THE EFFECT OF BROADBAND DIFFUSION ON U.S. INTERNAL MIGRATION

USING LIGHTNING STRIKES AS AN INSTRUMENT

by

Daniel Foster Malone

A thesis submitted in partial fulfillment
of the requirements for the degree

of

Master of Science

in

Applied Economics

MONTANA STATE UNIVERSITY
Bozeman, Montana

May 2025

TABLE OF CONTENTS

TABLE OF CONTENTS CONTINUED

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

ABSTRACT

This thesis seeks to explore the relationship between internet connectivity and internal migration rates in the United States. A two-stage least squares design with a long-difference technique took advantage of the variation of a cross-section of U.S. counties, between the years 2011 and 2019, to determine a causal effect. Average lightning strike density between the years 2000 and 2010 was used as an instrument to estimate internet connectivity in a county, which was then used to estimate the effect of internet connectivity on net migration, inflow migration and outflow migration. Regressions were used to establish the causal effect of internet connectivity on a change in wages. The results find that a 20-percentage point increase of broadband connectivity results in an almost 5% average annual drop in inflow and outflow migration rates over the time-period overall, and an annual average 0.7% influx of migrants into rural counties. Consistent with the literature, it finds that a 20-percentage point increase in broadband connectivity results in an increase in an overall annual average 1.8% growth in wages over the time-period.

CHAPTER ONE

INTRODUCTION

The internet has become the most important invention of our time. As of 2024, 65.6% of the world's population has internet access, and 95% of all information available has been digitized and made accessible (Giannelis, 2024). While the widespread availability of the internet allows the ability to connect people across great distances in real time, the population has clustered around tech hubs that have emerged in recent years which disproportionately attract investment and skilled workers. This has given rise to concerns that this clustering effect will result in a "digital divide", where less developed parts of the country will fall behind the more developed parts, which will increase wealth inequality and political polarization (Moretti, 2012).

The United States government has recognized the internet's significance and has invested extensive number of resources to disseminating broadband across the country through a series of subsidies. In 2009, Congress introduced the Broadband Technology Opportunities Program as part of the American Recovery and Reinvestment Act, providing $4 billion in grants intended to bridge the technology divide by deploying broadband infrastructure and encouraging the adoption of broadband service (NTIA, 2020). In 2021, a package called the Broadband Equity, Access and Deployment Program (BEAD) allocated $42 billion dollars which was intended to expand broadband to rural communities, though the rollout has faced delays (Hendel, 2024). Significant policy interest in expanding the internet underscores the need to determine the consequences of internet expansion.

The purpose of this thesis is to estimate the effect of varying levels of pre-existing broadband availability on internal migration. This estimation is complicated by the problem of omitted variable bias; both the rollout of broadband and internal migration is at least partially determined by the geography of an area. To alleviate omitted variable bias, a two stage least squares estimate will be used, using lightning strikes as the instrument. This strategy is used to estimate the level of broadband connectivity in the year 2011, which is then used to estimate the difference in annual net, outflow and inflow migration rates between the years 2011 and 2019.

This thesis aims to look at the effect of the internet on inter-state migration. It examines whether the mechanism for migration is at least partially dependent on the availability of high-speed internet at the migrant's destination. It examines whether a pre-existing level of broadband access in a county act as a key factor where prospective migrants decide to move to. It will also seek to determine whether the push effect of higher internet connectivity, associated with lower information costs when considering alternative locations to reside, is stronger than the pull effect, which is due to higher levels of productivity and more efficient worker-firm matching within one's own community. A separate set of regressions were conducted to determine the effect of internet penetration on wage growth.

This thesis finds that the effect of internet connectivity on overall net migration is insignificant, but there is a sharp decline in outflow and inflow migration rates, which suggests an overall decrease in mobility for counties with higher levels of internet access. It also finds, however, that rural counties with a higher baseline level of internet access experience a significant inflow of individuals moving into those counties, relative to counties with a relatively lower level of access. It also finds that, overall, wages in counties with higher levels of internet penetration grow significantly faster than in counties with relatively lower levels of penetration.

Migration Dynamics

The economic decision to migrate from a familiar area to an unfamiliar one is multifaceted and complex. Most models of migration are based on individual utility maximization where individuals weigh the relative benefits and costs for moving. These models are understood as centering around the "gravity" of origins and destinations, where the decision to move is directly related to the attractiveness of the destination, and inversely related to the distance between the origin and destination. Specifically, individuals are drawn to higher relative wages but are dissuaded by the distance that they would have to travel to reach their destination. This model, proposed by Larry Sjaastad in his 1962 paper, "The Relationship between Migration and Income in the United States", remains influential in the current literature (Greenwood, 1975).

The Effect of Distance on Migration.

One of the most important factors for individuals choosing to migrate is distance. All else being equal, the probability of an individual moving to an area which is far away compared to an area which is closer is significantly lower (Greenwood, 1975). According to Michael Greenwood (1975), three mechanisms explain how distance acts as a deterrent:

> [First,] In many instances the benefits associated with migration are not particularly high, and hence direct transportation expenses need not be great to discourage migration. [Second,] The psychic costs involved in migration are substantial and are closely related to distance. [Third,] Information declines perceptibly with distance, and hence uncertainty increases with distance [pg. 398].

The third mechanism, informational uncertainty, is especially relevant in the context of the effect of distance. Before the widespread adoption of the internet, information about destinations disseminated slowly over time, as was seen during the Great Migration from the

American South to the North between 1915 to 1960. During this time, young men with low migration costs migrated to the North first, then relayed information of their new homes back to their families in the South, which lowered the information barriers for subsequent migrants (Carrington, 1996). Today, the internet dramatically lowers information costs for those who have access to it, which allows individuals to more easily review employment, housing and other conditions in distant locations (Goldfarb, 2019). In this way, individuals with access to the internet would be able to "escape the gravity well" of their origin easier than those without, potentially facilitating out-migration.

The Effects of Relative Wage on Migration.

The principle of utility maximization is central to most economic models of migration. Individuals estimating the benefits of migration are assumed to compare the expected wages of the origin and the destination is used. In the existing body of literature, these potential incomes - defined as product of the probability of the worker finding a job in each location multiplied by the expected earnings of each location – are evaluated in both the origin and the destination (Todaro, 1969; Greenwood, 1975; Lee, 1999). The internet plays a role in this calculation by increasing the productivity of workers, which would theoretically translate to a corresponding increase in wages (Fabricant, 1962), as well as increasing the likelihood of finding employment in counties with higher levels of internet connectivity (Atasoy, 2013). However, as will be discussed later in this chapter, holding other things constant the increase in wages is heterogeneous in the population, which is dependent on the educational background and industry an individual is employed in. Therefore, the "pull" effect of better career opportunities may be weak or even inverted, so the effect of the internet on wages will be investigated in this thesis as well, to determine the effect of the relative level of and change in wages on the theoretical model, as it pertains to expected earnings.

The Internet

The Effect of the Internet on Productivity and Costs

The expansion of high-speed internet has been shown to dramatically increase the productivity of firms which employ it. A study of OECD countries between 1996 to 2007 found that a 10% increase in broadband penetration was associated with increased per capita GDP growth by up to 1.5% (Czernic, 2011). In New Zealand, firms which adopted broadband increased their productivity by up to 10% (Grimes, 2012). Beyond productivity, the internet also stimulates the development of new firms as well: the availability of ultra-fast broadband, (100+ megabytes per second download speeds), increases the rate of new business formation for up to 5 years after adoption (Biedny, 2024). The availability of high-speed internet also enhances the resilience of firms against economic downturns. Counties which were enrolled in the USDA's Broadband Initiative Program had a less severe decline in average employment after the program's introduction and rollout in 2009 and 2010 (Rupasingha, 2024).

Productivity gains of firms which employ high-speed internet stem primarily through the lowering of costs. Goldfarb (2019) identifies five primary cost reductions for firms: search, replication, transportation, tracking and verification costs. Additionally, the internet lowers costs for consumers by lowering transaction costs (Bajari, 2004).

Lowering Search Costs. The internet lowers search costs by being more efficient at finding information than traditional information sources. For example, a study which investigated airline and booking prices found that after the commercialization of the internet, consumers started purchasing airline tickets from the airline directly, eliminating the middleman of travel agents (Orlav, 2011). Firms also directly benefit from lowered search costs as they can source inputs and goods from a wider range of potential business partners, leading to more competitive pricing and reduced price dispersion. In both cases, the expanded access to information allows consumers and firms to access a broader range of goods and services (Goldfarb, 2019).

Lowering Replication and Transportation Costs Perhaps the most profound development comes in the form of the near elimination of replication and transportation costs. Before the internet, information such as music or software could only be distributed through physical media such as CD-ROM or Blu-Ray disks, which required the production of blank disks, the transport of the disks to firms to burn information onto the disks and then transport again to distribution centers and finally to a retailer, where the disk required physical display space. The internet has reduced these needs to a novelty; it allows information to be infinitely replicated and transported at virtually no marginal cost. The internet has therefore changed intellectual property from a rival good to a non-rival good, excluding attempts by copyright holders to restrict access to the intellectual property in question. This allows novel pricing models, such as club models where instead of purchasing specific instances of intellectual property, a consumer now has the option to pay a subscription fee for unlimited access to a catalog of properties, such as Netflix or Microsoft Game Pass (Goldfarb, 2019).

Shifts in Workforce Requirements Increased productivity for firms which employ the internet requires a shift in labor composition and education. A study investigating firms that invest in new equipment that provide real-time status updates of the manufacturing process found that firms require new staff which are trained to run and maintain the equipment, which is reflected in an increase in the proportion of white-collar workers employed (Bartel, 2007). Naturally, this shift in educational requirements has led to increases in wages and employment levels of college graduates in areas with access to broadband (Atasoy, 2013). However, sectors which are tangential to the internet experience fewer benefits from increased access to it; farmers in the Great Plains region of the United States who use the internet commonly report that they do not experience any returns on this ability, and only 30% of farmers report a positive financial return (Smith et al, 2004).

Labor market gains from broadband access are not limited to only college graduates. Access to broadband is linked to higher employment levels and wages for high school graduates, which is due to an increased need for services for the newly expanded cohort of white-collar workers (Moretti, 2012). The internet also makes novel forms of work possible, notably in the form of the "gig economy", where workers are employed for single projects or tasks, such as Uber. The gig economy is an example of a new industry which is only possible due to the reduction in search and verification costs (Anani, 2018). Furthermore, the facilitation of remote work improves labor market participation for demographics which were historically constrained by geography or caregiving responsibilities, particularly women with children (Dettling, 2017).

The Effect of the Internet on Wages

The internet has a heterogenous effect on wages, dependent on the industry that the individual using the internet is employed in. Industries which are expected to experience a significant increase in productivity, such as finance or information, enjoy higher wages after broadband is introduced into the area. Industries which would not experience an increase in productivity, or which render their workforce redundant through the introduction of the internet, tend to experience lower wages after the introduction of broadband (Atasoy, 2013). Other studies corroborate this claim that wage growth is enjoyed mainly by skilled workers and point out that the differential effect between skilled and unskilled labor would exasperate inequality, and result in the "digital divide" discussed earlier (DiMaggio and Bonikowski, 2008; Moretti, 2012). It is therefore important to determine how wages change over time to more completely understand the mechanisms underlying migration rates in a county.

Previous Uses of Instrumental Variables to Estimate Internet Penetration

The technique of instrumental variables to estimate the internet penetration in a region has been used several times in prior literature; this thesis will focus on the methods used by Dettling (2017) and especially Anderson (2012). Dettling (2017) notes that internet users are not randomly assigned, and that internet adoption is endogenous to labor market outcomes, as wealthier and more educated individuals are more likely to purchase internet subscriptions. She uses an instrumental variable strategy which exploits variation in the proportion of the population for each state which resides in multi-family dwellings such as apartment buildings. The reasoning she used was that multi-dwelling units are more likely to receive broadband earlier, since it is more cost effective for an internet service provider to connect a building with many potential customers, than it is to connect individual housing with few customers.

The methods which Anderson (2012) uses forms the inspiration for the instrumental variable used in this thesis. Anderson uses the flash density for each U.S. state to as an instrument to estimate the percent of households with a personal computer at home in the year 2003, manufacturing firm's IT investments in 2007 and, crucially, the percent of households with internet access at home in 2003. These estimates were then used by Anderson to estimate economic growth in the state due to IT industries. This use of lightning strikes to estimate household internet connectivity, which is then used to determine the effect of internet penetration on economic indicators, forms the empirical inspiration of this thesis. The use of lightning strikes as an instrument for internet connectivity will be discussed in more detail in the Econometric Model chapter of this thesis.

CHAPTER TWO

ECONOMIC THEORY

The decision to migrate from origin i to destination j can be modeled as a utility maximizing choice, expressed by the following model described by equation (1), is derived partially from Greenwood's (1975), Todero's (1969) and Carrington et al. (1996). Specifically, defining the present value of a move in terms of expected earnings in both the destination and the origin are integral to both Greenwood's and Todero's models, the inclusion of the relative costs of both locations is part of Greenwood's model, and the inclusion of a one-time moving cost is from Carrington et al.'s model. The theoretical model is as follows:

(1) $$PV_{ij} = \sum_{t=1}^{n} \frac{Eijt}{(1+r)^t} - \sum_{t=1}^{n} \frac{C_{ijt}}{(1+r)^t} - M_{ij0}$$

where PV_{ij} is the present value of migrating from the origin i to the destination j, E_{ijt} is the difference in earnings which are expected at the origin and the destination, C_{ijt} is the difference in net costs associated with both localities, and r is a discount factor which embodies the reluctance of an individual towards migration. M_{ij0} is a measure of the cost of migrating from the origin to the destination. This is a one-time cost which is contextual to the distance between the origin and the destination, and the effort required of the individual to move. n is the number of years that the individual is expected to live. If PV_{ij} is greater than zero, the value of moving is greater than the value of staying at the origin, and the individual will move. If PV_{ij} is less than zero, the value of moving is less than the value of staying at the origin and the individual will not move. If PV_{ij} equals zero, the individual is ambivalent. The presence of the internet affects both E_{ijt} by increasing the probability of finding a job and potentially increasing wages, and C_{ijt} by lowering information costs, so determining the aggregate effect of the internet on migration is the objective of the regression analysis.

The equations for earnings for an individual in any location is provided by Kennan and Walker (2011), and is defined as equation (2):

(2) $$E_{it} = \mu_i + \upsilon_i + G(X, a, t) + \eta + \varepsilon_i(a)$$

where μ_i is the mean wage for location *i* or *j*, υ_i is a permanent location match effect with the individual's abilities and skills, *G(X, a, t)* represents a linear time effect and the effects of observed individual characteristics, η is an individual effect across locations, *a* is the age of the of the individual and $\varepsilon_i(a)$ is a transient effect for each location. For the purposes of this study, it can be reasoned that μ_i is directly proportional to increases in internet connectivity, as increases in connectivity is reflected in an increase in productivity (Grimes, 2012; Solomon, 1962) as well as improvements in matching worker skills to suitable employment (Atasoy, 2013). Todero's model of migration (1969) states that value maximizing individuals will pursue higher wages and therefore will migrate to areas with a higher GDP per capita, conditional on their probability of finding work in the new area. Therefore, migration is plausibly affected by the level of internet connectivity in a county.

The one-time cost in migrating from *i* to *j* is defined by Carrington et al. (1996), and is described by equation (3):

(3) $$M_{ij0} = F(R_{t-1}, h)$$

where *h* summarizes the personal characteristics which would affect the cost of migration (i.e. age, family status, assets, etc.) and *R* is denoted as being representative of the number of workers from the origin to the destination. The main purpose of the original study was to ascertain the effect of information on migration costs, which was determined at the time by migrants to the North relaying information of their new homes to their families in the South, which allowed their families to move to the North. In this context, *R* would also be defined by the availability of the internet at location *i*, which would allow the prospective migrant to understand the economic situation in location *j*. Therefore, *R* would be inversely proportionate to the level of connectivity at location *i* which would lower the cost of migration.

CHAPTER THREE

ECONOMETRIC MODEL

Overview of the Model

Explanation of Model

The experiment utilizes a 2SLS model to estimate the causal effect of internet connectivity on migration, using the quasi-random distribution of lightning strike density as an instrument for internet connectivity in each county in the year 2011. The regression is on a long-difference of U.S. counties over the time-period between 2011 and 2019, which allows a non-biased estimate of the effect of different pre-existing levels of internet connectivity on migration over time. The use of a long difference model, as opposed to a panel study, is necessary as lightning strikes notably vary little year to year, particularly in areas which do not experience lightning frequently (Anderson, 2012). As such, panel analysis would cause collinearity between lightning strikes and county fixed effects, dramatically increasing the estimated standard errors and make estimations of the true effect biased. Additionally, the rollout of IT technologies is noted to take a significantly long time, so a lengthy, lagged observation of lightning strike density is necessary to estimate a baseline level of internet connectivity (Anderson, 2012).

The primary regressor of interest is the number of households connected to the internet to allow for effective use of innovations such as telecommuting, which captures actual access to the internet, rather than indirect measures to determine connectivity such as nominal levels of internet speeds or number of service providers.

To gain complete understanding of the mechanisms in play in determining net migration for each county, separate regressions were used to ascertain the net, outflow and inflow migration rates. Additionally, a separate set of regressions were used with an interaction term between internet connectivity and a rural dummy variable to determine the separate effect that increased internet penetration has on rural communities, relative to urban ones. In these regressions, an interaction term between lightning strikes and the rural dummy variable are also included to account for the unique impacts on internet connectivity due to lightning strikes between rural and urban counties. The use of interaction terms in the context of a 2SLS regression model has been used by other authors in the past (Aghion et al., 2005), and the overall validity of the method has been asserted in a study investigating the general application of this method (Bun and Harrison, 2018). Another separate regression to determine the effect of internet connectivity on relative wages was also used to gain further understanding of the underlying mechanisms that influence migration. As the relationship between migration and internet connectivity is highly heteroskedastic, robust standard errors will be used and will be clustered around internet connectivity categories.

Omitted Variable Bias

A two-stage least squares (2SLS) model is necessary as the relationship between internet connectivity and migration are both correlated with the geography of the area.

The cost of expanding broadband to an area is determined largely by the topography of an area, especially in mountainous and/or forested regions, and it is also more expensive to expand broadband into rural communities which are far from urban centers (Ryssdal and Leeson, 2024). The federal government, when assigning grants through the Broadband, Equity, Access and

Deployment program (BEAD) will prioritize areas which are below a cost threshold, where the cost of the internet is deemed to be too high to justify a rollout. Areas which are too costly to build broadband infrastructure are instead encouraged to use alternative sources of internet connection, such as low orbit satellites (Reliable Broadband Service & Alternative Technologies, 2024), which are notably less reliable than broadband, as well as more expensive for consumers (Young and Thadani, 2022). The decision to migrate is dependent on a host of variables related to these areas, including job prospects (Tickameyer et al., 2017; Todero, 1969), prevalence of tourism industries (Hyytiä and Kola, 2013), and, as noted in detail in the Migration Dynamics sections of the introduction, the remoteness of these areas which contribute to the one-time moving cost (Greenwood, 1975; Carrington, 1996), all of which are influenced by geography. Thus, the effect of internet connectivity on internal migration is confounded by an unobserved omitted variable, so to ameliorate omitted variable bias, a 2SLS strategy will be employed.

Lightning Strikes as an Instrument

For an instrument to be considered valid, the following conditions must be met: First, the instrument must be correlated with the endogenous variable that it is estimating, namely the level of internet connectivity in the year 2011. Second, the instrument must not be correlated with the dependent variable, in this case migration rate, except through the endogenous variable. The use of lightning as an instrument was established in “Lightning, IT Diffusion, and Economic Growth Across U.S. States” (Anderson, 2012). This novel instrument was determined to be a strong predictor for GDP growth after the diffusion of computer and internet technology across the United States.

The authors found a robust negative relationship between lighting strikes and internet connectivity, the mechanism for which is described as follows: Lightning strikes cause unpredictable power surges, which damage electronics connected to the power network. This damage both requires frequent electronic replacement as well as the installation of lightning protection, leading to higher maintenance costs. The higher maintenance costs result in less investment in IT infrastructure in areas which experience frequent lightning storms, as returns on investment in IT infrastructure are lower in these areas than areas which experience fewer lightning storms, which in turn slows diffusion of internet technologies and results in a lower level of connectivity in the affected areas. While Anderson's paper focuses on estimating household internet connectivity at a state level due to data restrictions, the model is applicable on a county level. The relationship between lightning strikes and internet connectivity is supported by circumstantial evidence: the U.S. states with the highest level of IT industries are California and Washington (Moretti, 2012), which experience very few lightning storms each year, whereas the American South, where lightning storms are frequent, have notably slower internet speeds than the rest of the country. A more detailed explanation of this mechanism, and the science of lightning strike mitigation in general, is available in Appendix A.

For the instrument to be valid, the plausible exogeneity of lightning strikes to migration must also be established. The data provided on property damage from lightning strikes is combined into a single variable, and the specific nature of the property damage is abstracted, frustrating efforts to separate electronic and non-electronic damage. However, a study in partnership with State Farm stated that over fifty percent of lightning damage claims were related to electrical surges damaging components or wiring, and that power surges from transformers or service line shorts were significant factors. It also noted that direct damage from lightning strikes, such as fires, are covered by standard insurance policies while only some policies cover electronic and electrical system damage specifically (Insurance Information Institute, 2017). Therefore, damages from non-electronic damage can be mitigated through a standard insurance policy, while damage to electronic damage requires a different policy with more coverage, and therefore higher premiums, which would be reflected in the relative costs of living in counties with frequent lightning storms, compared to counties with fewer storms. Additionally, more evidence comes from an analysis of the average lightning strike density and net migration; lightning prone Texas and Florida both enjoy an inflow of migrants into their states. Given that these states are considerably more affluent than their immediate neighbors (SSTI, 2016), it stands to reason that the ability to withstand higher costs associated with insuring, maintaining and replacing electronics for these states are relatively higher, and thus the decision to migrate to these areas is less affected by storms. With these points in mind, the exogeneity of lightning strikes to migration can be reasonably assumed.

Time Period Determination

The time period was determined by the constraints of data availability and the effects of the COVID-19 pandemic. Statistics on household internet connectivity are only available from the FCC from 2008 onwards, so the earliest possible start time is that year. Migration statistics through the IRS, while extending back to 1990, change data collection methodology in the year 2011. Before 2011, this data consisted of all tax returns collected by the end of September. Beginning in 2011, the data set was expanded to consist of all returns collected in a calendar year, resulting in an increase in the number of matched returns by 5%, and a 25% increase of high-income returns. While there is only a marginal increase in the number of collected returns, the time-period will be restricted to after this change in methodology to prevent any inaccurate estimates due measurement error of the migration rate over time caused by a discontinuous jump.

Although data for the variables of interest extend beyond 2019, the COVID-19 pandemic induced a series of shocks which disrupts estimates of a cross-sectional analysis including those time spans. During the pandemic, the federal government provided significant subsidies through the Emergency Broadband Benefit Program and Affordable Connectivity Program for both consumers and internet service (Federal Communications Commission, 2023). Subsequently, all counties experienced a sharp upturn in internet connectivity during this time. In terms of migration, widespread urban flight was observed American cities noticeably depopulated into the countryside (Rogers et al., 2023). Even the number of lightning strikes were affected; lightning was noted to be significantly lower across the country because of the pandemic. Lightning strikes are partially influenced by aerosol particles from car emissions are present, and during the closures and social distancing requirements over the pandemic, there were significantly lower aerosol emissions due to cars, which lead to a decrease in lightning during lockdown (Yusifandika et al., 2021). In summary, the pandemic simultaneously sharply decreased the number of lightning strikes, the level of internet connectivity increased dramatically, and internal migration from urban environments to rural areas accelerated. Therefore, the time-period covering the COVID-19 pandemic will be excluded from the study to preserve the identification strategy.

Regression Model

Overall Model

The overall model for the two-stage least-squares regression is described by equations (4) and (5), with equation (4) describing the first stage of the regression and equation (5) describing the second. The first stage of the regression estimates the level of internet connectivity in the year 2011 using the average lightning density between the years 2000 to 2010. The second stage uses that estimate to determine the effect of internet connectivity on migration rate, which represents the net, outflow and inflow migration rates in three different regressions. Additionally, regressions will be run incorporating an interaction term between the explanatory variables and a rural dummy variable, which will be elaborated later in this chapter.

First Stage Equation

The first stage of the econometric model is as follows:

$$(4)\ Internet\ \widehat{Connectivity}_{c,2011} = Log\ (Strike\ Density_{c,2000-2010})\beta + X_{c,2010}\beta + \eta$$

"Internet Connectivity" denotes the number of connected individuals in county *c* in the year 2011. "Strike Density" represents the average number of ground flashes per square kilometer for each county, which is log transformed to account for the wide range of recorded values across counties at any given point in time. *X* denotes a vector of control variables at the county level in the year 2010, and ε is the error term. The vector *X* includes the relative weekly wage of counties to the state they reside, the proportion of private industries engaged in the provision of services, the proportion of individuals in the county with at least a high school level education, the proportion with some college education and the proportion which had attained a four year degree, as well as whether the county is classified as rural by the Office of Management and Budget.

Second Stage Equations

The second stages of the econometric model are described as follows:

$$(5)\ Annual\ Net\ Migration\ Rate_{c,2011-2019} = Internet\ \widehat{Connectivity}_{c,2011}\beta + X_{c,2010}\beta + \varepsilon$$

$$(6)\ Annual\ Outflow\ Rate_{c,2011-2019} = Internet\ \widehat{Connectivity}_{c,2011}\beta + X_{c,2010}\beta + \varepsilon$$

$$(7)\ Annual\ Inflow\ Rate_{c,2011-2019} = Internet\ \widehat{Connectivity}_{c,2011}\beta + X_{c,2010}\beta + \varepsilon$$

The dependent variables across three separate regressions are the average annual net migration rate (5), the outflow rate (6) and the inflow rate (7). The estimated internet connectivity for each county in the year 2011 is the main independent variable of interest, and instrumented from the first stage. The vector of county level controls *X* is identical to the controls used in the first stage regression. Here the possibility of systemic differences between counties may bias the estimate of the coefficient of the regressor of interest is more troubling. It is reasonable to assume that county level policies and socioeconomic conditions could result in variation between counties that would affect both internet connectivity and migratory patterns and thus result in bias. However, the model risks overspecification with the addition of too many control variables, and the use of an instrumental variable with a long-difference regression model is designed alleviate bias associated with a systemic difference between counties, which allows a causal relationship to be established (Hahn et al., 2007).

Construction of Migration Rate Variables

The migration measures for the second stage of the two-stage least-squares model are constructed as follows:

(8) $Population_{c,2010} = Nonmigrants_{c,2010} + In\ Migrants_{c,2010} - Out\ Migrants_{c,2010}$

(9) $Annual\ Net\ Migration\ Rate_{c,2011-2019} = \left(\frac{\sum_{2011}^{2019} In\ Migrants_{ct} - \sum_{2011}^{2019} Out\ Migrants_{ct}}{Population_{c,2010}\ *}\right) * \frac{1}{9}$

(10) $Annual\ Inflow\ Migration\ Rate_{c,2011-2019} = \left(\frac{\sum_{2011}^{2019} In\ Migrants_{ct}}{Population_{c,2010}}\right) * \frac{1}{9}$

(11) $Annual\ Outflow\ Migration\ Rate_{c,2011-2019} = \left(\frac{\sum_{2011}^{2019} Out\ Migrants_{ct}}{Population_{c,2010}}\right) * \frac{1}{9}$

Equation (8) describes how the baseline population in the year 2010 for each county is constructed. This baseline is necessary to normalize the migration rate statistics to the population of each county. The baseline population is calculated by adding the number of in-migrants and subtracting the number of out-migrants from the number of non-migrants of each county in the year 2010. Equation (9) describes how the annual net migration rate of each county is calculated by combining the total number of in-migrants for each county over the years 2011 to 2019, subtracting the total number of in-migrants by the total number of out-migrants for the same time-period, then dividing the resulting number by the baseline population of the county in the year 2010. This result is also divided by the number of years in the cross-section time-period (which is to say, 9) to create the average annual net migration rate for each county in the time-period, which is done to enhance clarity when interpreting the results. Equations (10) and (11) describe how the average annual inflow and outflow migration rates are calculated; these are calculated by dividing the total number of in-migrants by the baseline population and the number of years in the cross-section, and the total number of out-migrants by the baseline population and the number of years respectively.

Regressions with Rural Interaction Terms

To ascertain the differential effect that lightning strikes have on internet connectivity for rural as opposed to urban areas, as well as the differential effect that internet connectivity has on migration on these areas, separate regressions which include interaction terms between these explanatory variables and a rural dummy variable were also conducted. For the sake of brevity, only the changes to the overall model will be displayed, with equation (12) showing the first stage, and equation (13) showing the second. For the second stage represented by equation (13), "Migration Rate" is replaced with the net, outflow and inflow migration rate statistics defined by equations (7), (8) and (9) to construct the second stage specifications used for these regressions.

(12) $Internet\ \widehat{Connectivity}_{c,2011} = Log(Strike\ Density_{c,2000-2010})\beta +$
$(Log(Strike\ Density_{c,2001-2010}) * Rural_{c,2010})\beta + X_{c,2010}\beta +\ \eta$

(13) $Migration\ Rate_{c,2011-2019} = Internet\ \widehat{Connectivity}_{c,2011}\beta +$
$(Internet\ \widehat{Connectivity}_{c,2011} * Rural_{c,2010})\beta + X_{c,2010}\beta +\ \varepsilon$

Change in Wages

Additionally, regressions to determine the effect of internet connectivity on the change in wages for each county were also ran to create a more robust understanding of the underlying mechanism of migration. While it should be noted that the pre-existing level in relative wages was used as a control variable for the main regressions on migration rates, this was done in those cases to determine the relative economic health for counties in the period before the study. These regressions are to determine if there is a significant impact on the change in wages over time due to prior variations of internet penetration in each county in the year 2011. In these regressions, the overall set up is a 2SLS model which is similar to the regressions described in equations (4), (5), (12) and (13). However, the dependent variable for the second stage described by equations (5) and (13), is the annual average percent change in wages in the time frame, and the relative wage was dropped from the vector of control variables. The construction of the dependent variable is described by equation (15). The second stage equations to estimate the effect of internet connectivity on wages using the main regression model and the interaction term model are equations (16) and (17), respectively.

(15) $Annual\ Avg.\%\Delta\ Wages_{c,2011-2019} = \left(\frac{Wage_{c,2019} - Wage_{c,2011}}{Wage_{c,2011}}\right) * 100\% * \frac{1}{9}$

(16) $Annual\ Avg.\%\Delta\ Wages_{c,2011-2019} = Internet\ \widehat{Connectivity}_{c,2011}\beta + X_{c,2010}\beta + \varepsilon$

(17) $Annual\ Avg.\%\Delta\ Wages_{c,2011-2019} = Internet\ \widehat{Connectivity}_{c,2011}\beta +$

$(Internet\ \widehat{Connectivity}_{c,2011} * Rural_{c,2010})\beta + X_{c,2010}\beta + \varepsilon$

CHAPTER FOUR

DATA

Primary Variables of Interest

Internet Connectivity

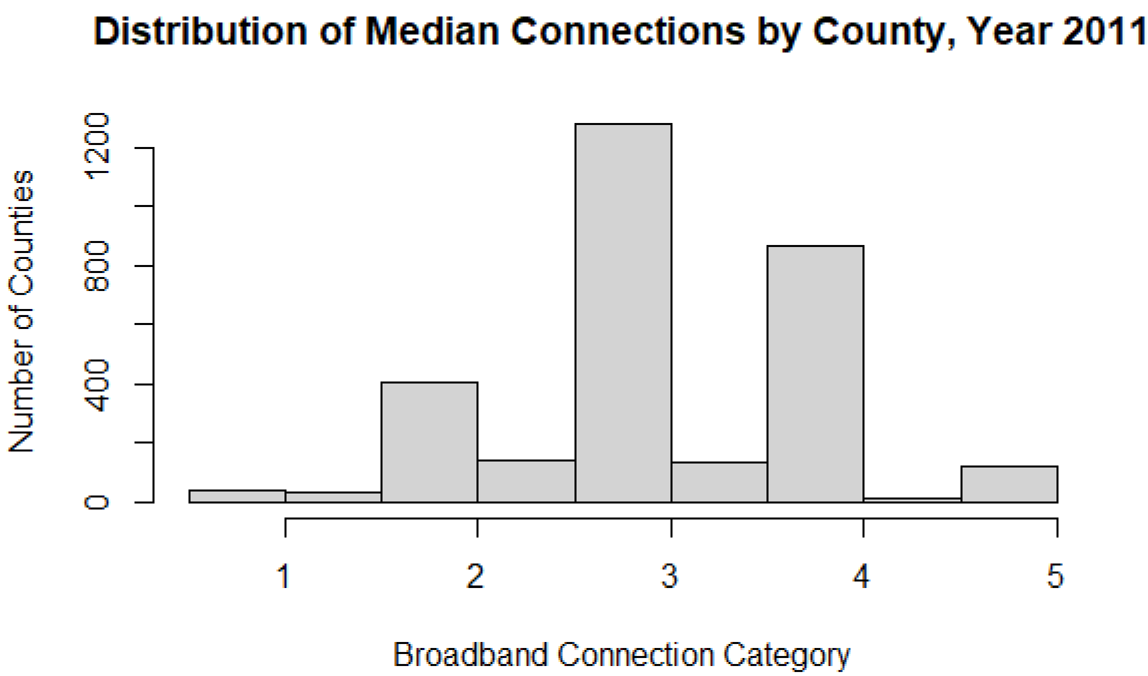
Internet connectivity data is sourced from FCC Form 477. Form 477 provides a record of the number of households with broadband connections of at least 200 Kbps download speeds. Connections are reported categorically on a scale from zero to five:

- A value of 0 denotes that no households are connected to the internet at 200 Kbps download speeds.
- Category 1 denotes 0-200 households connected per 1000 households,
- Category 2 denotes 200-400 households connected per 1000 households,
- Category 3 denotes 400-600 households connected per 1000 households,
- Category 4 denotes 600-800 households connected per 1000 households,
- Category 5 denotes 800-1000 households connected per 1000 households,

Beginning in 2008, the FCC began recording the categories of internet download speeds of at least 200 Kbps downstream at the census tract level, the distribution of which across U.S. counties in the year 2011 are illustrated in Figures 1 through 3. For the purposes of this thesis, the 200 Kbps benchmark in 2011 will be used as the primary measure of internet connectivity. While there is even more detail in the data set starting in 2014, time restrictions become an issue as the time-period would become restricted to the 6-year time span between 2014 to 2019. The benefits of using these more granular measurements would also be minimal, as there is already sufficient variation in the level of connectedness across counties at the 200 Kbps download speed threshold to allow for a regression analysis. As the data records the level of internet connectivity on the census tract level, the level of internet connectivity will be aggregated to the county level by taking the median connectivity of the county. This is done, as opposed to the mean, due to the prevalence of skewed estimates, especially in rural counties. To quantify this skewness, of the 2705 counties with more than one census tract, 591 counties have a skewness value less than -1 or greater than 1, which is 21.8% of counties with more than one census tract. 100 counties with more than one census tract have a skewness which is less than -2 or greater than 2. In many counties, it is the case that few census tracts in a county are connected to broadband at a higher level, while the rest of the county is not. Additionally, as the level of internet connection is categorical in nature and the purpose of the variable is to represent the typical level of internet connectivity in a county, therefore using the 50th percentile enhances the interpretability of the estimates. As such, out of an abundance of caution with respect to skewness and to maintain the clarity of the results, the median of each county's connectivity level is used for the regression estimates. Since the median for each county is being calculated, the resulting measure is occasionally 0.5, 1.5, 2.5, 3.5, or 4.5. These values would represent that the general level of connectivity is somewhere between two categories.

Figure 1 shows the distribution of internet connections by county in the United States in the year 2011. At this time, the most common connection categories were 3 and 4, which denotes a level of connectivity of 400-600 households and 600-800 households respectively. This shows that, while 200 kbps download speeds would be considered by consumers to be very slow even by the standards at the time, a significant portion of counties were unable to reach this threshold in the year 2011.

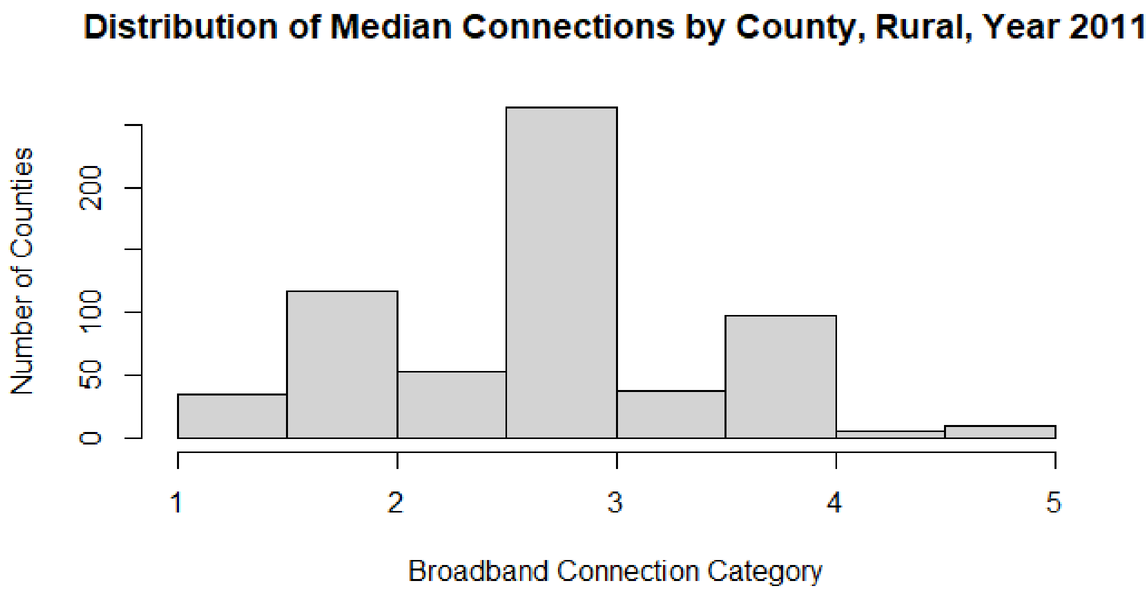
Figure 1: Histogram shows the distribution of broadband connections by connection category in the year 2011.



Source: The Federal Communications Commission, Form 477. The distribution is centered around 3. The data shows signs of right-ward skew.

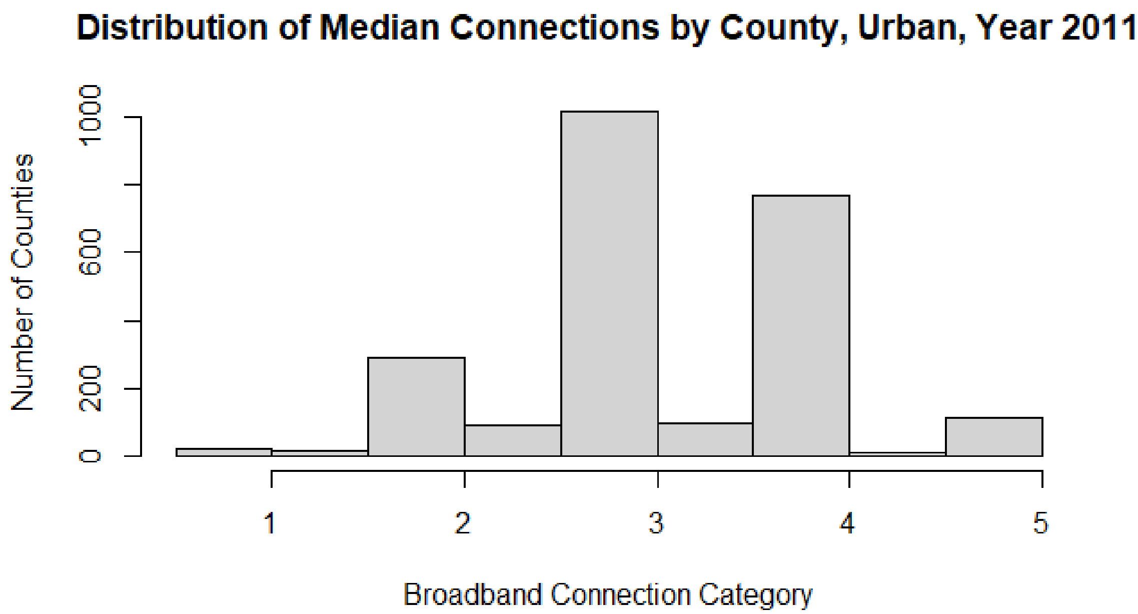
Figures 2 and 3 show the distribution of broadband connections by connection category for rural and urban counties in the United States, respectively. These figures show that rural counties overall have fewer household connections when compared to urban counties. Additionally, they show that rural counties had a more variation in household connections when compared to urban counties. However, it also shows that overall, the level of connections is distributed around categories 3 and 4 in the year 2011.

Figure 2: Histogram which shows the distribution of rural broadband connections by connection category in the year 2011.



Source: The Federal Communications Commission, Form 477. The distribution is centered around 3 and is shifted more to the left than the overall distribution.

Figure 3: Histogram which shows the distribution of rural broadband connections by connection category in the year 2011.



Source: The Federal Communications Commission, Form 477. The distribution is centered around 3 and is shifted more to the right than the overall distribution.

Internal Migration

Migration data for this study is sourced from the Internal Revenue Service (IRS) which compiles year-to-year address changes based on individual income tax returns filed in the previous year. [1]

The IRS data set consists of county-to-county migration flows, including:

- The number of tax returns filed, which approximates the number of migrating households.
- The number of personal exemptions claimed, which approximates the number of migrating individuals.

[1] The cleaned migration data set used for this thesis can be found on GitHub, at DOI: 10.5281/zenodo.15292816.

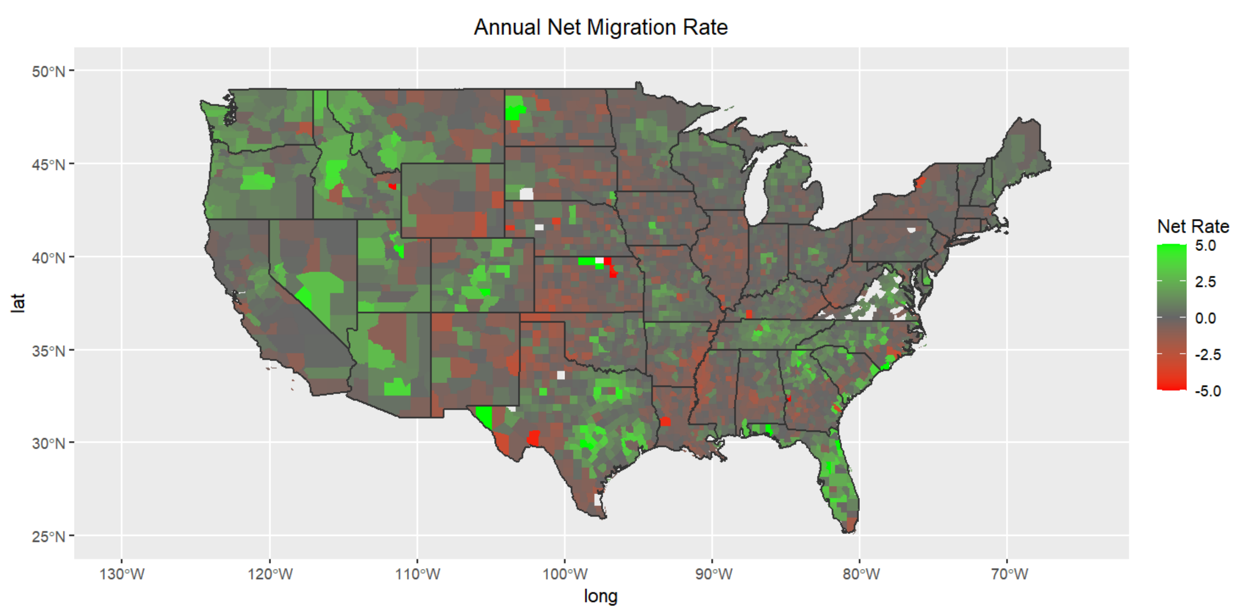
- Total adjusted gross income (AGI), which estimates total income-based mobility patterns.

A notable limitation of the data, aside from the methodological discontinuity in 2011, is the potential underrepresentation of those less likely to file tax returns, particularly low-income individuals, students, young adults, the elderly and the retired. Additionally, the IRS only records migration from each county if there 10 or more exemptions in either direction; for example, if a county received 9 in-migrants and 11 out-migrants, in-migration will be reported as zero for that county, while out-migration will be recorded as 11. As a result, counties with very low populations and counties with few migrants in or out may be systemically inaccurate. However, the number of is limited cases and should not interfere with the regression estimates; of the over 28,000 observations used to construct the IRS migration data over the 9-year timespan, only 240 observations have inflow statistics coded as zero, and 168 observations had outflow statistics of zero. In the year 2010 to calculate the baseline population, only 7 counties lacked information to construct the baseline statistic. For the main econometric analysis, personal exemptions are used to construct the migration statistics used in the regressions, as the IRS notes that this most closely records the net internal migrants for each county.

The migration data is divided into two files each year. Each file details the county of origin and the destination county, one data file specifically records the in-migrants for each county, and the other will record the out-migrants for each county. The net outflow and inflow of migrants for a county are calculated by taking the sum of each over 2011 to 2019. The net migration rate, outflow rate and inflow rate are calculated as described by equations (6) through (9) in the Econometric Model section.

Figure 4 shows the annual net migration rate for the continental United States between the years 2011 to 2019, where darker colors indicate a net flow out of the county, and brighter colors indicate a net flow into the county. It shows that there is overall a net flow of internal migrants out of the South, the Midwest and Rust Belt areas, and that counties in the Rocky and Appalachian Mountains and the West Coast have a net flow in. It also shows that rural counties in both regions have the highest rates in either direction, and that rural counties in the Rockies and the Appalachians overall had a positive net migration rate, while much of the lightning prone region had net negative migration. The exceptions are Florida and Texas, both of which have relatively large and robust economies compared to their immediate neighbors, as noted earlier.

Figure 4: Map illustrating the spatial variation in annual migration rate across U.S. counties between 2011 and 2019.



Source: Internal Revenue Service. Counties with net migration rates greater than -5 or less than 5 are truncated to -5 and 5 respectively for improved visual clarity. Counties in mountainous regions, the Pacific Northwest, Florida and Texas experienced net inflows of migrants, while much of the Great Plains counties experienced outflows.

Lightning Strike Data

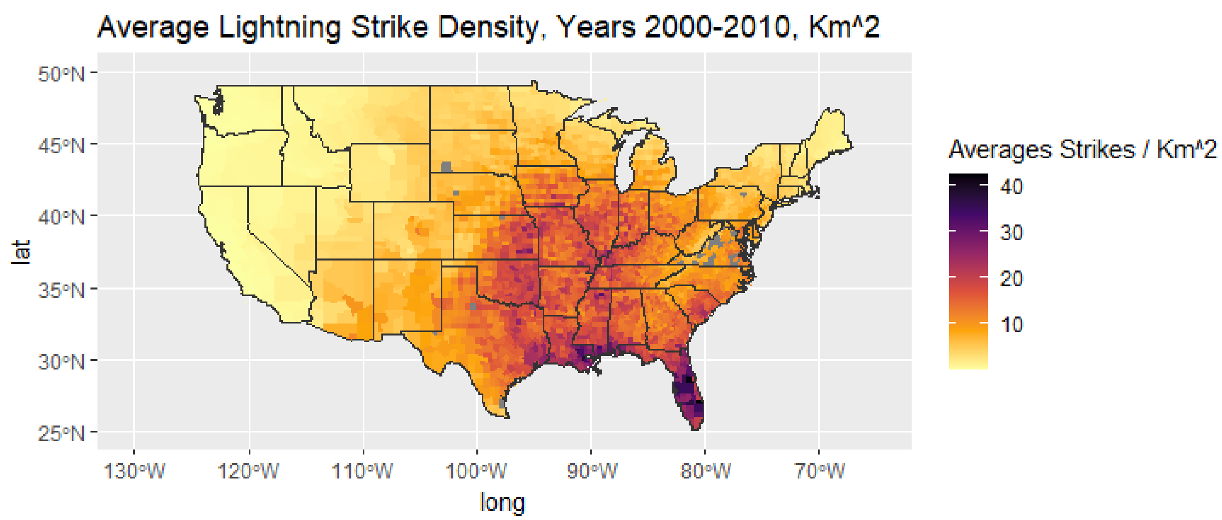
Data on ground lightning strikes is sourced from the U.S. National Lightning Detection Network (NLDN), a commercial lightning detection network operated by Vaisala. This data is made accessible to the public through National Oceanic and Atmospheric Administration's Severe Weather Data Inventory.

The NLDN lightning strike data consists of all cloud to ground lightning flashes detected across the United States. Each observation records the date, number, latitude and longitude of each strike. This data was spatially aggregated to the county level by mapping the latitude and longitude of the strikes onto each U.S county. This was done using shape files sourced from ArcGIS.

To compute strike density, the average number of strikes per square kilometer for each county was calculated and joined to the lightning dataset with the U.S. Census' shape files for the year 2010, to determine the county which each lightning strike fell in and the U.S. Census' estimated total land area in square kilometers for each county. The number of lightning strikes for each year between 2000 and 2010 in a county is divided by the total land area to determine the lightning strike density for that year. The average annual lightning strike density for each county is calculated, and that average for each county is to estimate internet connectivity.

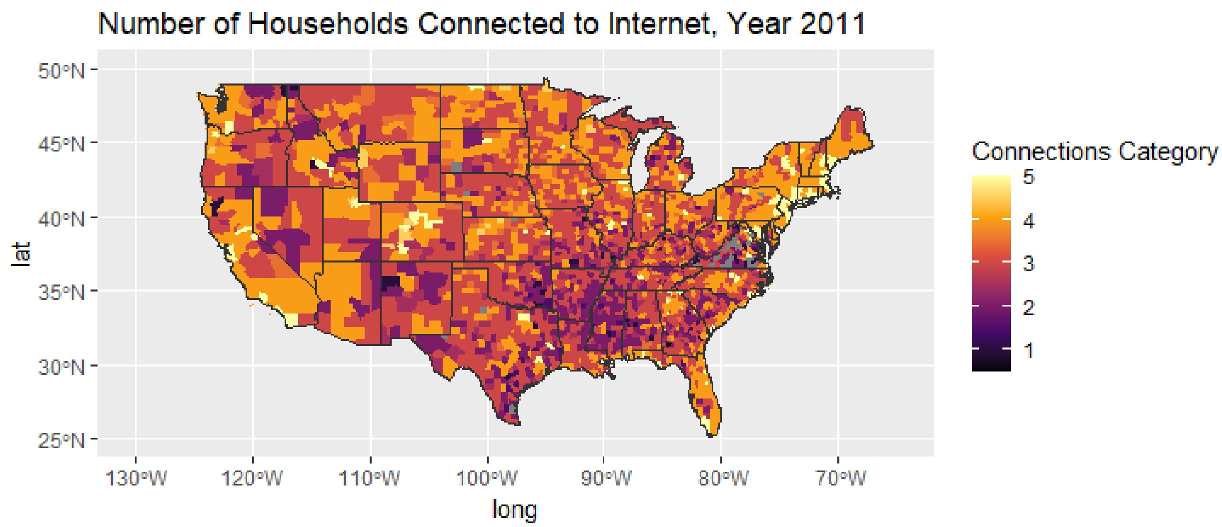
The comparison between lightning strike density and internet connectivity is shown in Figures 5 and 6. Figure 5 shows areas of less frequent lightning strikes in brighter yellow, and greater lightning frequency in dark purple / black. It shows that lightning strike density is the highest in the South and especially in Florida. The pattern of lightning strikes follows the Mississippi river basin, with lightning strikes decreasing in frequency the further north and the further west areas are relative to Florida. Figure 6 depicts the number of households connected to the internet at 200 kbps download speeds, using the categorical variable described above, where lighter areas reflect greater levels of connectivity. The Figures show that, while there is significant spatial variation in connectivity, areas that are prone to lightning have significantly lower levels of internet access compared to areas with less lightning.

Figure 5: Map depicting the average lightning strike density between the years 2000 and 2010, in number of ground flashes per km^2.



Source: National Oceanic and Atmospheric Administration. Lightning strikes are most frequent in the south, particularly in Florida and Louisiana, and spread north and west in a gradient. The areas with the least number of lightning strikes are the Pacific Northwest and northern New England.

Figure 6: Map depicting the number of households connected to the internet by FCC connection category, in the year 2011.



Source: Federal Communications Commission, Form 477. There is significant variation in connections between counties, but the South has significantly fewer connections overall than other parts of the country.

<u>Control Variables</u>

A set of control variables are included to account for county-level variation which may influence domestic migration or internet connectivity, to more accurately estimate the causal effect of internet access. All statistics are fixed to the year 2010 to provide a baseline of economic health, population skills and industry specialization at the beginning of the long-difference regression.

County level unemployment statistics are provided by the Bureau of Labor Statistics and represented as a percentage of the labor force which is unemployed. It should be noted that at this time, the world was still reeling from The Great Recession. As such, many of the unemployment statistics are sharply elevated compared to the other years when the unemployment rate was reported. Wage data is obtained from the Bureau of Labor Statistic's Quarterly Census of Employment and Wages (QCEW). This data is an average of weekly wages across all industries in each county for each year. For the main regression estimate, the relative wage of each county in relation to their state in the year 2010 will be used to gauge the value of migrating into that county. For the regression estimate to determine the effect of internet connectivity on wages, the change in wages will be used as the dependent variable. Education attainment statistics are sourced from the U.S. Census, the collected data of which is made available through the U.S. Department of Agriculture's website and reports the percentage of individuals as of the year 2010 who attained a less than a high school diploma, at least a high school diploma, some college education (equal to a two-year degree), or a four-year college degree.

Data on the industrial composition in each county is sourced from the Bureau of Economic Analysis (BEA). Private sector industries are aggregated into two broad categories:

- Private sector industries engaged in services and
- Private sector industries engaged in goods

Excluded industries are those which are classified as government industries, as well as industries which do not fit neatly into either category. Industries are classified as "government" industries if they are related to the executive, judicial, administrative and regulatory activities of federal, state or local governments, including the military. For the regression models, only the aggregated variable for private sector industries engaged in providing services will be included.

Table 1: Summary Statistics.

Variable	Year	Mean	SD	Max	Min	Source
Net Migration Rate	2011 – 2019	0.170	1.381	11.14	-37.90	IRS Migration Data
Outflow Rate	2011 – 2019	6.544	2.487	55.03	0.000	
Inflow Rate	2011 – 2019	6.714	2.743	42.59	0.708	
Connections Category at 200 Kbps	2011	3.197	0.809	5.000	0.500	FCC Form 477
Avg. Strike Density	2000 - 2010	11.03	6.555	42.39	0.015	National Lightning Detection Network
Rural	2010	0.204	0.403	1.000	0.000	U.S. Census
Unemployment Rate	2010	9.346	3.148	29.40	2.000	Bureau of Labor Statistics
Relative Wage	2010	0.989	0.182	2.429	0.543	Quarterly Census of Employment and Wages
Only High School Education	2010	35.11	6.884	55.10	9.069	U.S. Census
Some College Education	2010	29.69	5.299	48.82	11.24	
Bachelor’s Education or More	2010	19.30	8.526	71.25	5.432	
Proportion of Industries, Private and Providing Services	2010	0.469	0.229	0.966	0.000	Bureau of Economic Analysis

Summary statistics for main regressions.

CHAPTER FIVE

RESULTS

First Stage Regression Results

The results of the first stage of the two-stage least-squared regression, displayed in Table 2, show that average lightning strike density has a significant negative impact on the level of internet connectivity in a county. To help understand the magnitude of the coefficient, it is important to remember the wide range of lightning strike densities in the United States. For example, the lowest lightning density in the United States between 2000 and 2010 is San Juan County in Washington state, with a strike density of 0.0148 strikes per square kilometer in a year. The regression estimate implies that, for a county to have a full internet connectivity category point lower compared to San Juan, which corresponds with a 20-percentage point decrease in internet penetration in the county, a county would have to have a lightning strike density of 19.47. In the dataset, this would correspond to a county like Daviess, Kentucky. The highest lightning density in the United States is Martin County, Florida, with a lightning density of 42.39; the first stage regression estimates that, all else being equal, the difference in internet connectivity between these two counties should be -2.178, which would correspond to an over 40 percentage point decrease in internet penetration. Across the time-period examined, increases in lighting strikes are inversely proportional to the level of internet connectivity in a county, which supports the proposed mechanism, where the number of lightning strikes in a county actively suppresses the rollout of broadband internet in that county. This result corroborates Anderson (2012), which also found that the rate of IT diffusion in lightning prone regions is significantly slower. This lends credence to the overall validity of the instrumental variable to estimate the level of internet connectivity in a county.

Table 2: First-stage Regression Results: The Effect of Lightning Strikes on Internet Connections in the Year 2010

Variable	Estimate
Log(Average Strike Density)	-0.076***
	(0.012)
Unemployment Rate	-0.014***
	(0.004)
Relative Wage	0.159*
	(0.084)
High School Education	0.023***
	(0.003)
Some College Education	0.025***
	(0.002)
Bachelor's Degree or Better	0.063***
	(0.002)
Proportion of Industries	-0.020
Providing Services	(0.049)
Rural	-0.220***
	(0.029)
Observations	3,047
F-Stat	295.3

Note: OLS regression of log transformed average lightning density between 2000 and 2010 on internet connectivity in the year 2011. All control variables are fixed to the year 2010. Standard errors are in parentheses. *p < 0.1; **p < 0.05; ***p < 0.01.

The Effect of Internet Connectivity on Migration

The relationship between internet connectivity and migration seems to be dependent on whether the county receiving the internet is rural or not. The results of the main regression, as well as the main regression with the inclusion of a rural / internet connection interaction term, are recorded in Table 3. The estimated effect on overall net migration is that an increase of internet penetration by 20-percentage points results in an average annual increase in the net migration rate of 0.1%. This effect, in addition to not being of great magnitude, is indifferent from zero, so the null hypothesis that there is no effect on the net migration rate cannot be rejected. However, the effect of internet connectivity on the inflow and outflow migration rates are of great interest; the effect on inflow and outflow are average annual reductions of about 4.8% and 4.9% respectively. To wit, the effect of an increase of internet penetration by 20-percentage points corresponds to a yearly decrease in people moving at all into or out of a county by almost 5% each year. This is a dramatic drop of mobility, and it seems to imply that the introduction of the internet, all else being equal, results in a less migratory population overall. The fact that it also seems to diminish inflow migration rates is somewhat less intuitive; the results imply that people are less likely to move into areas where there is a higher level of internet connectivity. The exact mechanism for this phenomenon, while outside the scope of this thesis, would be an excellent focus for a follow-up study on this subject.

A robustness check to strengthen the findings of Table 3 were also specified by splitting the dataset into two, one subset containing only counties considered urban, and one with only rural counties. The results are shown in appendix B. The conclusions of those series of regressions are similar those listed in table 3; urban counties do not experience a change in net migration, but an overall decrease in mobility, while the annual average net migration rate of rural counties increases annually, as individuals move into counties with higher levels of internet connectivity.

The regression which includes the rural / internet connectivity interaction term illustrates the distinct effect that internet connectivity has on rural counties, relative to urban ones. In this case, the increase in internet penetration causes the net migration rate to rise on average by 0.7% annually, and this result is highly statistically significant. Likewise, the effect of internet penetration also shows that, while there is still an overall decrease in mobility, both the outflow and inflow rates for rural counties are greater than the overall change in mobility. The conclusion to be drawn from this result is that rural counties benefit dramatically more than urban counties from an increase in internet penetration. This may be due to the relatively greater effect on local job availability for rural counties; urban counties may not necessarily gain more career opportunities from the introduction of technologies designed to bridge geographic limitations, but rural counties which would not otherwise have had those opportunities would certainly benefit. This result should be of great interest to rural communities which seek to expand their population, and rural communities which are expanding their internet infrastructure that wish to prepare themselves for a potential influx of internal migrants.

As an additional robustness check, the same regression specifications were run on subsets of the data which contained only lightning prone states, defined as states with at least one county with an average lightning density of one strike per square kilometer per year. Table 4 displays the results of these regressions and shows that there is a significant decrease in the overall net migration rate of 1.2% annually with a 20-percentage point increase of internet penetration. This test, however, also finds that rural counties experience a lesser net migration outflow over the time-period as the main regression, which does not exclude states where lightning is infrequent. The difference between the net migration estimate of the lightning-prone subset and that of the nation-wide regression may be attributed to the exclusion of the west coast and New England from the lightning prone subset; the major tech hubs of Washington and southern California are excluded from the lightning prone subset, as well as the financial capital of New York City. Regardless, the results as they pertain to rural counties of the previous regression are reinforced by these findings, as is the finding that overall mobility dramatically decreases.

Instrument tests further support the validity of the model. Weak instrument tests are conducted for each regression to determine if the strength of the instrument is sufficient, with a score equal to or greater than 10 denoting a strong instrument. In each regression, the results of these tests reveal that the instrument is always strong, which mitigates concerns of bias due to potential weak instruments. Wu-Hausman tests for the net migration rate regression in the main regression indicate that there may be an issue with efficiency for that model. However, it should be noted that the Wu-Hausman tests have a considerable number of issues related to their reliability. It has been noted that the Wu-Hausman test determines not the exogeneity or endogeneity of specific components, but rather the effects on estimates of the coefficient of any possible endogeneity (Davidson and MacKinnon, 1989). Nevertheless, the results of Wu-Hausman test statistics are included in the regression summary tables to be thorough and to lend support to the efficiency of the models.

Table 3: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Migration Rates.

	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	0.097 (0.344)	-4.838*** (1.069)	-4.935*** (1.034)
Observations	3,047	3,047	3,047
F-Stat	42.233	42.23	42.23
Wu-Hausman	0.032	53.38	67.53
Regression with Connectivity / Rural Interaction Term			
	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	-0.180 (0.359)	-5.498*** (1.155)	-5.318*** (1.234)
Connections at 200kbps * Rural	0.693** (0.282)	1.648** (0.765)	0.955 (0.756)
Observations	3,047	3,047	3,047
F-Stat (Connections)	21.95	21.95	21.95
F-Stat (Connections * Rural)	160.81	160.81	160.81
Wu-Hausman	4.194	30.44	34.26

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Control variables used include the unemployment rate, educational statistics, whether the county is rural, the proportion of industries engaged in providing services, and the wage of each county relative to their state. Standard errors are clustered heteroskedasticity robust and in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01

Table 4: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Migration Rate, Lightning Prone States.

	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	-1.216*** (0.437)	-6.799*** (1.340)	-5.582*** (1.102)
Observations	2,621	2,621	2,621
F-Stat	43.26	43.26	43.26
Wu-Hausman	14.21	109.30	85.79
Regression with Connectivity / Rural Interaction Term			
	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	-1.604*** (0.544)	-7.651*** (1.745)	-6.047*** (1.382)
Connections at 200kbps * Rural	0.673** (0.310)	1.480 (1.031)	0.807 (0.822)
Observations	2,621	2,621	2,621
F-Stat (Connections)	24.08	24.08	24.08
F-Stat (Connections * Rural)	208.65	208.65	208.65
Wu-Hausman	8.024	51.98	40.08

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Control variables used include the unemployment rate, educational statistics, whether the county is rural, the proportion of industries engaged in providing services, and the wage of each county relative to their state. Standard errors are clustered heteroskedasticity robust and in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01

Effect of Internet Connectivity on Wages

The results of the two-stage least-squares regression estimating the effect of internet connectivity in 2011 on annual average wage changes over the period between 2011 to 2019 are displayed on Table 5. The result of the regression without the instrumental variable is recorded on the first column, and the result of the regression with the instrumental variable is listed on the second. The results show that an increase in internet penetration by 20-percentage points results in an annual average increase in wages by about 1.843% year over year. It also shows that when the effect of internet penetration on rural communities is separated from the overall effect, the overall effect is marginally higher. However, it should be noted that a partial F test between the two model specifications reveals that the two models are not statistically different. The model including an interaction term between internet connectivity and whether the county is rural is therefore not necessarily biased, but the estimator is less efficient than the model without the interaction term. These results imply that the earnings for individuals in counties with a higher pre-existing level of internet connectivity grow faster than those with fewer connections, yet in conjunction with the results of the migration regressions seem to imply that the faster growing wages are not enough to induce individuals to move into the area. This may be due to a well-reported inequality of wage increases due to internet diffusion, where the wages of skilled workers rise disproportionately compared to unskilled workers (Atasoy, 2013). Most workers would not perceive significant increases in earnings but would still have to pay more for housing and consumer goods (Hawk, 2013). Based on the theoretical model, those workers would be disinclined to move into city environments.

Table 5: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Annual Average Percent Wage Change

Dependent Variable: Average Annual Change in Wages (2011 – 2019)

Variable	(1)	(2)
Connections at 200kbps	1.843***	2.042***
	(0.386)	(0.441)
Connections at 200kbps * Rural		-0.435
		(0.327)
Unemployment Rate	-0.040***	-0.046***
	(0.011)	(0.013)
High School Education	-0.045***	-0.047***
	(0.011)	(0.012)
Some College Educations	-0.053***	-0.054***
	(0.015)	(0.015)
Bachelor’s Degree or Better	-0.126***	-0.137***
	(0.026)	(0.029)
Proportion of Industries	0.151	0.159
Providing Services	(0.125)	(0.129)
Rural	0.573***	1.869*
	(0.118)	(0.985)
Observations	3,047	3,047
		+321
F-Stat (Connections)	38.71	20.15
F-Stat (Connections * Rural)		163.39
Wu-Hausman	50.00	25.44
Partial F-Test		2.078

Note: Regression of internet connectivity on annual, average wage changes between 2011 and 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust, and in parentheses. F-Stat denotes the results of weak instrument tests. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

CHAPTER SIX

CONCLUSION

Analysis of the regression results reveal that the overall effect of increased internet penetration in the year 2011 is an insignificant change in net migration over the time-period between 2011 to 2019, but a dramatic decrease in mobility. However, rural counties experience an increase in average annual net migration over the time-period by an estimated 0.693% year-over-year, resulting from an additional 20 percentage point increase in internet penetration. The overall annual change in wages due to this increase in internet penetration is around 1.843% year-over-year.

This experiment has demonstrated that, for most counties, higher internet penetration results in the population moving around less but may not result in a net migration event one way or another. It has also shown that rural counties experience a significant inflow of migrants due to an increase in internet availability. This result is of great interest to those who wish to develop the infrastructure of rural counties, particularly the federal government, which has been noted to have spent considerable resources to expand internet infrastructure to rural counties. Developers must plan around the idea that, when there is a discrepancy in internet connectivity between rural counties and all else is held equal, individuals will be more likely to migrate into counties with higher internet access. This may result in strain being placed on public services and the local housing market, and the prudent planner would do well to prepare for this.

Future studies would benefit from analyzing the effect of new satellite technologies on rural broadband impact. While satellite internet in the past has been considered unreliable by the NTIA (Reliable Broadband Service & Alternative Technologies, 2024), and is currently much slower than broadband, advances in satellite technology and availability may make wireless internet transmission reliable enough to be competitive. At that point, it would be of interest to governments and ISPs whether the introduction of a viable competitor would change the calculus of migration. A future study which also analyzes the mechanism around the decrease of the overall as internet penetration increases would also be an excellent contribution to IT economic literature.

REFERENCES CITED

Aghion, Phillippe, et al. “The Effect of Financial Development on Convergence: Theory and Evidence.” *The Quarterly Journal of Economics*, no. 1, Oxford University Press (OUP), Feb. 2005, pp. 173–222. *Crossref*, doi:10.1162/0033553053327515.

Agrawal, David. “The Internet as a Tax Haven?” *American Economic Journal: Economic Policy*, Nov. 2021, doi:10.1257/pol.20170094.

Allegretto, Sylvia, and Steven Pitts. *The Great Recession, Jobless Recoveries and Black Workers*. Joint Center for Political and Economic Studies, https://laborcenter.berkeley.edu/pdf/2010/the-great-recession.pdf. Accessed 28 Mar. 2025.

Anani, Namir. “Paving the Way for the Future of Work.” *Canadian Public Policy*, no. S1, University of Toronto Press Inc. (UTPress), Nov. 2018, pp. S167–76. *Crossref*, doi:10.3138/cpp.2018-012.

Andersen, Thomas Barnebeck, et al. “Lightning, IT Diffusion, and Economic Growth Across U.S. States.” *Review of Economics and Statistics*, no. 4, MIT Press - Journals, Nov. 2012, pp. 903–24. *Crossref*, doi:10.1162/rest_a_00316.

Atasoy, Hilal. “The Effects of Broadband Internet Expansion on Labor Market Outcomes.” *ILR Review*, Elsevier BV, Apr. 2013. *Crossref*, doi:10.2139/ssrn.1890709.

Bajari, Patrick, and Ali Hortaçsu. “Economic Insights from Internet Auctions.” *Journal of Economic Literature*, no. 2, American Economic Association, May 2004, pp. 457–86. *Crossref*, doi:10.1257/0022051041409075.

Bartel, A., et al. “How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills.” *The Quarterly Journal of Economics*, no. 4, Oxford University Press (OUP), Nov. 2007, pp. 1721–58. *Crossref*, doi:10.1162/qjec.2007.122.4.1721.

Bazelyan, E. M., and M. M. Drabkin. “Scientific and Technical Basis for Preventing Lightning Strikes to Earthbound Objects.” *2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491)*, IEEE, pp. 2201–08. *Crossref*, doi:10.1109/pes.2003.1270960. Accessed 27 Mar. 2025.

Bhowmik, Sharit. “Determinants and Consequences OfMigration.” *Economic and Political Weekly*, no. Vol. 19, No. 14, Apr. 1984, pp. 589–90.

Bhuller, Manudeep, et al. “Broadband Internet: An Information Superhighway to Sex Crime?” *SSRN Electronic Journal*, Elsevier BV, 2011. *Crossref*, doi:10.2139/ssrn.1835314.

Biedny, Christina, et al. “Do Gigabits Mean Business? ‘Ultra-Fast’ Broadband Availability’s Effect on Business Births.” *Information Economics and Policy*, Elsevier B.V., 2024, doi:10.1016/j.infoecopol.2024.101094.

Boucher-Robinson, Summer, and Jake Varn. “States Reckon With Lapse of the Broadband Affordable Connectivity Program .” *Pew Trusts*, The Pew Charitable Trusts, 20 Sept. 2024, https://www.pewtrusts.org/en/research-and-analysis/articles/2024/09/20/states-reckon-with-lapse-of-the-broadband-affordable-connectivity-program.

“Broadband Technology Opportunities Program.” *Home Page | National Telecommunications and Information Administration*, National Telecommunications and Information Administration, 9 Dec. 2020, https://www.ntia.gov/category/broadband-technology-opportunities-program.

Bun, Maurice J. G., and Teresa D. Harrison. “OLS and IV Estimation of Regression Models Including Endogenous Interaction Terms.” *Econometric Reviews*, no. 7, Informa UK Limited, Mar. 2018, pp. 814–27. *Crossref*, doi:10.1080/07474938.2018.1427486.

Carpenter, Roy, and Mark Drabkin. *Lightning Strike Protection*. Lightning Eliminators & Consultants, Inc., 1990, http://www.lightningeliminators.com.

Carrington, William, et al. “Migration with Endogenous Moving Costs.” *The American Economic Review*, American Economic Association , Sept. 1996.

Czernich, Nina, et al. “Broadband Infrastructure and Economic Growth.” *The Economic Journal*, Oxford University Press, 2011. *Crossref*, doi:10.2139/ssrn.1516232.

Davidson, Russell, and James G. MacKinnon. “Testing for Consistency Using Artificial Regressions.” *Econometric Theory*, no. 3, Cambridge University Press (CUP), Dec. 1989, pp. 363–84. *Crossref*, doi:10.1017/s0266466600012573.

Detting, Lisa. “Broadband in the Labor Market: The Impact of Residential High-Speed Internet on Married Women’s Labor Force Participation.” *ILR Review*, Mar. 2017, doi:10.1177/0019793916644721.

DiMaggio, Paul, and Bart Bonikowski. “Make Money Surfing the Web? The Impact of Internet Use on the Earnings of U.S. Workers.” *American Sociological Review*, no. 2, SAGE Publications, Apr. 2008, pp. 227–50. *Crossref*, doi:10.1177/000312240807300203.

Elizaga, Juan C. “Internal Migration: An Overview.” *International Migration Review*, no. 2, JSTOR, 1972, p. 121. *Crossref*, doi:10.2307/3002132.

“Emergency Broadband Benefit | Federal Communications Commission.” *Federal Communications Commission | The United States of America*, 2023, https://www.fcc.gov/broadbandbenefit.

Frey, William. *Immigration and Domestic Migration in US Metro Areas: 2000 and 1990 Census Findings by Education and Race*. University of Michigan, 2005.

Frey, William H. “Immigration, Domestic Migration, and Demographic Balkanization in America: New Evidence for the 1990s.” *Population and Development Review*, no. 4, JSTOR, Dec. 1996, p. 741. *Crossref*, doi:10.2307/2137808.

Giannelis, Matthew. “Impact Of The Internet On Modern Society.” *Tech Business News*, 13 June 2024, https://www.techbusinessnews.com.au/blog/impact-of-the-internet-on-modern-society/.

Goldfarb, Avi, and Catherine Tucker. “Digital Economics.” *Journal of Economic Literature*, no. 1, American Economic Association, Mar. 2019, pp. 3–43. *Crossref*, doi:10.1257/jel.20171452.

Greenwood, Michael. “Research on Internal Migration in the United States: A Survey.” *Journal of Economic Literature*, no. Vol. 13, No. 2, American Economic Association, June 1975, pp. 397–433.

Grimes, Arthur, et al. “The Need for Speed: Impacts of Internet Connectivity on Firm Productivity.” *Journal of Productivity Analysis*, Springer, 2012, doi:10.1007/s11123-011-0237-z.

Gunderson, Ronald, and David Sorenson. “An Examination of Domestic Migration from California Counties.” *Journal of Regional Analysis and Policy*, 2010.

Guo, Wu, et al. *The Impact of Near-Strike Lightning on Computers: A Transient Magnetic Field Experiment*. IEEE, June 2024, doi:10.1109/ICCEM60619.2024.10559161.

Hahn, Jinyong, et al. “Long Difference Instrumental Variables Estimation for Dynamic Panel Models with Fixed Effects.” *Journal of Econometrics*, no. 2, Elsevier BV, Oct. 2007, pp. 574–617. *Crossref*, doi:10.1016/j.jeconom.2006.07.005.

Hawk, William. “Expenditures of Urban and Rural Households in 2011.” *Beyond the Numbers*, Bureau of Labor Statistics, Feb. 2013, https://www.bls.gov/opub/btn/volume-2/expenditures-of-urban-and-rural-households-in-2011.htm.

Hendel, John. “‘People Need to See It’: How Politics Hung up a $42B Biden Internet Buildout .” *Politico*, Politico LLC, 4 Sept. 2024, https://www.politico.com/news/2024/09/04/biden-broadband-program-swing-state-frustrations-00175845.

Hlavac, Marek. “stargazer: Well-Formated Regression and Summary Statistics Tables. R package version 5.2.3.” 2022, https://CRAN.R-project.org/package=stargazer.

“H.R.3684 - Infrastructure Investment and Jobs Act.” *Congress.Gov*, 15 Nov. 2021, https://www.congress.gov/bill/117th-congress/house-bill/3684.

Hyytiä, Nina, and Jukka Kola. “Tourism Policy as a Tool for Rural Development.” *Applied Economic Perspectives and Policy*, no. 4, Wiley, Oct. 2013, pp. 708–30. *Crossref*, doi:10.1093/aepp/ppt025.

Kennan, John, and James Walker. “The Effect of Expected Income on Individual Migration Decisions.” *Econometrica*, no. 1, The Econometric Society, 2011, pp. 211–51. *Crossref*, doi:10.3982/ecta4657.

Krider, E. Philip. “Benjamin Franklin and Lightning Rods.” *Physics Today*, no. 1, AIP Publishing, Jan. 2006, pp. 42–48. *Crossref*, doi:10.1063/1.2180176.

Lee, SeongWoo, and Curtis C. Roseman. “Migration Determinants and Employment Consequences of White and Black Families, 1985-1990.” *Economic Geography*, no. 2, JSTOR, Apr. 1999, p. 109. *Crossref*, doi:10.2307/144247.

Lightning and Static Effects on Industrial Electronics. Emerson Process Management, 2002, https://www.emerson.com/documents/automation/technical-data-sheet-technical-data-sheet-lightning-static-effects-on-industrial-electronics-rosemount-en-74342.pdf.

Metwally, Ibrahim A., and Fridolin H. Heidler. “Computation of Transient Overvoltages in Low-Voltage Installations During Direct Strikes to Different Lightning Protection Systems.” *IEEE Transactions on Electromagnetic Compatibility*, no. 3, Institute of Electrical and Electronics Engineers (IEEE), Aug. 2007, pp. 602–13. *Crossref*, doi:10.1109/temc.2007.902192.

Molloy, Raven, et al. “Internal Migration in the United States.” *Journal of Economic Perspectives*, Elsevier BV, 2011. *Crossref*, doi:10.2139/ssrn.1909383.

Moretti, Enrico. *The New Geography of Jobs*. Houghton Mifflin Harcourt, 2012, pp. 1–19.

“Number Of Lightning-Related Homeowners Insurance Claims Up In 2016, But Severity Down | III.” *III | We Are the Trusted Source of Unique, Data-Driven Insights on Insurance to Inform and Empower Consumers.*, Insurance Information Institute, 15 June 2017, https://www.iii.org/press-release/number-of-lightning-related-homeowners-insurance-claims-up-in-2016-but-severity-down-061517.

NVent ERICO Lightning Protection Handbook. nVent Engineered Electrical & Fastening Solutions, 2009, pp. 68–86.

Read, Anna, and Kelly Wert. “Broadband Access Still a Challenge in Rural Affordable Housing .” *Pew Trusts*, The Pew Charitable Trusts, 8 Dec. 2022, https://www.pewtrusts.org/en/research-and-analysis/articles/2022/12/08/broadband-access-still-a-challenge-in-rural-affordable-housing.

Reap, Ronald M. “Analysis and Prediction of Lightning Strike Distributions Associated with Synoptic Map Types over Florida.” *Monthly Weather Review*, no. 8, American Meteorological Society, Aug. 1994, pp. 1698–715. *Crossref*, doi:10.1175/1520-0493(1994)122<1698:aapols>2.0.co;2.

Reliable Broadband Service & Alternative Technologies. National Telecommunications and Information Administration, 2024, https://broadbandusa.ntia.doc.gov/sites/default/files/2024-01/BEAD_Reliable_Broadband_Service_Alternative_Technologies_Guidance.pdf.

Rogers, Luke, et al. “Two Years Into Pandemic, Domestic Migration Trends Shifted.” *Census.Gov*, Census Bureau, 30 Mar. 2023, https://www.census.gov/library/stories/2023/03/domestic-migration-trends-shifted.html.

Rupasingha, Anil, et al. “Broadband and Rural Development: Impacts of the U.S. Department of Agriculture Broadband Initiatives Program on Saving and Creating Jobs.” *Economic Inquiry*, U.S. Department of Agriculture, 2022, doi:10.1111/ecin.13200.

Ryssdal, Kai, and Sarah Leeson. “In National Broadband Rollout, Rural Landscapes Pose a Challenge.” *Marketplace - Business News & Economic Stories for Everyone*, 27 Aug. 2024, https://www.marketplace.org/story/2024/08/27/in-national-broadband-rollout-rural-landscapes-pose-a-challenge.

Sheldon, Tamara L., and Crystal Zhan. “The Impact of Hurricanes and Floods on Domestic Migration.” *Journal of Environmental Economics and Management*, Elsevier BV, Sept. 2022, p. 102726. *Crossref*, doi:10.1016/j.jeem.2022.102726.

Shivalli, Sanketa. “Lightning Phenomenon, Effects and Protection of Structures from Lightning.” *IOSR Journal of Electrical and Electronics Engineering*, 2016.

Smith, Aaron D., et al. “Computer and Internet Use by Great Plains Farmers.” *SSRN Electronic Journal*, Elsevier BV, 2004. *Crossref*, doi:10.2139/ssrn.711262.

Solomon, Fabricant. “Productivity and Wages.” *Challenge*, JSTOR, 1962, http://www.jstor.org/stable/40718586.

Tickamyer, Ann R., et al. *Rural Poverty in the United States*. Columbia University Press, 2017, pp. 439–48.

Todaro, Michael. “A Model of Labor Migration and Urban Unemployment in Less Developed Countries.” *The American Economic Review*, American Economic Association, 1969.

Uman, Martin A. *All about Lightning*. Courier Corporation, 1986.

“US Cities Database | Simplemaps.Com.” *Interactive HTML5 and JavaScript Maps for Websites | Simplemaps.Com*, https://simplemaps.com/data/us-cities. Accessed 29 Mar. 2025.

Useful Stats: Per Capita Personal Income by State, 2010-2015 | SSTI. SSTI, 31 Mar. 2016, https://ssti.org/blog/useful-stats-capita-personal-income-state-2010-2015.

Young, Makena, and Akhil Thadani. *Low Orbit, High Stakes: All-In on the LEO Broadband Competition*. Center for Strategic and International Studies (CSIS), 2022, pp. 5–9.

Yusfiandika, Fazandra, et al. “Lightning Behaviour during the COVID-19 Pandemic - PMC.” *PMC Home*, National Library of Medicine, 9 Sept. 2021, https://pmc.ncbi.nlm.nih.gov/articles/PMC8573680/.

APPENDICES

APPENDIX A

LIGHTNING STRIKE SCIENCE REVIEW

Basic Mechanics

Lightning is a natural phenomenon that introduces a mechanism by which part of the variation in broadband diffusion across counties can be explained (Anderson, 2012). To fully understand the mechanism by which lightning strikes suppress broadband rollout, the basic science of a lightning strike will be explained in this section.

Lightning strikes occur when a negative electrical charge accumulates in a thunderhead, and a corresponding positive charge builds on the ground. As the air is a poor conductor, for this imbalance to resolve itself, an exceptionally large charge differential is necessary between the thunderhead and the ground. When a sufficiently large charge is accumulated, which can be up to 30,000 kilovolts, a charge neutralization (a strike) occurs which results in an instantaneous electrical current between thunderhead and ground of up to 500 kiloamps (Emerson Process Management, 2002). As a comparison, power distribution lines carry between 7 to 14 kilovolts of electricity. The charge neutralization creates a hemispherical strike zone around the point where the strike occurs, known as the point of discrimination (Uman, 1986).

Electrostatic Damage

Suspended Power Lines.

Aside from the obvious physical damage that can result from a lightning strike, the discharge of a strike results in an electrostatic pulse which travels across any conductor which is within the strike zone. Overhead wires will become charged with the potential of the strike based on the height of the wire relative to the ground. This charge potential will then find the shortest path to the ground, often through electrical equipment connected to that power distribution line. This sudden, extreme charge delivered into equipment can result in the electronics of that equipment short circuiting, destroying it (Carpenter, 1990).

Buried Power Lines.

Underground transmission lines are not immune to damage from an electrostatic pulse. After a strike contacts the ground, the electricity resolved will travel along or just below the earth's surface from where the charge is induced, to where the strike terminates. As buried transmission lines are much better conductors than soil, voltage will travel through the buried line based on the proximity to the strike terminus, which can also ground itself by traveling through electronics (Carpenter, 1990).

Electromagnetic Pulse

Beyond electrostatic effects, the flow of current through the lightning stroke channel induces a transient magnetic charge as well, which results in a rising neutralization current. This current pulse rises on average by 100 kiloamps per second, though they have been measured to be as high as 510 kiloamps per second. This current induces a transient magnetic field and the induced voltage from this magnetic pulse can be significant enough to short circuit electronics and can induce further electrical currents within powerlines. An electromagnetic pulse (EMP) from a lightning strike occurs on a wide spectrum of frequencies, some of which are low enough to penetrate shielding around both raised and buried power lines. Although this pulse is not strong enough to physically damage the electrical network, it is powerful enough to damage systems which operate at much lower voltages, such as data circuits and consumer electronics (Guo et al., 2024).

Methods of Preventing Damage

Strike Prevention.

Preventing lightning strikes revolves around the reduction of electric potential between the thunderhead and the earth. Preventing a strike from occurring requires constructing structures which facilitate a slow and continuous resolution of charge. Much of a storm's energy is bled off naturally through ambient objects such as trees, grass, fences or other objects, which are ionized by the electrostatic field created by a storm cell. A Dissipation Array System (DAS) enhances this process by drawing away charge from a protected site, which leaves that site with a lower potential than its surroundings, preventing a strike. This drawn away charge also ionizes the air around the DAS, which creates a Faraday shield around the site, further reducing the probability of a strike (Bazelyan and Drabkin, 2003).

A DAS also requires a Ground Charge Collector (GCC) to function. This system consists of buried wires or copper tubing approximately 25 centimeters below the earth's surface and then radiating grounding rods 10 meters away from the DAS site. The GCC can be augmented by connecting the DAS to surrounding buildings, or even by sowing the ground with mineral salts which lowers the resistance between the GCC's radiating wires and the earth. The installation of these system requires trained specialists, enough space for the assembly, and excavation of the ground around the main DAS to install the GCC (Carpenter, 1990).

Surge Protection.

Surge Protection Devices (SPD), also known as lightning arresters, are designed to limit the damage caused by transient voltages. During normal operation, an SPD is in a high impedance state which does not interfere with the flow of electricity. When a voltage spike occurs on the connected circuit, the SPD changes to a low-impedance state, which redirects the current and sends it into the ground or back to its source. After the surge has passed, the SPD returns to its normal state of high impedance (Emerson Process Management, 2002; Ibraham, 2007).

Different ratings exist for SPDs, which are classified on the severity of the transient voltage that the SPD can safely dissipate. In regions where lightning strikes are more frequent, a higher rating is required to adequately protect the circuit. Naturally, more robust SPDs are more expensive to produce and install, and in all cases require trained personnel to integrate into the system (NVent, 2009).

Lightning Rods.

Lightning rods are one of the oldest methods of preventing damage from a lightning strike. Lightning rods provide a conductive target which is designed to be struck and send the current into the ground. Invented by Benjamin Franklin in 1750 and first used for protection in 1752 in France, lightning rods are attached to buildings or the tallest object near the structure is designed to protect. This creates a cone of protection which radiates from the top of the lightning rod and protects the structure within it. High voltage wires are run from the rod to the ground, where the charge is dissipated. While lightning rods provide adequate protection against strikes in most cases, it should be noted that they do not provide absolute protection; The Empire State Building has a lightning rod installed at the top of its structure, yet strikes have been recorded 50 ft. below the top of the skyscraper (Krider, 2006).

All methods of preventing damage from lightning strikes – whether through dissipation, surge protection, or lightning rods - require specialized equipment and specialist technicians in order to function properly. As a result, regions which experience a high frequency of lightning have significantly higher maintenance costs associated with protecting their IT infrastructure. Higher maintenance costs suppress investment in IT technologies, particularly in rural or low-income areas where profit expectations are limited. Therefore, all else being equal, areas with more lightning have less internet penetration.

APPENDIX B

RURAL AND URBAN SUBSET REGRESSION RESULTS

Table B1 and B2 display the results of a robustness check, where the rural counties and the urban counties were separated into two distinct datasets and the same main regression run with them. The results are consistent with the findings of the main regression, with rural counties experiencing a few of migrants into the county at an increase of 1.4% annually between the years 2011 – 2019. Likewise, urban counties did not experience a significant shift in net migration over the time-period, but they did experience a significant drop in both the migration outflow rate and the migration inflow rate, with the estimates of both being 6.9% and 7.3% annually in the same time-period. These results are consistent with the conclusions of this thesis, that overall mobility decreases in urban counties with a higher level of internet penetration, and that individuals move into rural counties with greater access to broadband.

Table B1: Two-Stage Least Squares Regression Results, Internet Connectivity on Migration Rate, Rural Subset.

	Net Migration Rate	Migration Outflow Rate	Migration Inflow Rate
Connections at 200kbps	1.397**	-0.253	1.144
	(0.707)	(0.770)	(0.865)
Unemployment Rate	0.145***	0.070*	0.216***
	(0.041)	(0.043)	(0.046)
Relative Wage	0.003	2.700***	2.703**
	(0.666)	(1.016)	(1.305)
High School Education	0.003	-0.046	-0.043
	(0.024)	(0.030)	(0.033)
Some College Education	-0.008	-0.027	-0.035
	(0.023)	(0.026)	(0.029)
Bachelor's Degree or Better	-0.011	0.043	0.032
	(0.039)	(0.048)	(0.057)
Proportion of Industries Providing Services	0.710*	0.691	1.400**
	(0.408)	(0.455)	(0.639)
Observations	621	621	621
F-Test	20.67	20.67	20.67
Wu-Hausman	4.325	0.00	2.63

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust and in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01

Table B2: Two-Stage Least Squares Regression Results, Internet Connectivity on Migration Rate, Urban Subset.

	Net Migration Rate	Migration Outflow Rate	Migration Inflow Rate
Connections at 200kbps	-0.447	-6.879***	-7.326***
	(0.432)	(1.631)	(1.751)
Unemployment Rate	0.048***	-0.065*	-0.017
	(0.010)	(0.036)	(0.039)
Relative Wage	-0.756***	1.783**	1.027
	(0.170)	(0.742)	(0.765)
High School Education	0.031***	0.016	0.047
	(0.011)	(0.040)	(0.043)
Some College Education	0.057***	0.238***	0.295***
	(0.016)	(0.061)	(0.063)
Bachelor's Degree or Better	0.062**	0.429***	0.492***
	(0.031)	(0.111)	(0.120)
Proportion of Industries Providing Services	0.460***	-0.827*	-0.367
	(0.119)	(0.446)	(0.462)
Observations	2,426	2,426	2,426
F-Stat	27.49	27.49	27.49
Wu-Hausman	1.732	86.68	79.59

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust and in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

APPENDIX C

COMPLETE MIGRATION RATE REGRESSION RESULTS

This appendix shows the full results of the migration rate regressions, with the coefficients and standard errors of all control variables included.

Table C1: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Migration Rates, Complete Results.

	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	0.097 (0.344)	-4.838*** (1.069)	-4.935*** (1.034)
Unemployment Rate	-0.064*** (0.010)	0.005 (0.027)	-0.059** (0.026)
Relative Wage	-0.728*** (0.165)	1.041* (0.548)	1.769*** (0.540)
High School Education	0.022** (0.009)	0.011 (0.026)	-0.011 (0.025)
Some College Education	0.037*** (0.013)	0.192*** (0.039)	0.155*** (0.040)
Bachelor’s Degree or Better	0.030 (0.024)	0.331*** (0.072)	0.301*** (0.069)
Proportion of Industries Providing Services	0.539*** (0.105)	0.103 (0.334)	-0.435 (0.333)
Rural	-0.018 (0.096)	-1.474*** (0.305)	-1.456*** (0.291)
Observations	3,047	3,047	3,047
F-Stat	42.233	42.23	42.23
Wu-Hausman	0.032	53.38	67.53

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust and in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01

Table C2: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Migration Rates, Including Interaction Term, Complete Results.

Regression with Connectivity / Rural Interaction Term			
	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at	-0.180	-5.498***	-5.318***
200kbps	(0.359)	(1.155)	(1.234)
Connections at	0.693**	1.648**	0.955
200kbps * Rural	(0.282)	(0.765)	(0.756)
Unemployment Rate	0.074***	0.029	-0.045
	(0.012)	(0.030)	(0.030)
Relative Wage	-0.611***	1.319***	1.930***
	(0.184)	(0.586)	(0.568)
High School	0.025***	0.017	-0.007
Education	(0.009)	(0.029)	(0.027)
Some College	0.038***	0.157***	0.195***
Education	(0.013)	(0.041)	(0.041)
Bachelor’s Degree or	0.044*	0.363***	0.320***
Better	(0.024)	(0.080)	(0.075)
Proportion of	0.533***	-0.444	0.089
Industries Providing Services	(0.107)	(0.342)	(0.349)
Rural	-2.065**	-6.343***	-4.279*
	(0.856)	(2.311)	(2.274)
Observations	3,047	3,047	3,047
F-Stat (Connections)	21.95	21.95	21.95
F-Stat (Connections * Rural)	160.81	160.81	160.81
Wu-Hausman	4.194	30.44	34.26

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust and in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01

Table C3: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Migration Rate, Lightning Prone States, Complete Results.

	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	-1.216*** (0.437)	-6.799*** (1.340)	-5.582*** (1.102)
Unemployment Rate	0.035** (0.014)	-0.048 (0.040)	-0.083** (0.034)
Relative Wage	-0.481** (0.206)	1.490** (0.719)	1.971*** (0.611)
High School Education	0.046*** (0.013)	0.043 (0.037)	-0.003 (0.030)
Some College Education	0.064*** (0.015)	0.220*** (0.046)	0.156*** (0.041)
Bachelor's Degree or Better	0.113*** (0.031)	0.449*** (0.088)	0.336*** (0.072)
Proportion of Industries Providing Services	0.593*** (0.138)	0.476 (0.454)	-0.117 (0.386)
Rural	-0.317*** (0.109)	-1.807*** (0.374)	-1.490*** (0.301)
Observations	2,621	2,621	2,621
F-Stat	43.26	43.26	43.26
Wu-Hausman	14.21	109.30	85.79

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust and in parentheses. *$p < 0.1$, **$p < 0.05$, ***$p < 0.01$

Table C4: Two-Stage Least Squares Regression Results, Internet Connectivity in 2010 on Migration Rate, Lightning Prone States, Including Interaction Term, Complete Results.

	Net Migration Rate	Migration Inflow Rate	Migration Outflow Rate
Connections at 200kbps	-1.604***	-7.651***	-6.047***
	(0.544)	(1.745)	(1.382)
Connections at 200kbps * Rural	0.673**	1.480	0.807
	(0.310)	(1.031)	(0.822)
Unemployment Rate	0.044***	-0.029	-0.029
	(0.015)	(0.042)	(0.042)
Relative Wage	-0.334	1.813**	2.147***
	(0.246)	(0.818)	(0.677)
High School Education	0.053***	0.058	0.005
	(0.015)	(0.044)	(0.035)
Some College Education	0.069***	0.231***	0.162***
	(0.016)	(0.053)	(0.045)
Bachelor's Degree or Better	0.134***	0.496***	0.362***
	(0.036)	(0.110)	(0.087)
Proportion of Industries Providing Services	0.598***	0.487	-0.111
	(0.146)	(0.483)	(0.401)
Rural	-2.314**	-6.199*	-3.884
	(0.969)	(3.171)	(2.521)
Observations	2,621	2,621	2,621
F-Stat (Connections)	24.08	24.08	24.08
F-Stat (Connections * Rural)	208.65	208.65	208.65
Wu-Hausman	8.024	51.98	40.08

Note: Regression of internet connectivity on average annual net, outflow and inflow migration rates over the period between 2011 to 2019, using average lightning density over the years 2000 to 2010 as an instrument. Standard errors are clustered heteroskedasticity robust and in parentheses. *p < 0.1, **p < 0.05, ***p < 0.01