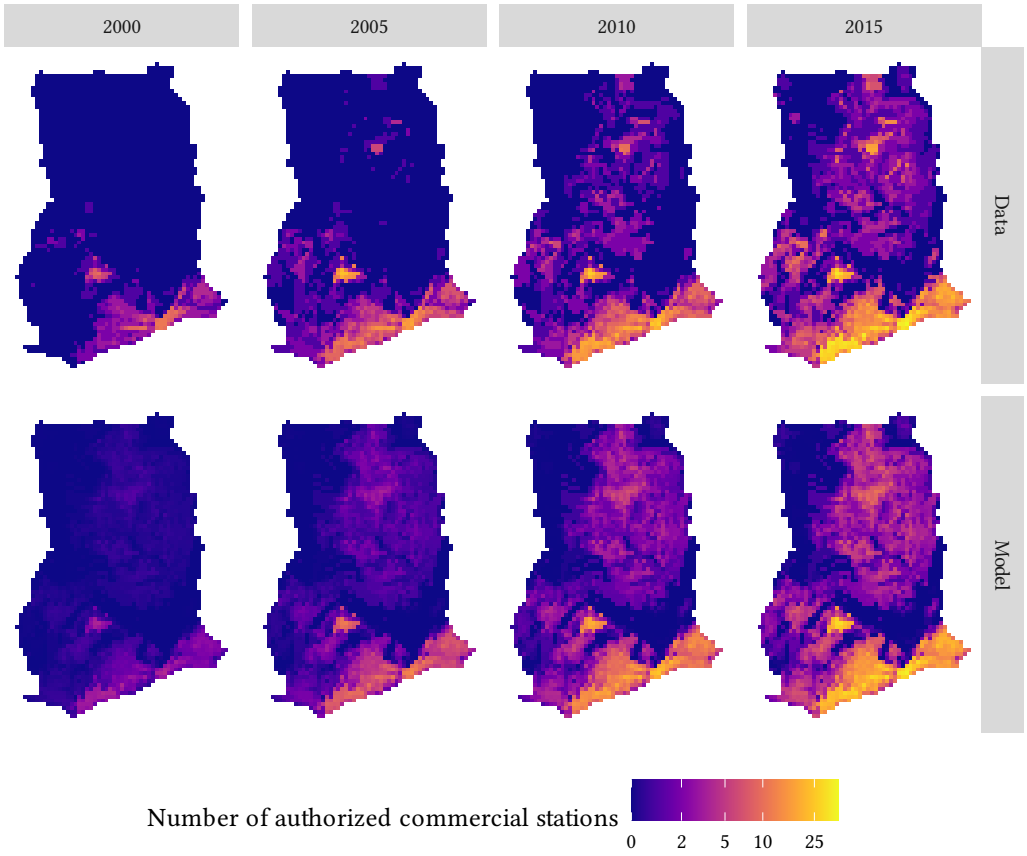


FIGURE A.11: Model fit across the country.

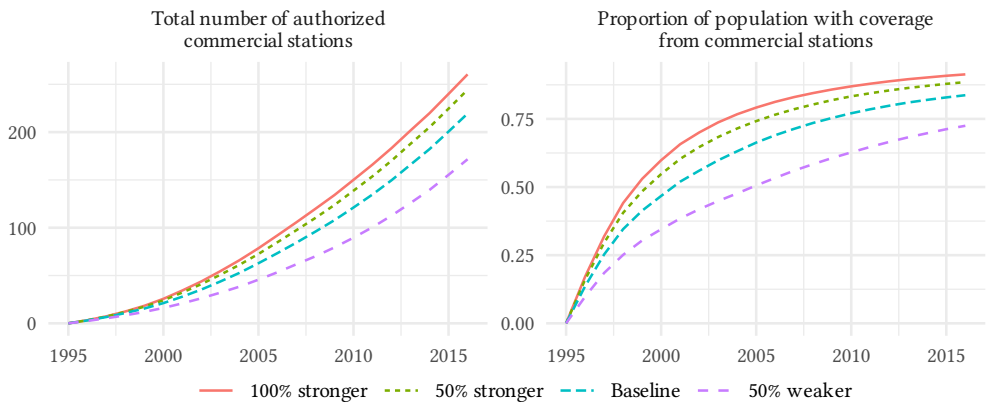
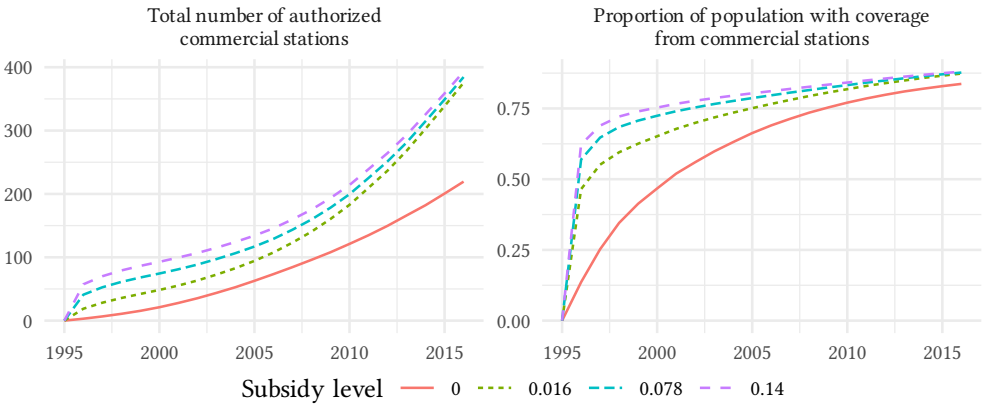
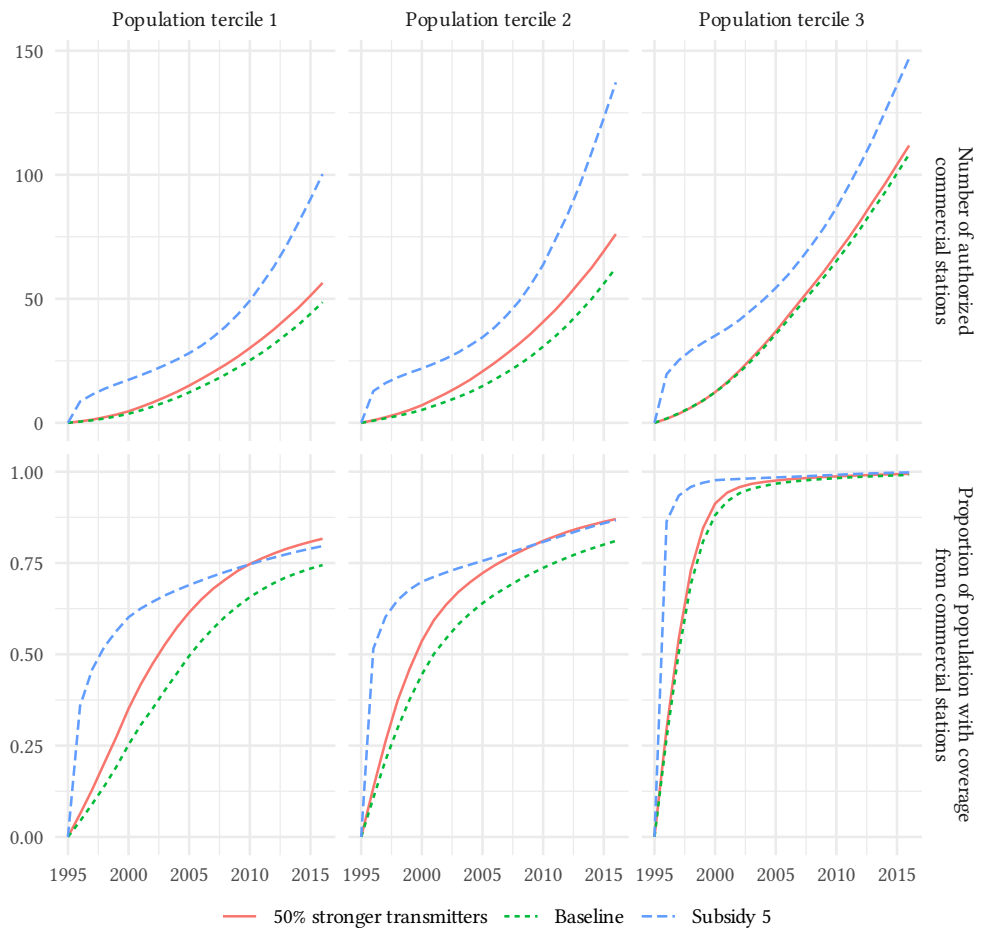


FIGURE A.12: Counterfactual signal strength policies.



The subsidy level denotes the proportion of the average annual profits a station would receive once-off for every additional 10,000 people newly served by radio after entry.

FIGURE A.13: Targeted entry subsidies.



Each population tercile has an equal total population by differ by population density. Population tercile 1 is the most rural, and population tercile 3 is the most urban.

FIGURE A.14: Counterfactual simulation results split by population terciles.

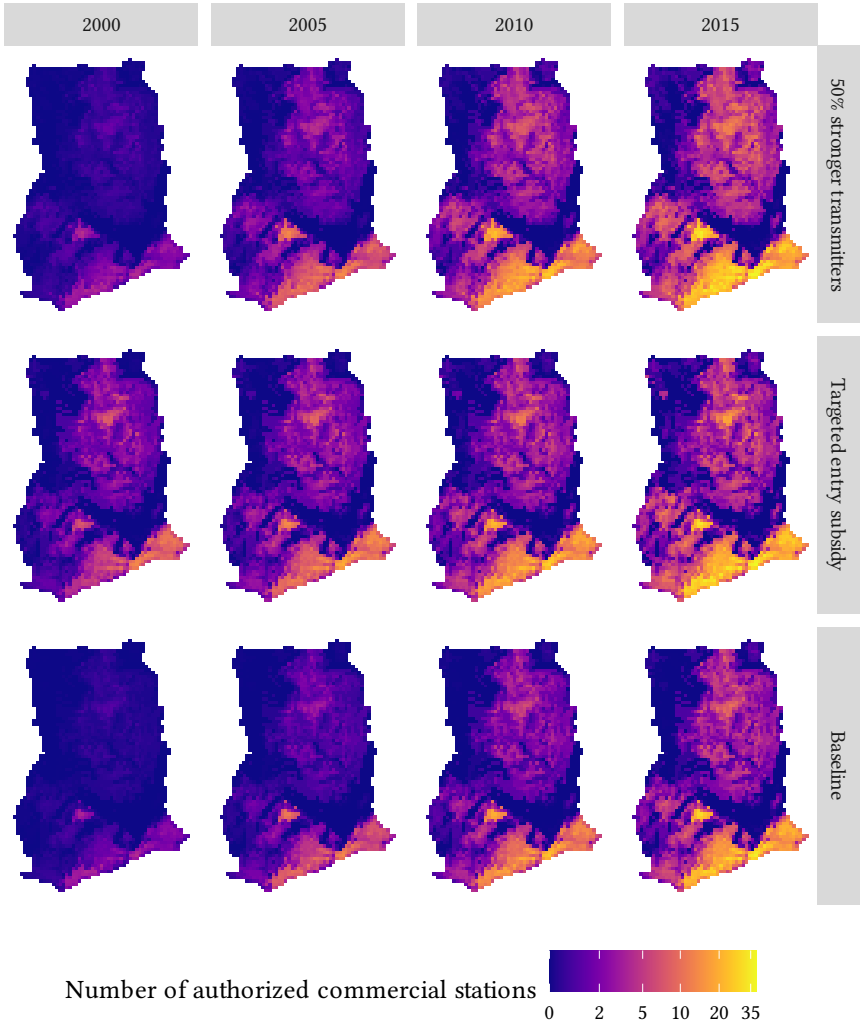
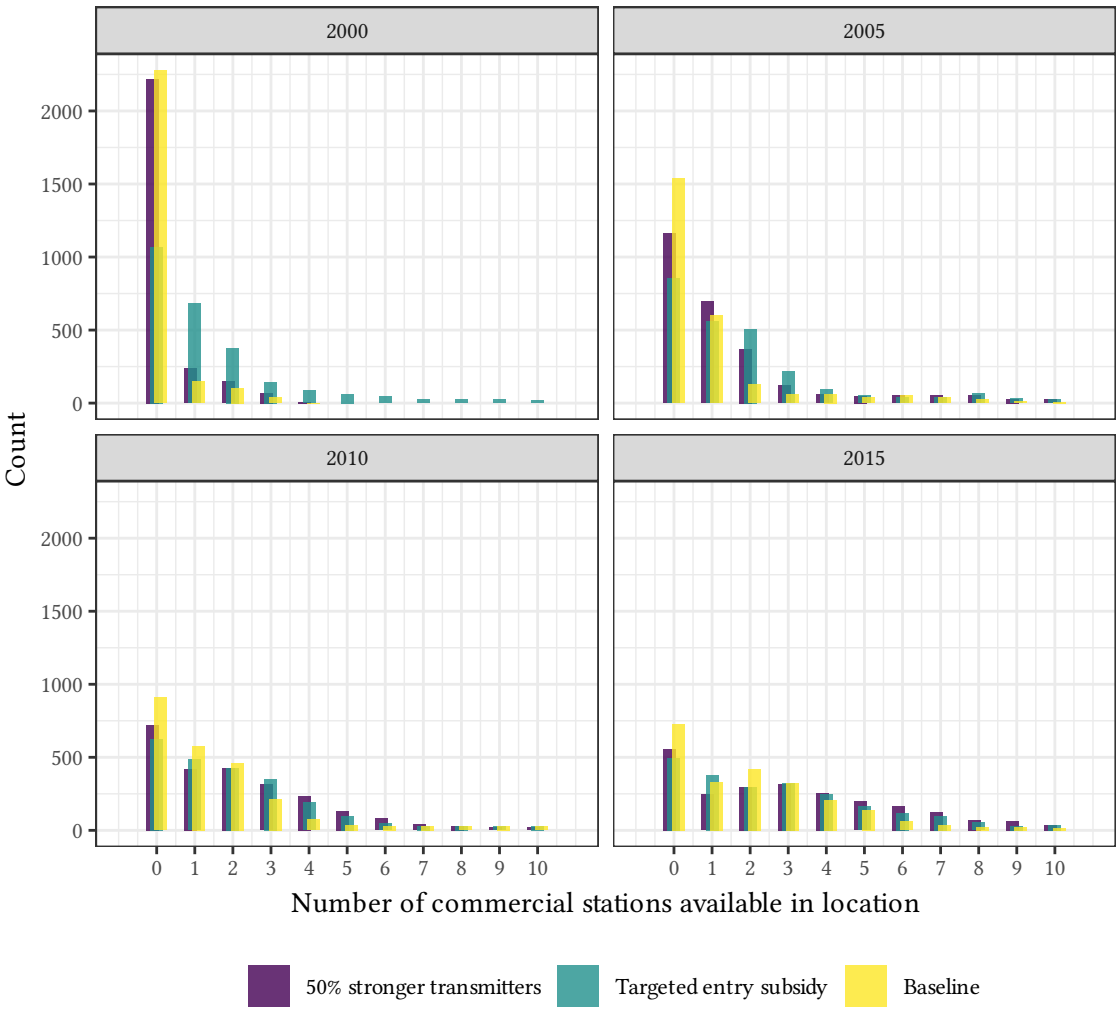


FIGURE A.15: Counterfactual simulation results by geography.



The bars in the histograms represent the number of cells (locations) where each number of commercial stations are available in a year, for each counterfactual simulation. Values above 10 are omitted to maintain legibility.

Figure A.16: Histograms of the counterfactual simulation results.

A.2 License Type Definitions

The National Communications Authority describes each of the different license types as follows. *Commercial* stations are those that are privately owned, controlled and operated for profit by independent commercial groups or individuals. *Public* stations are stations that are owned and operated by the Ghana Broadcasting Corporation (GBC) or any other stations established by the Government of Ghana whereas *Public Foreign* stations are stations established by foreign governments, such as the BBC. *Community* stations are non-profit and provide service for a specific marginalized community. Ownership and management of community stations are representative of the community. Finally, *Campus* stations are stations operating within the ambit of educational institutions.

A.3 Station Language and Content

We searched for information on the stations in our data set through a number of sources, such as their websites, Facebook pages, and live streams. We were able to ascertain the languages for 197 of our 436 stations. Table A.8 shows the frequency of different languages among these stations. 107 of these listed English as the only language and a further 66 stations were multilingual with English as one of the languages. Akan, Twi and Ewe are the most common languages other than English. There were 13 stations with other languages that no other station broadcasted in. The total sum of stations in Table A.8 adds up to more than 197 because of multilingual stations.

Language	No. of stations	Language	No. of stations
English only	107	Wale	4
Partially English	66	Sissali	4
Akan	22	Gonja	4
Twi	21	Briffor	3
Ewe	20	Kokomba	2
Dagare	9	Hausa	2
Fante	8	Efutu	2
Dagbani	5	Other languages	13

Table A.8: Station languages for a subset of stations in the data.

We were also able to obtain a short description of the station's genre for 153 stations. 103 of these stations describe themselves as partly music or entertainment stations

whereas 91 describe themselves as news or talk stations. 64 of the stations describe themselves as doing both news and music. Therefore music and talk stations make up 85% of all stations we have information on. There are 12 religious stations out of these 153 stations, which is the most common alternative type of station. A smaller number of stations detailed the music genres the played, with most broadcasting several genres. The most common genres are pop, local and African music.

A.4 Distribution of Languages

I use data from Glottolog to obtain the distribution of local languages. This contains the geographic coordinates of 64 distinct local languages in Ghana. These coordinates "represent the geographical center-point of the area where the speakers live." These are shown in the map in Figure A.17. The precise language border for these languages, however, are not provided by Glottolog. For each location in the country, I assume that the local language is the language of the closest centroid provided by Glottolog. The means that the language is a Voronoi tessellation within the country's borders. This is represented by the gray polygons in Figure A.17.

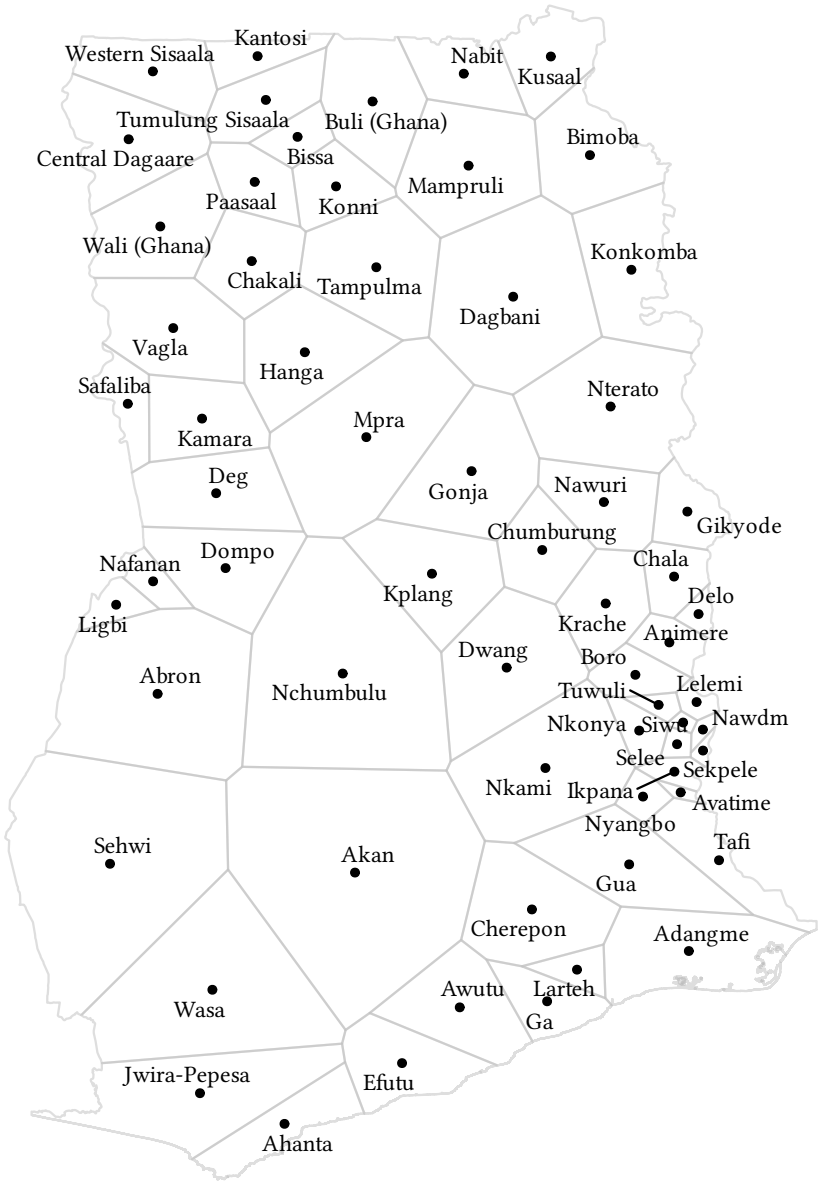


Figure A.17: Distribution of Local Languages in Ghana.

A.5 Example Coverage Streak

Figure A.18 shows an example of a part of Ghana that was included in the reduced-form regressions. The top-left figure shows the cells with radio coverage in that year. Each square in the figure is a cell and each cell is approximately 900m×900m in area. The top-center figure shows the cells that form part of a coverage streak. These are cells where the coverage-land ratio is below 0.2 and the cells are in the outer quintile of distance away from a station's transmitter. Locations that are included in the regression are shown in the top-right figure. This includes the coverage streak itself (in purple), as well as tiles that do not have coverage but are are within 1km of the edge of the coverage streak (in yellow).

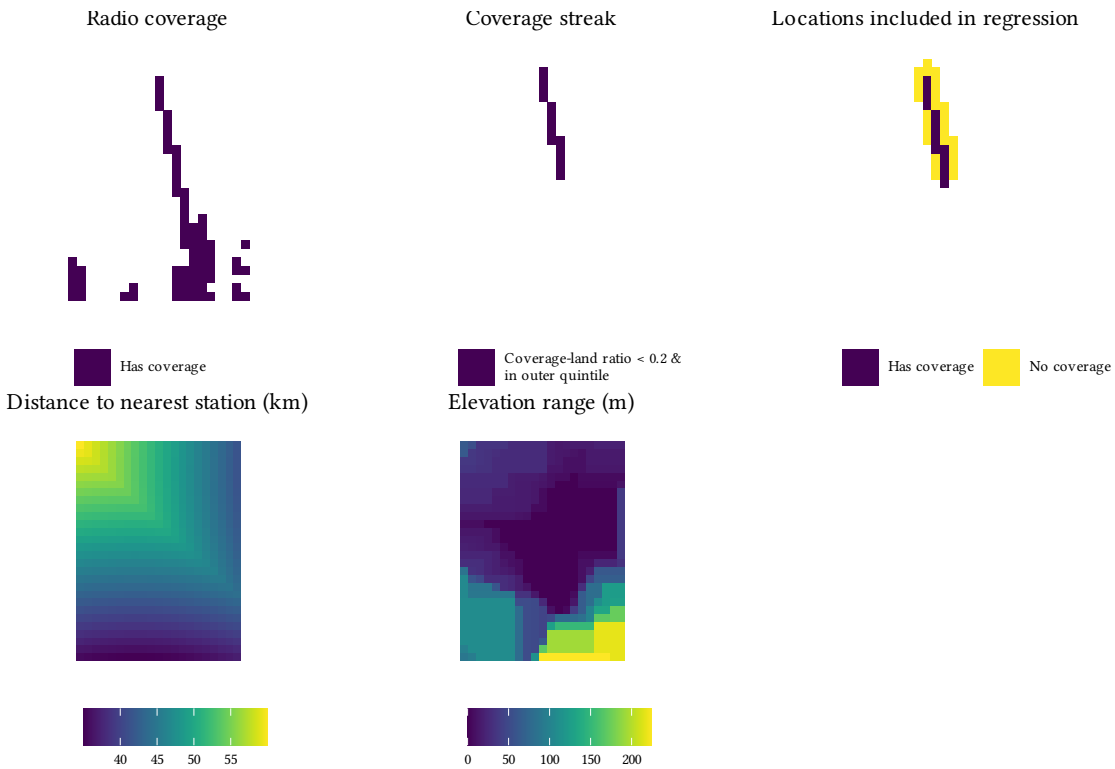


Figure A.18: Example coverage streak

If the distance to the nearest station is less than 10km or the maximum elevation range within a 9km×9km grid around the cell is larger than 50m, then the cell is omitted from the regression. This is because locations that are close to the source of coverage are more likely to be within the intended coverage area of the station. We also omit locations

that are very close to abrupt elevation changes as this may affect outcomes directly. The distance to the nearest station and the elevation range within a 9km×9km grid in the same area are shown in the bottom-left and bottom-center panels. From this we can see that the example coverage streak satisfies both conditions.

A.6 Demographic Health Surveys

Mosquito bed nets are a simple but highly effective method to prevent malaria. To investigate if individuals in locations receiving radio coverage are also more likely to use mosquito bed nets, I make use of the Demographic Health Survey (DHS) data. The DHS program has conducted more than 300 surveys of population, health and nutrition in over 90 countries. In Ghana, they have conducted six repeated cross sections from 1988 to 2014. However, the questions vary year-by-year and information on mosquito bed net usage is contained only in the surveys since 2003. The DHS provides approximate coordinates of each survey cluster, which allows matching with the coverage data. The 1,250 cluster locations are shown in Figure A.19. We can see that the survey locations are spread throughout the entire country. To preserve the anonymity of the survey respondents, the DHS purposefully adds an error of up to 2km for urban clusters and 5km for rural clusters. Thus matching the coverage data with the DHS data using the coordinates that they provide introduces large measurement error. To reduce this measurement error I take the average amount of coverage available within a radius around the cluster locations, rather than taking the number of available radio stations at the coordinates given in the data. The radius I use for each survey cluster corresponds to the maximum possible error introduced by the DHS (2km for urban clusters and 5km for rural clusters).

Given the noise purposefully introduced to survey locations, I cannot implement the coverage streak identification strategy because we do not know if a survey cluster is inside or outside of a coverage streak. Even without this anonymization, there are very few survey clusters that are near coverage streaks. Therefore I use both a fixed effects and an instrumental variables estimation strategy to estimate the impact of coverage on bed net usage. One possible instrument would be to use local hilliness as this is negatively correlated with radio coverage. However, using such an instrument is problematic if hilliness affects outcomes directly. For example, areas that are hilly could be more disconnected from society and as a result obtain less information from other villages. Instead of using local hilliness as an instrument, I use distant hilliness. Elevation changes far away from

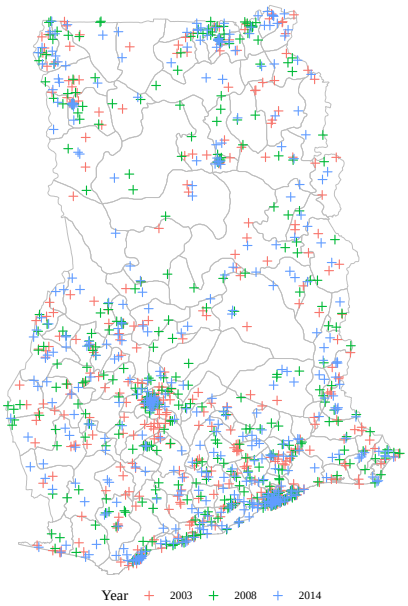


Figure A.19: Demographic Health Survey cluster locations by year.

the village can also affect the amount of coverage the village receives. This is because it does not matter if the obstruction between the radio tower and the village is closer to the village or further away and closer to the radio tower: both will affect the signal reaching a village in the same way. However, the distant elevation from the village is less likely to affect the village directly compared to local elevation. Based on this idea, I construct the instrument as follows. I draw two circles around each survey cluster in the DHS data with radii 10km and 15km respectively. These two circles form a ring. I then take the standard deviation of elevation values within the ring as a measure of distant hilliness. Figure A.20 shows an example ring around a survey cluster on top of an elevation heatmap.

Table A.9 shows the regression results. Because coverage only varies at the cluster-year level, I first aggregate the household data to that level. The first column shows the results from a linear model regressing the proportion of households where children use a mosquito bed net on radio station coverage. I control for year fixed effects, region fixed effects and a host of demographic controls. I also control for local malaria presence using the average incidence within a 5km radius of the survey cluster. Coverage increases the probability of mosquito bed net usage on average by 4.2 percentage points over a baseline average usage of 29.7%. The second column shows the results from the first stage of the instrumental variables regression. The standard deviation of elevation in the ring

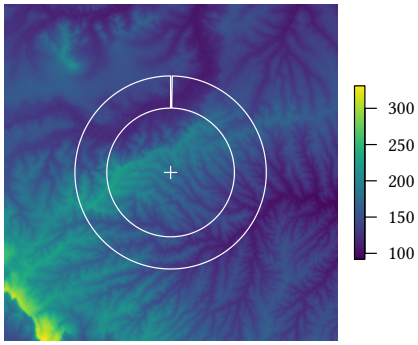


FIGURE A.20: Elevation Instrument

Example ring around a survey cluster overlaid on an elevation heatmap (measured in meters). The instrument for coverage is the standard deviation of elevation within the ring, measuring distant hilliness from a survey cluster.

	Children sleep under bed net OLS (1)	Coverage from commercial station OLS (2)	Children sleep under bed net IV (3)
Coverage from commercial station	0.042		0.169
	(0.020)		(0.072)
Std. dev. of elevation in ring around cluster		−0.003	
		(0.000)	
Demographic controls	Yes	Yes	Yes
Region Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
N	1142	1142	1142
Adjusted R^2	0.389	0.504	0.367
F-statistic	25.250	39.607	

Robust standard errors in parentheses. Data is collapsed at the survey cluster-year level. Demographic controls include gender, age, education attainment, religion, ethnicity, literacy, newspaper readership, television viewership, urban/rural. Local malaria presence is also controlled for using Malaria Atlas Project data.

TABLE A.9: DHS Regression Results.

around the survey cluster has a significant negative effect on coverage. In the instrumental variables regression, coverage on average increases the probability of mosquito bed net usage by 16.9 percentage points. This result is consistent with the results found for malaria prevalence and provides a supporting mechanism. Areas receiving radio coverage are informed of ways to reduce the risk of malaria, and more households use mosquito bed nets. This in turn leads to a lower prevalence rate in those areas.

A.7 Increased Transmitter Strengths: Population Served and Competition Intensity

In this section I discuss how the increased transmitter strengths affect stations' profits through the increased population served and differing competition intensity. I also discuss the impact of a single station unilaterally increasing its signal strength on profits.

With a 50% increase in transmitter strengths, the population living within a station's

coverage area is 37.6% larger for the median firm. However, in general the policy also impacts the amount of competition faced by stations as, holding the entry configuration fixed, the majority of listeners have a wider range of stations to choose from. To compute how increased signal strengths affect competition, I compute how a station's average listener share across locations (conditional on listening) changes. The average listener share conditional on listening across locations is given by:

$$\frac{1}{\mu_t} \frac{\sum_{\ell=1}^{L} \mathfrak{s}_{f\ell t}}{\sum_{\ell=1}^{L} \gamma_{f\ell}}$$

At the entry configuration in the data, this falls by 7.2% for the median station, indicating an increase in competition from increased transmitter strengths. 71.4% of station-years experience an increase in competition. Most stations earn higher profits with stronger transmitters because the population effect is larger than the competition effect. But at the entry configuration in the data, 9.3% of station-years would have smaller profits under stronger transmitters, indicating that the competition effect dominates the population effect for these stations.

I now discuss the impact of a unilateral increase in a station's transmitter strength. For each station, one-by-one, I increase their transmitter strength holding the coverage areas of other stations fixed and compute the profits they would receive if they were active. I do this at the entry configuration observed in the data. Compared to the baseline case, the median firm has a 33% increase in profits. All stations would earn at least as much profits compared to the baseline case. This is because the listener share is the same in all locations where they had coverage in the baseline case (as all rivals have the same coverage area), but they earn additional profits from the new locations that their coverage reaches. Compared to the situation where all stations have stronger transmitters, the unilateral increase leads to a 4% increase in profits for the median firm. This is because although the population within the station's coverage area is the same in both situations, the station competes with fewer stations for listeners as rivals have smaller coverage areas in the unilateral increase scenario.