THREE ESSAYS ON REGIONAL AND DEVELOPMENT ECONOMICS

By

LUYI HAN

Bachelor of Arts in Economics
University of Minnesota - Duluth
Duluth, MN
2016

Master of Science in Economics
Oklahoma State University
Stillwater, OK
2019

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
DOCTOR OF PHILOSOPHY
May, 2021

THREE ESSAYS ON REGIONAL AND
DEVELOPMENT ECONOMICS

Dissertation Approved:

Mehtabul Azam

Dissertation Adviser

Harounan Kazianga

Dan Rickman

Brian Whitacre

Name: LUYI HAN

Date of Degree: MAY 2021

Title of Study: THREE ESSAYS ON REGIONAL AND DEVELOPMENT ECONOMICS

Major Field: ECONOMICS

Abstract:

My dissertation comprises three chapters. The first chapter examines the impacts of the U.S. shale boom on local patents. The second chapter assesses how more competitive political competitions in Sub-Saharan African countries and positive birth-year rainfall shocks affect child mortality rates. The third chapter explores the effects of access and adoption of broadband on self-employment and work-from-home.

The first chapter examines the impacts of the U.S. shale boom on local patenting at a commuting zone level. I expect that the shale boom will negatively affect patents because shale development may crowd out labor and capital investments in other non-energy industries. My findings show that a one standard deviation increase in non-vertical drilling well density decreases patent intensity by 3.6% of the mean. Areas with higher drilling densities have lower levels of patented innovation compared to their counterfactuals. This study contributes to the existing literature related to the "natural resource curse." I provide new evidence based on local patenting, which is an important indicator for regional innovation and long-term economic growth.

In the second chapter, I empirically test three hypotheses that affect child mortality based on the rural sample in Sub-Saharan African countries. In the first hypothesis, I assess the effects of more competitive presidential elections on child mortality. In the second hypothesis, I investigate the impacts of birth year rainfall shocks on child mortality. In the third hypothesis, I argue the effects of political competition can be heterogeneous due to different environment conditions. So I interact the presidential election variable with the rainfall variable to examine the heterogeneous effects when there are good rainfall shocks during a more competitive presidential election period. The results show that both competitive elections and positive rainfall shocks reduce child mortality. Their interaction indicates positive rainfall shocks may be less effective to reduce child mortality during a more competitive election time period.

In the third chapter, using the American Community Survey and the Federal Communications Commission data, I examine how broadband affects self-employment and work-from-home for married women. Based on different sources of internet variables, I investigate the impacts of internet from both the adoption and access to broadband. I find that adoption and access to high-speed broadband have significantly positive impacts on self-employment and work-from-home. This study contributes to the existing literature that examines how Information and Communications Technology affects the labor market.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

THE IMPACTS OF THE SHALE BOOM ON LOCAL PATENTING: A NATIONAL STUDY

1. Introduction

Innovation is a vital factor for long-term economic growth (Verspagen, 2005; Gorodnichenko and Roland, 2011). This effect is not only aggregated at the country level but can also be localized at the regional level. For example, using spatial data in Mexico, Torres-Preciado et al. (2014) establish that technological innovation has a positive effect on Mexico's regional economic growth. Of the several innovation indicators, the use of granted patents has been widely supported in previous literature (Acs et al., 2002; Hall et al., 1984; Jaffe, 1986). Determinants of patenting are various, and they include the level of research and development (R&D) expenditure, the stock of human capital inputs, sector activities, and the size of firms (Simonen and McCann, 2008). In most cases, these factors are endogenously determined. For instance, firms that hire additional high-skilled workers are more likely to have more patents granted, but it is hard to precisely estimate the effects of additional high-skilled workers on patents, since the hiring process is correlated with many unobserved factors (e.g., firm atmosphere), which are endogenous. Although exogenous shocks are rare, they provide opportunities to make a more reliable causal inference. The U.S. shale boom starting from the early 2000s has created relatively exogenous shocks that have been widely used as quasi-experimental designs (Brown, 2018). This study examines the impacts of the U.S. shale boom on local patenting at a commuting zone (CZ) level.

The unprecedented shale boom in the U.S. since the early 2000s has been well assessed in previous works. With the advancement of fracking technology, firms can extract oil and gas trapped in deep shale rocks more efficiently, while minimizing environmental impacts compared to conventional drilling techniques. From 2004, the beginning of the shale boom, to 2014, wherein oil prices dropped significantly, both oil and gas production increased more than four times in the U.S. (Feyrer et al., 2017). On the one hand, previous research works have demonstrated the positive effects of the shale boom on population, employment, wages, and local economic growth.[1] On the other hand, several side effects have been accompanied the shale boom. For example, environmental issues, such as water and noise pollution, increase in drilling-intensive areas, while some studies have established that crime rates increase in places with more drilling activities (James and Smith, 2017; Komarek, 2018; Rozell and Reaven, 2012).

I expect that the shale boom will negatively affect local patenting. This expectation is related to the "natural resource curse," which implies that countries or areas with rich natural resources have lower economic growth (Van der Ploeg, 2011). Specifically, I discuss several possible mechanisms whereby shale development may negatively affect local patenting. First, the shale boom decreased local college attainment, which could result in a lower level of patenting (Black et al., 2005; Kumar, 2017; Rickman et al., 2017). Second, the shale boom reduced the number of entrepreneurs and business start-ups, thereby lowering patent applications from these potential newly established businesses (Partridge et al., 2019). Third, the high pay in the oil and gas industry altered the composition of workers across different industries, possibly decreasing the share of innovative workers in shale-rich areas. Fourth, a substantial amount of capital was invested in the oil and

[1] For examples, Weber (2012) finds a large increase in gas production causes modest increases in employment, wages, and median household income in Colorado, Texas, and Wyoming. Paredes et al. (2015) find robustly significant employment effect but less significant income effect in Marcellus region. Munasib and Rickman (2015) study net economic impacts of oil and gas production from shale formations in Arkansas, North Dakota, and Pennsylvania, and conclude that large and statistically significant positive effects exist in oil and gas counties across wide range of regional labor market measures. Maniloff and Mastromonaco (2017) find over 500,000 local jobs are attributable to the shale boom in the U.S.

gas industry, thus crowding out capital investment in other more innovative industries (Gilje et al., 2020; Popp et al., 2020). Finally, the shale boom was detrimental to local amenities, including air quality. Therefore, shale-rich areas may become less attractive to high-skilled workers or decrease workers' productivity, consequently lowering patented innovation (Zhang and Chung, 2020).

My findings show that a one standard deviation increase in non-vertical drilling well density decreases patent intensity by 3.6% of the mean. This suggests that areas with more drilling activities have lower levels of patented innovation compared to their counterfactuals. I also empirically assess the mechanisms using evidence from the labor market, and I find that the share of creative classes decreases in shale-rich areas.

This paper contributes to the broad spectrum of literature related to the "natural resource curse." To the best of my knowledge, this is the first paper that examines the effects of the U.S. shale boom on local patenting, which is an important indicator of regional innovation and long-term economic growth. Specifically, this paper answers an important question: "Did the shale boom enlarge the innovation gap between shale-rich areas and shale-infeasible areas in the U.S.?" The empirical evidence showed that before the shale boom, shale-rich areas had less patenting compared to shale-infeasible areas, and the shale boom widened the gap. This implies that the shale-rich areas embraced the boon of the shale boom in the short run, but in the long term, it may have hampered innovation-induced economic growth.

The remainder of this paper is organized as follows: The basic conceptual framework as well as relevant literature are presented in Section 2, while the data sources and main analytical sample are explained in Section 3. The empirical methods and models to address the research question are discussed in Section 4, while the empirical results with explanations are presented in Section 5. Lastly, the summary and concluding remarks are provided in Section 6.

2. Conceptual Framework

2.1. Background of Fracking Technology and the Shale Boom

The recent shale boom is associated with technological advancements in fracking (hydraulic fracturing). For a long time, scientists have known that a substantial amount of oil and gas is trapped within deep shale rocks in shale formations. With the expansion and advancement of fracking techniques, trapped oil and gas in deep shale rocks can be safely and efficiently extracted and produced. Drilling wells constructed using advanced fracking technologies are referred to as non-vertical drilling wells. Before the advancement of fracking technology, most drilling wells built were vertical. Understanding the differences between non-vertical and vertical drilling wells is crucial for the empirical designs presented herein.[2] The initial steps involved in constructing non-vertical and vertical drilling wells are similar; a hole is drilled straight down through the fresh water aquifer to a targeted area. However, subsequently the construction processes for vertical and non-vertical wells differ. For vertical drilling wells, only a simple zone of interest, which is limited to a certain spherical area, is available to do a single stage fracking job to produce oil and gas. For non-vertical drilling wells, the pipeline is horizontally or directionally extended, and multiple stages of fracking jobs are available to produce oil and gas. The unique design of non-vertical wells is more efficient and may minimize the environmental impacts compared to vertical drilling wells. The oil and gas production of one non-vertical drilling well can be compared to that of several vertical drilling wells. However, although non-vertical wells are more lucrative in the long run, they are much more expensive to construct than conventional vertical wells (Bartik et al., 2019). Therefore, building non-vertical wells involves more intense labor and/or capital investments. As noted in Fitzgerald (2012), the average cost of building non-vertical drilling wells is approximately three times higher than that of vertical drilling wells. For instance, in the Woodford Shale, the shift from vertical to non-vertical wells increased the average well

[2] Non-vertical drilling wells are also known as unconventional drilling wells. Vertical drilling wells are also called conventional drilling wells.

cost from US$2 million to US$5–6 million (Bartik et al., 2019; Fitzgerald, 2012).

2.2. Shale Boom and Resource Curses

Summarily, the resource curse refers to the fact that countries with rich natural resources may have lower economic growth (Van der Ploeg, 2011; Van der Ploeg and Poelhekke, 2010; Brunnschweiler and Bulte, 2008). For example, "oil revenues per capita in Nigeria increased from US$33 in 1966 to US$325 in 2000, but income per capita has stagnated at approximately US$1000 in purchasing power parity terms since its independence in 1960 (Van der Ploeg, 2011)". The resource curse also occurs at a disaggregated level within a country. For example, crime rates increase in shale-rich areas due to factors, such as higher income inequality (James and Smith, 2017; Komarek, 2018; Shakya, 2019). Using data from Pennsylvania, previous works have established that nearby shale development has a high negative impact on housing values (Muehlenbachs et al., 2015). The likelihood of water contamination increases from natural gas extraction in shale-rich areas (Rozell and Reaven, 2012). Herein, I provide new evidence on the resource curse from another perspective—the shale boom and local patenting. There are two papers that can elucidate the hypothesis of this paper, and using cross-country data, the authors find that the shale boom decreases clean energy patenting (Acemoglu et al., 2019; Popp et al., 2020). My paper provides evidence at the regional level in the U.S. and focuses on overall patents. My hypothesis is that for regions with more shale development, labor and capital are more concentrated in the oil sector, thereby resulting in the lower investment of labor and capital in other sectors. The crowding-out effects of labor and capital in other sectors are reflected in the lower level of patenting in shale-rich areas. I provide descriptive evidence in Section 2.4.

2.3. Innovation and Patenting

The main determinants of patenting include the level of R&D expenditure, the stock of human capital inputs, sector activities, and the size of the firms (Simonen and McCann, 2008; Pfister et al., 2021; Agrawal et al., 2010; Xiao et al., 2021). All of these determinants may be affected during the shale boom. Carlino and Kerr (2014) report that although invention is a vital indicator of innovation, invention is not equal to innovation. The first stage of innovation is invention, and patents have been widely used to measure invention. Subsequent investment is needed to nourish invention into innovation. Therefore, using patents to measure innovation is a rather crude method and it has a certain degree of caveats. For instance, some important innovations may never apply for patents, and the quality of patents can significantly differ. In this work, I use patents to measure local innovation while acknowledging that limitations do exist.

2.4. Descriptive Evidence

Figure 1 shows the average new non-vertical and vertical drilling well densities at a CZ level from 2000–2015. From 2005, the non-vertical well density increased rapidly until 2008 and reduced in 2009 temporarily, whereafter it then increased again until 2014. However, the vertical drilling well density increased from 2000 until 2008 and thereafter reduced considerably. Figure 2 presents the mean patents per 100K population (patent intensity) between 2000 and 2015. Before 2009, patent intensity fluctuated between 12 and 14, and it skyrocketed thereafter from 12 to 19.5 in 2014. The potential crowding-out effects of the shale boom on patenting are depicted in Figure 3. I define the control group as CZs in 48 lower U.S. states without any non-vertical wells from 2000–2015. Respectively, other CZs with at least one new non-vertical drilling well are in the treatment group. It is evident that the CZs in the treatment group have a lower level of patent intensity across all the years from 2000–2015. Notably, 2005 or thereabouts is perceived as the onset of the shale boom in the U.S. Before 2005, the mean difference in patents per 100K population (patent intensity) between the control and treatment groups was less than two. But this difference has amplified over the years after 2005. In Figure 4, I plot the differences in the patent intensity between the treatment and control groups shown in Figure 3. It is more intuitive to observe that the patent intensity between the treatment and control groups increases over the years. Figure 3 and 4 may indicate that shale-rich areas have always had a lower level of patent intensity, and the gap widened after the shale boom.

2.5. Potential Mechanisms

In this section, I discuss potential mechanisms through which shale development may crowd out local patenting. The first mechanism entails that the shale boom increases the opportunity costs of obtaining more education. Hence, local average levels of education are reduced, which indicates that there are fewer higher educated people to innovate. This mechanism is widely supported by previous literature (Black et al., 2005; Kumar, 2017; Rickman et al., 2017; Han and Winters, 2020). For instance, a previous study reported that the Appalachian coal boom in the 1970s reduced high school enrollment significantly (Black et al., 2005). Furthermore, Kumar (2017) discovered that the oil boom in the 1970s lowered the college premium, leading to a significant drop in college enrollment and completion. One work examines the effects of the shale boom on high school and college attainment in Montana, North Dakota, and West Virginia, and the authors find that the shale boom significantly reduces high school and college attainment in all three states (Rickman et al., 2017). For the second mechanism, if the shale boom decreases the number of entrepreneurs or business start-ups, patenting activities may be simultaneously reduced. Partridge et al. (2019) find that the formation of new firms as well as sales initially decrease after exposure to the shale boom.

The third mechanism may be attributed to the high pay in the oil and gas industry that has altered the composition of workers across different industries, thereby possibly decreasing the share of innovative workers in shale-rich areas. Using the American Community Survey (ACS), I find that the shale boom has decreased the share of creative class workers. I provide more details in Section 5.3. The fourth mechanism involves the capital market. The advancement of fracking technology during the shale boom has resulted in the construction of numerous non-vertical drilling wells. Although non-vertical drilling wells are more efficient than the vertical ones, they are much more expensive to construct (Bartik et al., 2019). This may suggest that in shale-rich areas, a substantial amount of capital is invested in the oil and gas industry, thereby crowding out capital investment in other more innovative industries (Gilje et al., 2020; Popp et al., 2020). Moreover, the energy industry is a less innovative industry. Over the past decade, the share of the average R&D expenditure over net sales revenue across all industries has been between 2.5 and 3.5%; for the manufacturing industry, the share ranges from 3.1 to 3.9%, while for high-tech industries, such as the computer industry, it has been approximately 10%. However, for mining and extraction industries, which include the energy sector, the share has been less than one percent until 2015 (Popp et al., 2020). The last mechanism derives from the local amenity perspective. Zhang and Chung (2020) find that poor air quality has been detrimental to regional innovation based on city-level evidence in China. As has been documented in previous works, the shale boom has had negative effects on local housing values and increased local water contamination (Muehlenbachs et al., 2015; Rozell and Reaven, 2012). The deterioration of local amenities may be less attractive to high-skilled workers and workers may become less productive, and this may indicate lower patented innovation.

3. Data

I combine different types of data for the empirical analyses. Herein, the main geographical unit is at a CZ level. Notably, CZs are clusters of U.S. counties that are characterized by strong within-cluster and weak between-cluster commuting ties (Autor et al., 2019; Autor and Dorn, 2013). Using CZ-level data can mitigate spillover effects across counties (Tolbert and Sizer, 1996). The information on new drilling wells is obtained from Enverus [3]. The data reports detail information on new drilling wells, including the coordinates of each drilling well. I use the coordinates to identify the CZ that each drilling well belongs to. Enverus also provides information on the types of drilling wells; hence, it

[3] https://www.enverus.com/

is possible to distinguish between vertical and non-vertical drilling wells. As discussed in the conceptual framework section, due to the technological advancement of fracking techniques during the shale boom, an increasing number of non-vertical drilling wells have been built after 2004 (Figure 1). Additionally, non-vertical wells are much more expensive to build (Bartik et al., 2019). I test different treatment variables based on the non-vertical or vertical drilling well density to observe where the significant negative effects on local patents originate.

Patent information is obtained from the U.S. Patent and Trademark Office. I extract the total utility patents for inventions granted from 2000 to 2015 at the county level, and thereafter aggregate them to the CZ level. According to the data description, "the geographic distribution of the patents at the county level is based on the residence, at grant, of the inventor whose name appears first on the printed patent (i.e., the first-named inventor)."

The main data for control variables are obtained from the ACS via the Integrated Public Use Microdata Series (Ruggles et al., 2020). The ACS is an annual survey of one percent of the U.S. population wherein demographic, education, work, income, and other information is collected. To comply with the sample period of the patents, I use the 2000–2015 ACS data.[4] The smallest identifiable geographic area in the ACS is the Public Use Microdata Area (PUMA). Based on PUMA to CZ crosswalk files, the variables are converted to the CZ level.

Another source of control variables is the Quarterly Census of Employment and Wages (QCEW) data. The QCEW data provide industry-specific employment and wages at the county level. I use these data to control employment outside the mining industry. Again, the county level variables are aggregated to the CZ level.

Combining all the above-mentioned data sources, the main analytical sample extends from 2000 to 2015 and contains 705 CZs covering the entire lower 48 U.S. states (excluding the District of Columbia).

[4] I use 2000 5% census data and 2001-2015 1% ACS data.

4. Methods and Models

4.1. Baseline Two-way Fixed Effects Model

The paper examines how new drilling activities affect the patent intensity at a CZ level. In the baseline specification, I estimate linear regressions in the form provided below for a CZ c in year t:

$$Patent_{ct} = \beta_0 + \beta_1 WellDensity_{ct} + \gamma X_{ct} + \pi_c + \varphi_{rt} + \delta_c T_{ct} + \mu_{ct} \tag{1}$$

As noted in the data section, I choose CZs as the analytical units because they are clusters of counties with strong within-cluster and weak across-cluster commuting ties (Autor et al., 2019; Autor and Dorn, 2013). Using CZs can mitigate spillover effects across counties (Tolbert and Sizer, 1996). Because CZs vary according to the population and land area sizes, directly comparing the total patents across CZs is undesirable. It is very likely that larger CZs have higher patenting activities simply due to their larger populations. Thus, the dependent variable is defined as patent intensity, which is measured as the number of total patents per 100K population [5]. The main explanatory variable of interest—$WellDensity_{ct}$—is defined as the new non-vertical well density. I calculate the variable as the number of new non-vertical wells divided by the land area in 1000 square miles in a CZ c. X_{ct} is a set of time-variant control variables, including demographic, education, and non-mining industry employment variables. Details of the construction of X_{ct} are reported in Appendix. I also include CZ fixed effects π_c and region-by-year fixed effects φ_{rt} to absorb time-invariant heterogeneities across CZs and shocks to different regions in each sample year.[6] In addition, I include $\delta_c T_{ct}$ to capture a CZ-specific linear time trend.

[5]Population refers all population not only working age population. To account for endogenous population growth, I use year 2000 population to generate the patent intensity variable for all the years. For example, year 2001 patent intensity in CZ code 100 is year 2001 patents divided by 100K population in year 2000.

[6]The four US Census regions are the Northeast, Midwest, South, and West.

4.2. Interactive Fixed Effects Model

In Section 4.1, I control for a linear time trend in addition to the two-way fixed effects baseline model. I also include a test for a quadratic time trend (see footnote 9). However, the actual time trend may neither be linear nor quadratic; it can be in other forms, which are not fully observable. As depicted in Figure 2, there is no apparent trend for patent intensity prior to 2008, but there is a sizeable upward trend after 2008. The overall trend for full sample years from 2000–2015 may be perceived as a quadratic trend. This elicits one of the challenges of the baseline model — time-varying unobserved heterogeneities.

To address this, I employ the interactive fixed effects (IFE) approach in addition to the two-way fixed effects model to enable a flexible and explicit modeling of the non-linear time-varying heterogeneities, as shown in Equation 2 below (Bai, 2009; Czarnowske and Stammann, 2020; Kejriwal et al., 2018; Totty, 2017):

$$Patent_{ct} = \beta_0 + \beta_1 WellDensity_{ct} + \gamma X_{ct} + \pi_c + \varphi_{rt} + \lambda_{sc} F_{st} + \mu_{ct} \tag{2}$$

Compared to Equation 1, the linear time trend term $\delta_c T_{ct}$ is dropped in Equation 2, while it includes an additional term $\lambda_{sc} F_{st}$ to model the IFE. The IFE model assumes that the dependent variable patent intensity can be modeled as a function of s unobserved linear factors, F_{st}, which are common across all CZs c in year t. Furthermore, λ_{sc} represents the CZ-specific factor loadings, which are constant over year t. These factors can be considered as omitted variables, including $\lambda_{sc} F_{st}$, which allow the unobserved common factors to be correlated with the regressors (Bai, 2009; Totty, 2017). To understand this, assume that there is a national policy change which encourages patenting by giving monetary rewards. The effects of the policy may be heterogeneous across CZs depending on CZ-specific characteristics. For example, CZs with higher shares of older populations may be less motivated by this national policy because older people may be richer; thus, they are less attracted by the monetary rewards.

Finally, as noted in the second paragraph on page 1716 in Totty (2017), "the IFE model will produce similar estimates as the ordinary least squares model if the two-way fixed effects specification is correct." I compare different estimates in the subsequent sections, and I implement the IFE model using the Stata "regife" command developed by Gomez (2017).

4.3. Border Contiguity

I utilize the two-way fixed effects model and the IFE model described in the previous sections to address the unobserved time-varying heterogeneity. However, using the full U.S. sample to compare shale-rich CZs with shale-infeasible CZs may not be ideal as it may not be a comparable control group. To facilitate a more appropriate comparison, I choose control CZs based on spatial contiguity (Dube et al., 2010). Rather than comparing the patent intensity of all the shale-rich CZs to CZs without any non-vertical drilling wells, I compare the patent intensity of all the shale-rich CZs to their contiguous CZs without any non-vertical drilling wells.

4.4. First-Difference Model

Although using CZs as an analytical unit can help to control spillover effects across counties, there may still exist year-variant spillover effects across CZs that possibly contaminate the main specification. Because patenting is highly human capital intensive, proximity to large metropolitan areas can induce more patenting. This effect may change over time. To account and quantify the year-variant spillovers, I control for urban hierarchy variables, which are incrementally distant to the nearest metropolitan statistical area (MSA) with a population of at least 250K, 500K, 1500K, and 2000K (Partridge et al., 2008). However, the baseline two-way fixed effects model in Equation 1 absorbs all the time-invariant variables, including incremental distance controls. Alternatively, I estimate a first-difference model in Equation 3, as follows[7]:

$$\Delta Patent_{ct} = \beta_0 + \beta_1 \Delta WellDensity_{ct} + \gamma \Delta X_{ct} + \xi IncDis_c + \theta Patent_c^{90-99} + \varphi_{rt} + \mu_{ct} \quad (3)$$

where $\Delta Patent_{ct}$ = $Patent_{ct}$ - $Patent_{ct-n}$;

$\Delta WellDensity_{ct}$ = $WellDensity_{ct}$ - $WellDensity_{ct-n}$;

$\Delta X_{ct} = X_{ct}$ - X_{ct-n}; subscript c refers to a CZ, t is a year with n = 1, 2, 3, 4, and 5, which indicates the period of time differences used for each model; $IncDis_c$ represents the incremental distance to the nearest MSA with a population of at least 250K, 500K, 1500K, and 2000K; while $Patent_c^{90-99}$ refers to the average patent intensity between 1990 and 1999. This term captures the time-variant effect of initial patenting for each CZ. For example, CZs with more patenting between 1990 and 1999 may have high patenting growth rates in subsequent years.

4.5. Instrumental Variables

To alleviate endogeneity issues due to the fact that drilling activities are not randomly assigned to different CZs across the U.S., I use instrumental variables (IVs) based on historical geological formation for shale plays. The IV incorporates time-invariant geological and temporal information for a given CZ (Feyrer et al., 2017). I construct this IV using two steps in Equations 4 and 5 below. I first estimate:

$$Ln(Well_{ct} + 1) = \alpha_c + \lambda_{jt} + \epsilon_{ct} \quad (4)$$

where $Well_{ct}$ represents the number of non-vertical drilling wells for CZ c and year t, α_c represents a dummy of each CZ, and λ_{jt} represents a set of dummy variables for shale play-year combination.[8]

[7] Fix effects and first-difference are similar but each has its own merits. One of the benefits using a first-difference is it allows regressions to include time-invariant control variables (Wooldridge, 2010).

[8] According to the U.S. Energy Information Administration (EIA), "shale play is a set of discovered, undiscovered or possible natural gas accumulations that exhibit similar geological characteristics. Shale plays are located within basins". There are around 20 shale plays in the U.S.

In the second step, I generate predicted wells density as follows:

$$\widehat{WellDensity}_{ct} = (e^{\widehat{\alpha_c}+\widehat{\lambda_{jt}}} - 1)/TotalArea_c \tag{5}$$

Where $TotalArea_c$ is the total land area in 1000 square miles for a CZ c. This IV uses $\widehat{\alpha_c} + \widehat{\lambda_{jt}}$, which contains geological formations in a given CZ and shale play-year information, to predict new non-vertical drilling well densities. These predicted values for new non-vertical drilling well densities are based on the timing of new drilling wells for all the CZs within a particular shale play. As a CZ represents a small part of the shale play's drilling wells' construction, the instrument is exogenous with respect to the idiosyncratic roll out of fracking within individual CZs (Feyrer et al., 2017).

5. Results

5.1. Main Results

Table 1 shows the descriptive statistics for all the variables. The construction of control variables is reported in Appendix. The mean CZ-level patenting per 100K population (patent intensity) over the full sample period is 14.66. The means of non-vertical and vertical wells per 1,000 square miles are 3.31 and 8.92, respectively. Figure 5 shows the distribution maps of the mean non-vertical drilling well densities for full sample years between 2000 and 2015. Drilling wells are concentrated in energy states, including Oklahoma, Texas, and Wyoming. Figure 6 shows the distribution of the mean patent intensity, wherefrom it is indicated that energy states have lower levels of patent intensity.

Table 2 Panel A presents the baseline regression results obtained from Equation 1 along with the IFE described in Section 4.2 for the full sample. In Table 2 Panel B, the sample CZs are restricted based on the border contiguity discussed in Section 4.3. Columns (1) and (2) present the baseline regression results, and column (3) presents the IFE results estimated using Equation 2. In Panel A, column (1) indicates that a one unit increase in non-vertical drilling well density decreases patent intensity by 0.0294, which means that a one standard deviation increase in drilling well density decreases the patent intensity by 3.6% of the mean. In column (2), I include a linear time trend to serve as a control for the linear growth in patent intensity over years, and it reveals that a one unit increase in non-vertical well density decreases the patent intensity by 0.0286, which is similar to column (1).[9] In column (3), the IFE is incorporated to explicitly estimate and control the time trend. It shows that an increase in non-vertical drilling well density reduces patent intensity, though the magnitude (0.0221) is a little less than those in columns (1) and (2). Table 2 Panel A indicates that areas with more drilling activities measured by the non-vertical drilling well density have lower levels of patent intensity compared to other areas without non-vertical drilling wells. As noted in Section 4.3, directly comparing drilling-rich areas with all other areas without any drilling may not be ideal. I narrow down the control group CZs based on the border contiguity in Table 2 Panel B.

Table 2 Panel B replicates Panel A regressions using border contiguity CZs as the estimation sample. Specifically, the sample contains CZs with at least one non-vertical drilling well between 2000 and 2015 and their border CZs. All other CZs are excluded in the sample. In Panel B, column (1) shows that a one unit increase in the non-vertical well density decreases the patent intensity by 0.0306. This is very similar to that in column (1) in Panel A, but with a marginally higher magnitude. Columns (2) and (3) in Panel B are similar to Panel A. Because restricting sample CZs based on border contiguity does not qualitatively change the results, I use the full sample, including all the CZs, in the subsequent tables unless otherwise noted.

In Table 3, I present an estimation of the first-difference regressions in Equation 3. Using first differencing can remove time-invariant factors that may contaminate the estimates, and it serves a similar function as CZ fixed effects. The benefit of using first differencing compared to fixed effects is that the latter provides researchers with the flexibility to include time-invariant control variables. Consequently, I can observe

[9] Instead of including a linear time trend, I test including a quadratic time trend. Results are similar to Panel A, both column (1) and (2).

whether the time-invariable attributes of interest have any effects over time. I include urban hierarchy incremental distance controls and average patent intensity between 1990 and 1999 (Partridge et al., 2008). As I discussed in the previous sections, CZs are clusters of counties with strong within-cluster ties and weak across-cluster ties. Using CZs can mitigate spillover effects across counties; however, some spillovers may still exist from the nearby large MSAs, and the spillovers may change over years. For example, CZs near the largest New York–Newark–Jersey City metropolitan area may receive strong spillovers, which may increase over years as this metropolitan area further grows. By including the average patent intensity between 1990 and 1999, the initial levels of patent intensity can be captured to observe how the effects of the initial levels of patent intensity evolved over the years. I use different time periods for first differencing from one year to five years in columns (1) to (5). Therefore, I can capture whether the crowding-out effects in drilling-rich areas amplify or shrink from the short-term (1-year) to the median-term (5-year). In column (1), one unit change in the non-vertical drilling well density reduces the change in patent intensity by 0.022. These crowding-out effects increase over years and arrive at the peak in the 4-year difference, wherein one unit change in non-vertical drilling well density reduces the change in patent intensity by 0.034. This is approximately 50% larger than the 1-year difference. In addition, from the incremental distances to MSA controls, I observe that CZs have a lower patent intensity when they are located away from MSAs with at least populations of 500K and 1500K. This is consistent with urban hierarchy models (Partridge et al., 2008). Specifically, it indicates that CZs near large MSAs receive positive spillover effects for higher patent intensity growth. Finally, the significant positive coefficients from the average patent intensity between 1990 and 1999 reveals that the higher initial level of patent intensity in the foregoing time period can predict higher patenting in subsequent years.

The results using IVs are presented in Table 4. First stage results are shown in Panel A. Both the first stage results presented in column (1) and (2) have expected positive signs and exceed conventional values to identify the weak instrument. In column (1)

Panel B, one unit increase in non-vertical well density reduces the patent intensity by 0.022. Incorporating a CZ linear trend gives a similar estimate in column (2). One unit increase in the new drilling well density is associated with a 0.025 decrease in patent intensity.

5.2. Robustness Checks

I conduct several robustness checks, as presented in the Tables in the Appendix. Table A1 reports the lagging effects of the non-vertical well density. In Panel A, I test the main treatment variables separately, and it shows that there are significant crowding-out effects of building new non-vertical drilling wells up to a two-year lag. The strongest effects are the contemporaneous ones. In Panel A, I control the contemporaneous and lagging effects separately because the construction of new drilling wells for continuous years may be highly correlated, and controlling them together may trigger perfectly collinear issues. The results reported in Panel B include the simultaneous contemporaneous and lagging effects of new drilling wells. All columns in Panel B provide evidence that the crowding-out effects of new drilling wells on patent intensity are the strongest contemporaneously.

In Table A2, I test the main results restricting the sub-sample. Panel A column (1) replicates the baseline two-way fixed effects in column (1) Panel A of Table 2. In Panel A column (2), I exclude the top one percent CZs with the highest new drilling well density. It shows that a unit increase in the new drilling well density is associated with a 0.039 unit decrease in patent intensity. Similarly, column (1) in Panel B excludes the top one percent CZs with the highest patent intensity. Panel B column (2) excludes the top one percent CZs with the highest new drilling well density and patent intensity. The main results still hold. New York became the first state to ban fracking around 2014; thus, I excluded all CZs in New York in Panel C column (1). Although fracking is not completely banned in California, drilling activities are strictly regulated. Furthermore, California as a state with many technology centers, such as Silicon Valley, may dominate patenting activities. In Panel C column (2), I exclude CZs in California. In Panel D column (1), I exclude all CZs in New York or California. All these exclusions do not alter the main results qualitatively. Finally, from Figure 3, I notice that the difference in the mean patent intensity in the control and treatment CZs diverged in a larger magnitude after 2011. In Panel D column (2), I estimate Equation (1) using the 2000–2010 sample only. The estimates are close to the main baseline results in Panel A column (1) with a marginally smaller size. This means that the crowding-out effects are not only due to the most recent years from 2011 to 2015 but start in the early period of the shale boom.

In Table A3, I estimate Equation (1) for the MSA sample CZs in Panel A and for non-MSA sample CZs in Panel B. The crowding-out effects on patent intensity are stronger in the MSA sample but are still sizeable in the non-MSA sample. Notably, CZs with MSAs have more innovative activities and they may be more affected during the shale boom.

The technological innovation during the shale boom increased the number of new non-vertical drilling wells radically. Meanwhile, conventional vertical wells were still constructed over years especially in the early shale boom years (Figure 1). In Table A4, I estimate Equation (1) using the new vertical drilling well density as the main treatment variable. Vertical drilling well density is calculated as the number of new vertical drilling wells divided by the total land area in 1000 square miles. No significant crowding-out effects are observed in any columns in Table A4. Compared to Table 2, using new non-vertical wells, the insignificant effects in Table A4 suggest that the crowding-out effects ensue from building non-vertical wells rather than vertical ones. This helps to confirm the hypothesis—building non-vertical wells costs more resources, including labor and capital. Therefore, the allocation of resources in other industries is crowding out.

5.3. Testing Mechanisms in the Labor Market

I test potential mechanisms from the labor market based on the share of different classes of workers. Following previous literature, I define the share of workers as the creative, service, and working classes (Florida et al., 2008; Florida, 2014; Gabe et al., 2013). If the share of creative class workers decreases in areas with more drilling activities, it may indicate that there are fewer innovative people to obtain granted patents. Similar to Equation 1, I regress the share of different classes of workers on the new non-vertical drilling well density. The results are presented in Table A5. In column (1), a unit increase in the non-vertical drilling well density is associated with a 0.0029 (0.29%) decrease in the share of creative class workers. Column (3) shows that a one unit increase in the non-vertical drilling well density is associated with a 0.0078 (0.78%) increase in the share of working-class workers. These results indicate that the share of creative class workers declines, whereas the share of the working-class workers grew as more drilling wells are built. The shale boom induced local labor composition changes. However, share of workers depend on both denominator and numerator. The decrease in the share of creative class workers may simply reflect an increase in total number of workers and the number of creative classes may remain unchanged. It is not surprising that share of working class workers increases in drilling areas because working class consists extraction and other oil and gas related occupations.

6. Summary and Concluding Remarks

The recent U.S. shale boom due to the advancement in fracking technology has been attracting considerable attention from researchers. In this study, I examine the effects of the U.S. shale boom on local patenting. I find that a one unit increase in the non-vertical drilling well density is associated with more than a 0.029 unit decrease in patent per 100K population. This implies that a one standard deviation increase in the non-vertical drilling well density decreases patent intensity by 3.6% of the mean. The crowding-out effects do not exist when the vertical drilling well density is used as the treatment variable.

I provide possible mechanisms thorough which the crowding-out effects exist based on evidence from human capital, labor, capital market, and local amenity perspectives. I empirically access the share of creative class workers, and I find that shale development reduces the local share of creative classes, though the evidence is not ascertained. I find that contemporaneous effects are stronger than lagging effects. Some patents can take up to several years to process, whereas some are processed presumably quicker. Drilling wells also take time to be built, although this generally takes less than a year.

I admit the limitation that patents are a rather crude measure of innovation, and the quality of patents can be very different across products. The empirical evidence presented herein suggests that shale-rich areas had lower levels of patent intensity compared to shale-infeasible areas before the shale boom, and this gap has widened over years after the shale boom. This indicates that innovation-related phenomenon was being crowded out as measured by patents. More research is needed to identify the resources that have been crowded out, thereby resulting in lower patent intensity; however, this paper provides evidence on a new channel through which the natural resource curse may be detrimental to local areas. The development of natural resources reduces local patenting, thereby suggesting that areas with natural resources have reduced innovation and, consequently, less long-term economic growth.

As for the policy implications, people should be aware that regional innovation may be impeded in areas where natural resources are intensively developed. Specifically, shale resource-rich states, such as Oklahoma and West Virginia, may need more place-based policies for innovation-related regional economic growth in the long term. Otherwise, the innovation gap between shale-rich and shale-infeasible areas may grow indefinitely and never reach economic convergence, that is, poor areas may never catch up with rich areas (Barro and Sala-i Martin, 1992).

CHAPTER II

RETURNS TO POLITICAL COMPETITION AND LIVING CONDITIONS: PRESIDENTIAL ELECTIONS, RAINFALL, AND CHILD MORTALITY IN SUB-SAHARAN AFRICA

1. Introduction

Child mortality in SSA countries is of research interest in many studies. One reason is that how to reduce the high child mortality in SSA countries is an important question (see Table 5). Another reason is child mortality can be used as a proxy for economic well-being and social welfare, examining factors that affect child mortality can provide information for policy makers to improve economic outcomes. Cutler et al. (2006) investigate the determinants of child mortality, and they find "the long-term reach of early-life factors" is one of the determinants of mortality. Therefore, I propose three hypotheses below to examine how economic shocks affect child mortality in Sub-Saharan African (SSA) countries.

In the first hypothesis of this paper, I assess empirically the effects of political competition on economic outcomes in Sub-Saharan Africa. I exploit the increased number of elections in SSA countries since 1990 and the availability of detailed micro level data to explore the effect of political competition on household welfare. I use presidential election outcomes in SSA countries between 1990 and 2013. Political competition is measured by the margin of the elections. I measure economic performance by child mortality. I then estimate the effects of the intensity of presidential election competitiveness on child mortality and I expect more competitive elections indicate the president is more likely to be competent, hence lowers child mortality. There are some published works in which can elucidate the first hypothesis. For example, Battaglini (2014) shows theoretically that whether increased political competition produce more efficient resource allocation depends on the electoral rules and the underlying fundamentals of the economy. Besley et al. (2010) use the Voting Rights Act of 1965 to associate an increase in political competition in the US Southern states with an improved economic performance measured by income per capita.

Inspired by Deaton (2007) and Bozzoli et al. (2007), in the second hypothesis, I investigate the impacts of birth year rainfall shocks on child mortality in Sub-Saharan Africa countries. I expect that positive rainfall shocks in the birth year improve living conditions of new-born infants. Therefore, their probability of dying early may decrease. Specifically, Deaton (2007) and Bozzoli et al. (2007) investigate environmental determinants of height in developing countries. They find high childhood mortality is associated with taller adults in Africa, which suggests "selection effects" are stronger than "scarring effects". This means better birth year living conditions contribute to lower child mortality rates, which implies larger numbers of weaker children survive and they are more likely to be in lower height.

In the third hypothesis of this paper, I argue the effects of political competition can be heterogeneous due to different environment conditions. In light of the vast rainfall literature on both developing and developed countries, I test the third hypothesis using standard rainfall data covering the entire sample countries and years from the first hypothesis. Specifically, political competition during rich rainfall years may have different impacts compared to political competition during normal rainfall years. For example, if a competitive political election is complementary to higher rainfall, I may anticipate additional decrease in child mortality during rich rainfall years. However, if a competitive political election is not complementary to higher rainfall years, I may observe statistically insignificant additional mortality reduction.

The results show that both competitive elections and positive rainfall shocks reduce child mortality. Their interaction indicates a positive rainfall shock may be less effective to reduce child mortality during a more competitive election time. This paper contributes to the existing literature related to child mortality in SSA countries.

2. Conceptual Framework

2.1. Political Completion and Its Effects

In the earlier works, Stigler (1972) and Wittman (1989) lay out the theoretical background and argue that increased competition would lead to more efficient governments. Although the literature on this topic is growing, but due to endogeneity issues, it is a nascent area that should get increased attention.

Carbone and Pellegata (2017) discuss how multiparty elections in African countries after 1990s affect welfare policies. They test whether voters exercise their power to push policy makers to improve their performance and the delivery of social welfare goods. They examine indicators for welfare investments including health spending, primary school enrolment rates, and child mortality. They find more competitive elections increase social welfare. Similarly, Gerring et al. (2015) suggest that electoral democracy contributes to human development, because politicians are incentivized to enhance social policies. Harding and Stasavage (2014) pay a special attention to the connections between African elections and African education. Based on national and individual African data, their work shows democracies have higher rates of school attendance than non-democracies, since democracies are more likely to abolish school fees. Wullert and Williamson (2016) find hybrid regimes in 47 African countries cause higher infant mortality due to factors such as political instability. Kjelsrud et al. (2020) focus on India and show that a rise in employment inequality induces more post-neonatal infant deaths, only when the political competition is weak. Their work informs that a more efficient government is more likely to issue pro-business policies to mitigate the adverse effects of employment inequality. In Addition, existing literature point out quality of politicians are capable for competent government, consequently economic performance (Ferraz and Finan, 2009; Grofman and Selb, 2009; Selb, 2009).

Two of the close antecedents to the study are Kudamatsu (2012) and Harding (2020). Kudamatsu (2012) exploits the transitions from one-party or military regimes to multi-party systems in the 1990 in Sub-Saharan Africa to test the effects of democratization on child mortality, and the author finds that shifting to multiparty system in SSA countries since 1990 drops child mortality significantly. Another study by Harding (2020) use individual data from 27 African countries. The author finds democratic elections increase access to primary education and reduce child mortality rates only in rural areas, and conclude that competitive elections incentivize governments to provide pro-rural policies. This study is different from their papers in two ways. First, I take advantage of the detailed elections results over 25 years to precisely measure the competitiveness of over 100 presidential elections as "multi-shocks". Unlike Kudamatsu (2012), who focuses on "the one time shock", defined as the shift from one-party to multi-party systems. Second, I interact the political competition with rainfall shocks to test whether a more competitive election may have a stronger or weaker impact on child mortality during the period with more rainfall.

In light of the existing works, I formulate two main mechanisms whereby political competition can improve welfare and hence reduce child mortality in SSA countries. In the first mechanism, a more contested political completion refers to higher democracy and lower corruption (e.g., one-party vs. multiparty system), which could result in social welfare improvements. In the second mechanism, a more contested political competition is accompanied with stronger electoral incentives, which could be reflected as welfare improvements.

2.2. Rainfall Shocks and Its Effects

Exposure to different shocks at the early age may have short-run and long-run effects (Baird et al., 2011; Kazianga, 2012). These shocks can be various, including weather shocks, income shocks, war shocks, etc. Among weather shocks, the rainfall shock has been widely accessed in the existing literature. One of the main reasons could be that the rainfall shock is perceived as an exogenous shock. Therefore, identification of causality is more convincing even without a rich set of control variables. For example, Maccini and Yang (2009) investigate the impact of early-life rainfall shocks on health, schooling, and economic outcomes using Indonesian data. They find higher levels of rainfall at the birth year has positive effects on health conditions, height, and grades of schooling for women but no significant effects for men. The positive effects are identified only at the birth year. One of the mechanisms is higher precipitation increases the agricultural conditions in Indonesia where rice production is one of the main industries.

Shah and Steinberg (2017) study the factors that affect human capital investment in India. They find children are more likely switching out of school into productivity work when exposing to higher rainfall shocks. The opportunity cost of foregone wages for schooling is high when productivity is high. They drop out the school and take a lucrative job instead. BenYishay (2013) investigates the impacts of early rainfall on trust behavior using SSA countries. They show that higher rainfall shocks in the first five years after birth substantially increase localized trust of one's relatives and neighbors in adult life. They also mention rainfall shocks is one of the primary shocks in early childhood since it is one of the determining factors for agriculture and living conditions.

Another work by Jayachandran (2006) uses rainfall as an instrument for agricultural productivity because rainfall is exogenous. The author offers an alternative way to construct rainfall variable. In particular, the rainfall shock variable equals one if the annual rainfall is above the eightieth percentile for the district, zero if it is between the eightieth and twentieth percentiles, or minus one if it is below the twentieth percentile. In this study, I prefer to construct rainfall shock variable as suggested by Maccini and Yang (2009) since this rainfall variable is continuous and more likely to capture the rainfall variations. Similar to the sample used in the study, Hyland and Russ (2019) examine the effects of long-term drought in SSA countries, and they find children who were exposed to drought in the birth year received lower education and were significantly less wealthy as adults.

Another strand of literature discusses the consequences of higher rainfall on precipitation abundant areas. For instance, abnormal high levels of rainfall can result in flood, which have negative effects on social economic outcomes (Carrillo, 2020). The author investigates the long-term effects of in-utero exposure to abnormal rainfall in Columbia, and finds in-utero exposure to adverse rainfall shocks leads to poorer long-term adult outcomes. Specifically, increase in prenatal flood is associated with an increase in mental disability rates, a decline in years if schooling, an increase in illiteracy rates, and a reduction in the likelihood of working. In addition, the author finds that the effects on health and educational outcomes are larger for males than for females. The paper sheds light on the fact that rainfall shocks can have heterogeneous effects based on different areas. Carrillo (2020) finds higher rainfall can generate flood in Columbia, which negatively affects long-term adult outcomes. However, this hardly provides legitimate explanation for this paper, because most SSA countries are covered by desert and suffered from water shortage.

Finally, I want to answer whether rainfall shocks affect political competition? If the answer is true, the interaction design of political competition and rainfall shocks may be problematic, because now the political competition becomes an outcome variable of rainfall shocks. Brückner and Ciccone (2011) can elucidate the answer. Their study is based on a country level rainfall and democratic institutions data, and show that "negative rainfall shocks are followed by significant improvements in democratic institutions". It is clear that country level rainfall conditions (e.g., drought) have an impact on political competitions. Thus, I exploit a local rainfall variable which is measured at a DHS cluster level.[10] Because each cluster only represents a small part of the country, it is less likely that the local rainfall shock will affect country level presidential elections.

[10]DHS cluster is the smallest geographic area identified in each round of DHS survey. This variable identifies the residence area when a mother was surveyed.

3. Data

I use three main data sources that are publicly available for the empirical analyses. The information on election outcomes is taken from the African Election Database that archives election outcomes for SSA Countries.[11] The information includes presidential elections, national assemblies' elections and national referendums, starting from the earliest elections until the latest in each country. First round and runoff elections are included when relevant. I use presidential election outcomes in 23 Sub-Saharan Africa (SSA) countries between 1990 and 2013. During the period I consider, the sample countries ran 101 presidential elections involving more than one candidate.

I extract the rainfall data from Global Climate Resources data housed at University of Delaware.[12] This data provides monthly precipitation level at 0.5 degree resolution satellite imaginary, covering the entire Africa, with more than 100 years data from 1900. I aggregate the monthly rainfall data to obtain the yearly rainfall data.

I am primarily interested in the effect of political competition on household level outcomes. I proxy household economic well-being by child mortality, and child and maternal health care including immunization and births assisted by health professionals. The demographic and health surveys (DHS) provides measures of these variables that are consistent across countries and across time.[13] I only use the rounds of DHS surveys that provide GPS coordinates to merge with the nearest rainfall data.

Combining all the data above, the main analytical sample consists 20 SSA countries with 38 rounds of DHS surveys. The sample covers children born from 1988-2013.

[11] http://africanelections.tripod.com/index.html
[12] http://climate.geog.udel.edu/ climate
[13] http://www.dhsprogram.com/

4. Methods and Models

To empirically access three hypotheses listed above, I estimate linear regressions in the forms below:

$$y_{imlcat} = \beta_0 + \beta_1 compet_{ca} + \beta_2 X_{imlcat} + \beta_3 Z_{mlct} + \theta_a + \gamma_{mlct} + \varepsilon_{imlcat} \tag{6}$$

The first hypothesis is estimated by equation (6). Child is indexed by i, mother by m, DHS cluster by l, country by c, birth year by a, and survey year by t. The outcome variable y describes the mortality status for child i, born to mother m, in country c, in year a, and who was surveyed in year t. The variable $compet_{ca}$ measures how strongly the presidential election was contested in year a, i.e. in the year when the child was born. I define $compet_{ca}$ as 100 minus the share of votes of the presidential election winner in percentage. Intuitively, using this measurement of $compet_{ca}$, a lower share of votes of the presidential winner refers to a more competitive election. X_{imlcat} and Z_{mlct} represents child level and mother level observed characteristics, respectively. I include variables child gender and child birth order. θ_a is child birth year fixed effects, which is used to control for birth-specific shocks. γ_{mlct} represents mother fixed effects, which can absorb time-invariant unobserved heterogeneities up to a specific mother Kudamatsu (2012). To illustrate, let consider two siblings, each born at the beginning of a presidential term of five years. Sibling A is born when the president is freshly elected with a large margin. In contrast, sibling B is born when the president is elected with a smaller margin. I am interested in testing whether sibling B is less likely to die in childhood than sibling A, as a result of being born in a period of a more competitive presidential mandate. The central hypothesis is that more contested elections would have induced the president to promote more effective public policies when sibling B was born (references).

The main identification strategy is based on mother fixed effects. However, because time-varying country level variables (say δ_{ct}) are likely to affect child mortality rates and elections outcomes, OLS estimate of β_1 in equation (6) may be biased. For example, improving economic condition over time could both affect child mortality and political competitiveness. Likewise, increasing literacy rates or a growing middle class could make elections more contested and reduce child mortality simultaneously. Alternatively, a negative economic shock (especially if badly handled by the sitting government) could increase child mortality but may influence political outcomes (Baird et al., 2011; Brückner and Ciccone, 2011). To deal with this issue, I include a rich set of controls to proxy for time-varying country level variables. I first include two additional variables to capture runoff and incumbent presidential candidates' status. I then argue that given the structure of African economies, annual prices of main exports and imports and time interacted with regional and colonial background fixed effects can control most of the annual changes in each country. Specifically, I consider the main following main regions: West Africa, East Africa, Central Africa and Southern Africa. Colonial background includes "French", "British" and "Spanish". Additionally, I also include the main agro-climatic regions (e.g., the Sahel). Each of these regions usually includes several countries and more than one agro-climatic region could be found in many countries (e.g., Cameroon, Cote d'Ivoire, Ghana, etc.). These variables interacted with time trend could control for most unobserved time-varying effects and still leave enough variations in the competition variable to be identified. Finally, I also control for annual prices of the main export and import goods, including oil, aluminum, bananas, cocoa, coffee, cotton, groundnuts, log, maize, rice, tea and uranium. Controlling for these proxy variables should reduce the partial correlation between δ_{ct} and $compet_{ca}$ and hence the correlation between $compet_{ca}$ and ε_{imlcat}, which in turn will reduce the bias in the estimate of β_1. In main results section, I compare estimates with different combinations of these covariates to provide an insight about the bias of the estimates.

$$y_{imlcat} = \beta_0 + \beta_1 rainfall_{la} + \beta_2 X_{imlcat} + \beta_3 Z_{mlct} + \theta_a + \gamma_{mlct} + \varepsilon_{imlcat} \tag{7}$$

The second hypothesis is estimated by equation (7). Child is indexed by i, mother by m, DHS cluster by l, country by c, birth year by a, and survey year by t. The outcome variable y describes the mortality status for child i, born to mother m, residing in DHS cluster l, in year a, and who was surveyed in year t. The main treatment variable $rainfall_{la}$ is birth year rainfall deviation, which is defined as rainfall at the cluster l in logarithm minus its long-term average in logarithm (Maccini and Yang, 2009). X_{imlcat} and Z_{mlct} represents child level and mother level observed characteristics. Similar to equation (6), θ_a is child birth year fixed effects, and γ_{mlct} represents mother fixed effects. Compared to equation (6) using $compet_{ca}$ as the main treatment variable, the rainfall variable $rainfall_{la}$ is less likely to be endogenous, as the rainfall is widely used as an exogenous shock in many previous works (Hyland and Russ, 2019; Maccini and Yang, 2009). However, to test the sensibility of rainfall regressions in equation (7), I include similar covariates to capture country level time-varying factors as equation (6).

$$
\begin{aligned}
y_{imlcat} = \beta_0 + \beta_1 compet_{ca} + \beta_2 rainfall_{la} + \beta_3 compet_{ca} \times rainfall_{la} \\
+ \beta_2 X_{imlcat} + \beta_3 Z_{mlct} + \theta_a + \gamma_{mlct} + \varepsilon_{imlcat} \quad (8)
\end{aligned}
$$

Finally, I estimate equation (8) for the third hypothesis. Child is indexed by i, mother by m, DHS cluster by l, country by c, birth year by a, and survey year by t. I include both $compet_{ca}$ and $rainfall_{la}$ from equation (6) and (7), and their interaction term $compet_{ca} \times rainfall_{la}$. I want to measure whether the effects of political competition on child mortality rates can be heterogeneous respected to different rainfall in the birth year. To make the interpretation clearer, I report the marginal average treatment effects. It is possible that rainfall may affect election outcomes. For example, in rich rainfall years, farming workers have better harvest, hence more likely to cast vote on the incumbent president. This issue may bias the estimate. However, since I measure the rainfall at DHS cluster level, in which each cluster only contains a very small part of the country. It is less likely that the cluster rainfall will affect country level presidential elections.

5. Results

5.1. Main Results

Table 5 shows the descriptive statistics for all the variables. The sample is restricted to rural areas because climate factors (e.g., rainfall) are more likely to have a larger impact in rural areas. In general, child mortality in SSA countries are relatively high. Child mortality rates by age one, three, and five are 8.7%, 12.4%, and 13.1%, respectively. The main variables of interest are political competition and rainfall deviation, and the means of them are 0.0303 and -0.0098, respectively. Summary statistics of other control variables are also listed in Table 5.

Results for the effects of political competition on child mortality by age one are reported in Table 6, and they are estimated by Equation (6). I include different specifications of control variables and fixed effects to test the sensitivities of different estimates. In column (1), I control for mother fixed effects and I find a one unit increase in political competition decreases child mortality at age one by 0.4859. This indicates a one standard deviation increase in political competition reduces child mortality at one by 9.4% of the mean.[14] In column (2), I include a dummy for girl to capture gender difference, birth order variable, and two dummies to consider runner-off and incumbent president impacts. In addition, as noted in Section 4, I also control for annual prices of the main export and import goods, including oil, aluminum, bananas, cocoa, coffee, cotton, groundnuts, logs, maize, rice, tea and uranium. These variables are served to reduce the partial correlation between the error terms and the main political competition variable of interest. I find a one unit increase in political competition decreases child mortality at age one by 0.3109. In column (3), I include additional time interacted with regional and colonial background fixed effects to control most of the annual changes in each country, and I find a one unit increase in political competition is associated with child mortality by 0.2265. In column (4) (5), and (6), I test birth year effects in addition to mother fixed effects,

[14] $(-0.4859 \times 0.0168/0.0866) \times 100\% = -9.42\%$

control variables, and regional specific time trend to capture birth year specific shocks. All estimates are highly statistically significant, though with slightly lower magnitudes. Summarily, in Table 6, I find strong effects of political competition on child mortality by age one.

In Table 7, I repeat the empirical specifications in Table 6 for child mortality rates by age three. In column (6), I include mother fixed effects, all other control variables, and regional time trends. I find a one unit increase in political competition is associated with 0.2127 unit decrease in child mortality by age three; that is, a one standard deviation increase in political competition is associated with 4% decrease in child mortality. In Table 8, I find similar effects that more competitively political competition reduces child mortality by age five statistically significantly across all the specifications from column (1) to (6).[15]

Next, I examine how positive rainfall shocks affect child mortality by age one using Equation (7), and results are present in Table 9. In column (1), controlling for mother fixed effects, a one unit increase in rainfall deviation is associated with 0.0057 unit decrease in child mortality. This indicates a one standard deviation of a positive rainfall shock decreases child mortality by 1.17% of the mean.[16] I include additional control variables in column (2) and control for regional specific time trends in column (3), but results are not statistically significant. I also investigate imposing birth year fixed effects, though results are not statistically significant in column (4), (5), and (6). A possible explanation can be that mother fixed effects have already absorbed a lot of variation, so that the additional birth year fixed effects wipe out most yearly variations for the rainfall variable.

Similar to Table 9, I explore effects of rainfall shocks on child mortality by age three in Table 10. In column (1), I find a one unit positive rainfall shock reduces child mortality at three by 0.0175 (2.5% of the mean). Including control variables in column (2) and

[15] I also test including per capita GDP (current US$) or its lagging value as another control. Main results do not change qualitatively. However, in main Tables 6-14, I do not include per capita GDP as a control variable. Because per capita GDP is more likely to be an outcome variable of political competition, so including it can cause "bad controls" problems. Using lagging value cannot deal with this issue since a freshly elected president will be in office for multiple years.

[16] $(-0.0057 \times 0.178/0.0866) \times 100\% = -1.17\%$

regional time trends in column (3) still have significant negative estimates, but with smaller magnitudes. However, controlling for birth year fixed effects in column (4), (5), and (6) yield insignificant estimates. Finally, I test how rainfall shocks affect child mortality by age five in Table 11. I find strong effects that a positive rainfall shock decreases child mortality without birth year fixed effects. Including birth year fixed effects still pertain some significance but not very strong.

I now focus on the interaction effects of political competition and rainfall shocks. I want to investigate whether the effects of political competition on child mortality rates can be heterogeneous respected to different rainfall in the birth year. Table 12 reports the interaction effects on child mortality by age one. In Panel A column (1), I find the competition variable is still a significant negative number (-0.4851) and almost identical to column (1) in Table 3, where the political competition effects are estimated independently. But the rainfall variable becomes insignificant. The interaction is a significant point estimate of -0.3605. To make the interaction term interpretative, I report the marginal average treatment effects of political competition and rainfall deviation variable in Panel B column (1). The point estimate is a significant -0.4814 for the political competition variable and an insignificant -0.0044 for the rainfall variable. Since the marginal average treatment effects of political competition are almost identical to column (1) in Table 6, I do not find any significant interaction effects on child mortality by age one. I also explore different combinations of control variables, regional time trends, and birth year fixed effects through column (2) to (5), and I obtain similar outcomes as column (1).

In Table 13, I examine the interaction effects of political competition and rainfall for child mortality before age three. Generally speaking, most estimates for child mortality at age three are similar to age one shown in Table 12. The only notable difference is the marginal average treatment effects of rainfall become significant in Panel B column (1), (2), and (3). But the marginal effects of rainfall shocks are much smaller than political competition shocks (e.g., in column (1), -0.7204 for political competition and -0.0156 for rainfall). This shows positive rainfall shocks are less effective to reduce child mortality when political competition is more competitive. It may suggest that positive rainfall shocks are not complementary to more competitive political competitions. Therefore, when a positive shock is added to another one, a basic economic principle "diminishing return to scale" may be the reason to explain the findings. This reconciles the work by Brückner and Ciccone (2011). They find democratic institutions improve significantly following negative rainfall shocks (e.g., drought). Table 14 shows the results for child mortality by age five. Again, I obtain similar results.

5.2. Additional Results

All results in Table 6 through Table 14 are based rural sample in SSA countries. Can I find similar results using urban sample? In Table A6, I test the effects of political competition on child mortality by age three for urban sample. It is noticeable that observations are only about one third of urban sample, whereby most African residents were living in rural areas during the sample period 1988 – 2013. Compared to Table 7 for the rural sample, I still find strong effects that a more competitive political competition decreases child mortally, while the magnitudes are little smaller. This is consistent with the hypothesis about urban/ rural bias, which refers that welfare improvement is stronger in rural Africa.

In Table A7, I test the effects of rainfall shocks on child mortality by age three in urban sample. I do not find any statistically significant effects. Table A8 reports the interaction effects of political competition and rainfall on child mortality before age three.

In Table A9 and A10, I explore the gender differences. As shown in main results, the coefficient on the girl dummy is a significant negative number. This refers that in SSA countries, girls have a lower child mortality than boys. From Table A9 and Table A10, I find in general, political competition has large effects on boys than girls to reduce child mortality.

Finally, in Table A11, I explore how political competition can affect country level variables like GDP capita and health related expenditure. I find more competitive presidential elections have significant effects on most health related expenditure variables but not GDP capita. This provides suggestive evidence on potential mechanisms.

6. Summary and Concluding Remarks

In this study, I empirically test three hypotheses that affect child mortality rates based on rural sample in Sub-Saharan African countries. I find a more competitive presidential election has significant effects to reduce child mortality in subsequent years during the period when the freshly elected president is in the office. This demonstrates convincingly the extent to which political competition can affect economic policy and economic outcomes. I find a positive birth year rainfall shock is associated with lower child mortality rates. This means the birth year environment has significant impacts in later years after the birth. I also discuss the interactive effects between competitive presidential completion and positive rainfall shocks, and I find the marginal average treatment effects are smaller for rainfall shocks. This indicates a positive rainfall shock may be less effective during a more competitive election time. I interpret this cautiously that the overall impact of favorable environmental and political conditions on child mortality exhibits a certain degree of "diminishing return to scale". It is stronger at the initial levels and the marginal effects are going down as environmental and political conditions become more beneficial.

CHAPTER III

BROADBAND, SELF-EMPLOYMENT, AND WORK-FROM-HOME — EVIDENCE FROM THE AMERICAN COMMUNITY SURVEY

1. Introduction

With the technological expansion of broadband since 2000s, high-speed internet becomes accessible and adoptable for more population. However, the adoption to high-speed internet is not evenly distributed in the U.S., even for the recent years. Figure 7 shows the distribution of adoption to high-speed broadband in the U.S. from 2013 - 2017. It is clear that the adoption of broadband is mature in metropolitan areas but still lagging in rural areas.

Previous works find access to high-speed broadband affects many economic factors including the labor market, local economic growth, and the college enrollment (Dettling et al., 2018; Kolko, 2012; Xu et al., 2019). For example, using slope of terrain steepness as an instrumental variable for broadband, Kolko (2012) shows broadband expansion is associated with higher local economic growth and employment growth. Dettling et al. (2018) examines the impact of high-speed broadband on college applications and they find access to high-speed broadband has positive effects on college applications. They argue that the main mechanism is the development of broadband technology considerably reduce students' effort to apply for schools.

In this paper, the main research question I want to address is how the adoption and access to broadband affect the labor market. Specially, I examine how adoption and access to broadband influence self-employment (SE) and work-from-home (WFH) based on married women. The existing works show that married females are more likely to be self-employed due to various reasons (Cai et al., 2019; Patrick et al., 2016). For instance, Cai et al. (2019) point out that married women spend more time on household work, so self-employment provides the flexibility to choose when to work and how much to work. Patrick et al. (2016) suggest a comprehensive set of factors determining self-employment status and they find different factors for married and unmarried women. For married women, household burdens associated with children are more like to be the driven factors. But for unmarried women, the local business climate and individual characteristics are important factors.

The closest work to this study is Dettling (2017). The author examines the effects of high-speed internet on married women's labor force participation. The main conclusion is that the use of high-speed internet leads an increase in labor force participation of married women. The author briefly tests the effects of broadband on self-employment but does not obtain statistically significant effects. My paper is different from Dettling (2017) in many perspectives and contributes to the existing literature in several ways. First, the dependent variable in this paper incorporates SE, WFH, and their combinations, so I can investigate the effects of broadband on different labor market outcomes. Second, I use the most recent 2013-2017 American Community Survey (ACS) as the main data. In this way, I can provide up-to-date inference for policy makers based on one of the largest and most reliable census survey data in the U.S. Third, my empirical analyses are based on the ACS self-reported internet data as well as the Federal Communications Commission (FCC) internet data. The effects of adoption of broadband can be estimated using the ACS self-reported internet data and FCC household connections to high-speed broadband data, while access to broadband can be estimated using FCC numbers of high-speed broadband providers' data. With different sources of broadband data, this study can capture effects of broadband on SE and WFH from both the adoption and access to broadband.

2. Related Literature

2.1. Broadband and Its Impacts

The relationship between broadband and economic outcomes are correlated at the local level (Kolko, 2012). A majority of literature finds the availability of local broadband connections have strong positive effects on local economic indicators, such as population growth, employment growth, college enrollment, and innovation (Dettling et al., 2018; Kolko, 2012; Xu et al., 2019). The main mechanism whereby broadband contributes to local economic development is the availability of internet connections allows higher levels of economic activities.

Kolko (2012) examines the impacts of access to broadband on local employment growth and other indicators, and the author shows that access to broadband is strongly associated with local employment growth, but with lower median household incomes. It suggests that the availability of broadband attracts more migrants, from which increase the local labor supply, leaving the employment rate and average pay unchanged. In addition, this paper proposes a new instrumental variable to deal with endogeneity issues for access to broadband. Specifically, broadband construction can be both demand and supply driven, and it is correlated with many local indicators. Kolko (2012) argues that average slope of the local terrain is a valid instrument for broadband availability, as it satisfies both correlation and exclusion for a good instrument. Construction of broadband is associated with higher cost in steeper terrain, so broadband is negatively correlated with the slope of terrain. This instrument also satisfies the exclusion criteria, since the slope of local terrain is not likely directly associated with local economic growth. In addition to Kolko (2012), a large number of studies find access to broadband improves local economic growth (Ivus and Boland, 2015; Van Gaasbeck, 2008; Whitacre et al., 2014).

Besides local employment and economic growth, a wide local economic indicators are examined by the existing literature. Dettling et al. (2018) investigate whether the availability of broadband can improve students' postsecondary outcomes. They find the availability of broadband during students' junior year encourage students to send more college applications, apply to more out-of-state universities, and get higher SAT scores. The authors suggest that the main mechanism is that the availability of broadband technology directly decreases efforts and costs associated with the application process. Some papers find broadband diffusion affects entrepreneurships, business start-ups, R&D collaborations, and Innovation. Using data in Ohio as a case study, Mack (2014) finds fast-speed broadband is an important factor for agricultural and rural business establishments. Tranos and Mack (2016) show that an increase in number of broadband providers is associated with more knowledge-intensive business services. Parajuli and Haynes (2017) examine the impacts of broadband on new firm formation using data from Florida and Ohio, and the authors find positive effects exist but vary spatially. Based on spatial econometric models, Mack et al. (2011) find broadband infrastructure is important for knowledge intensive firms to locate in certain areas. Forman and Zeebroeck (2012) study the effects of adoption of internet on firm patents and find significant effects that internet adoption increases the likelihood of collaborative patents from geographically dispersed teams. Finally, Xu et al. (2019) find access to internet is associated with increased patents, and they suggest the mechanism is internet connection reduces disseminating costs of filling patents locally.

Apart from the positive effects of broadband, a narrower amount of works studies that broadband is negatively correlated with some local economic indicators, shedding light on the social challenges from increased internet access. Conley et al. (2016) find that a high level of broadband adoption may have negative impacts on number of entrepreneurs and creative class employees in rural America. Chan et al. (2016) examine the relationship between internet and racial hate crime and show that broadband access increases racial hate crimes. Similarly, using data from Norway, Bhuller et al. (2013) find that access to broadband increases sex crime propensity, possibly due to factors like increased consumption of pornography.

2.2. Self-Employment and Work-from-Home

Why workers leave wage employment and select into SE? Earlier works by Boden Jr (1996, 1999) discuss several possible reasons. First, workers' wage employment experience affects the probability of self-employed. Specifically, workers employed in small businesses are more likely to leave wage employment and become self-employed. Second, as for gender differences, responsibility for childcare is an important factor that pushes women into self-employment, while for men, presence of children is not an important determinant of self-employment. Third, workers in certain occupations such as clerical and administrative are more likely to switch to self-employment. Budig (2006) also argue that the effects of family structure on self-employment differ by occupational classes. Bruce (1999) show that husbands can affect wives' decision of self-employment. The author finds having a husband with self-employment experience increases the likelihood that a wife is self-employed.

Besides family and occupations, individual characteristics are strong predictors of self-employment. For example, Carr (1996) shows that human capital factors like education, age, and past working experience are important factors that influence self-employment for both men and women. In addition, some psychological factors may also affect self-employment. Clain (2000) mentions that women who are engaging in full self-employment may be less highly valued as women who work for full-time wage employment in the marketplace. Fairlie and Meyer (1996) investigate ethnic and racial factors that affect self-employment, and they find that racial groups immigrated from countries with high self-employment rate are less likely to be self-employed in the U.S.

Besides all the factors and determinants of self-employment across different groups, Hamilton (2000) analyses the returns of self-employment compared to wage employment. The author argues that the returns of self-employment are not only from earnings, but also from nonpecuniary benefits, such as "being your own boss". This implies that self-employed workers are willing to sacrifice substantial amount of earnings in exchange to nonpecuniary benefits such as flexibility of working and owning a business. Hundley (2000) discusses earning differences between women and men respect to household factors. The author finds that self-employed women's earnings decline with marriage, family size, and hours of housework, while for self-employed men, earnings increase with marriage and family sizes.

Compared to self-employment, working from home is more related to the location of work, but not the intrinsic characteristics of work — self or wage employment. Therefore, it is intuitive that the adoption of broadband is likely to increase the probability of work-from-home, because of the easiness and convenience people receive from the home internet service. The existing literature related to work-from-home is scant but emerging in recent two years, especially after the outbreak of the COVID-19 pandemic.

One of the earlier works by Felstead et al. (2001) propose seven hypotheses related to WFH. They suggest internet is a key facilitator for working from home and working from home is predominately a female activity. Felstead et al. (2002) find working from home is more likely to achieve work-life balance. Sarbu (2015) point out the determinants of WFH contain education, tenure status, use of computers, and presence of young children. However, Hill et al. (2003) hold a different view that WFH may blur the boundaries of work and family balance for certain occupations, which would result to less successful personal or family life.

Using Chinese travel company data, Bloom (2014) and Bloom et al. (2015) find workers who are allowed to telecommute and work-from-home are less likely to leave the company and more productive. The author suggests that one to two days working from home per week is beneficial to workers' well-being, especially for the occupations that remote work can be accomplished easily such as working at call centers. But the positive effects may not exist for some occupations, as it is not feasible for some type of workers to work from home. Moreover, the authors also suggest there are some potential problems of WFH, such as it is hard for employers to evaluate employee's performance.

A recent work by Dingel and Neiman (2020) investigates what jobs can be done at home. They find like computer, education, training, business, and financial occupations can be done at home, but construction, extraction, food preparation, and maintenance occupations are almost impossible to be accomplished at home. Belzunegui-Eraso and Erro-Garcés (2020) suggest during the Covid-19 pandemic, allowing employees to telework is a good way to ensure employees' safety while provide continuity to economic activities. Gottlieb et al. (2021) propose a measure of the ability to WFH from developing countries. They suggest workers are less educated and in low-wage occupations are less likely to work remotely, and WFH ability is not only related to occupations but also the nature of tasks within occupations.

3. Data

The main data is obtained from the American Community Survey (ACS) via IPUMS (Ruggles et al., 2020). The ACS is an annual survey of one percent of U.S. population, and collects labor force and employment, demographics, education, migration, and other information. From 2013, the ACS starts to report two internet related variables — access to internet and subscription to high-speed broadband. These two variables are at a household level. I use these two variables to measure adoption of broadband and high-speed broadband. In particular, I use ACS 2013-2017 from which these two internet related variables are reported. To protect confidentiality, the smallest identified geographic areas are Public Use Microdata Area (PUMA).[17]

In addition to two self-reported internet variables, I obtain a second source of internet data from the Federal Communications Commission (FCC). Previous literature notes FCC broadband data is one of the highest quality data for broadband (Grubesic and Mack, 2015). Specifically, I use two variables from FCC form 477 data on internet access services. The first variable reports share of population with access to high-speed internet, and this variable is used for the adoption of broadband. I aggregate this data from a

[17] To protect confidentiality, the smallest identified geographic areas in ACS 2013-2017 is Public Use Microdata Area (PUMA). PUMA is generally following county boundaries. For small counties with less than 100K residents, they are combined as a PUMA. There are 2,351 unique PUMAs covering the entire U.S.

census tract level to a PUMA level. The second variable reports number of high-speed broadband providers, and it is used to measure the local access to high-speed broadband. I aggregate this data from a census block level to a PUMA level. The last data source used in this study is USDA topography index. I use this variable as instrumental variables for high-speed broadband providers.

Combining all the data sources, the main analytical sample is restricted to married women, covering the entire U.S. at a PUMA level, from 2013-2017. Because broadband providers' information is not compatible across the entire 2013-2017, regressions concerning access to broadband are based on 2014-2016 data only.

4. Methods and Models

In the main specification, I estimate a linear probability model (LPM) to investigate the probability of SE and WFH:

$$y_{imt} = \beta_0 + \beta_1 Broadband_{imt} + \gamma_1 X_{imt} + \gamma_2 P_{mt} + \pi_m + \varphi_{rt} + \mu_{imt} \tag{9}$$

The dependent variable is a dummy variable equal to one if an individual i living in PUMA m in year t is self-employed, work-from-home, or combinations of these two. Two ACS reported internet related variables are used as the main treatment variable $Broadband_{imt}$. The first one is a dummy equal to one if an individual has subscription to high-speed internet. The second one is a dummy equal to one if an individual has adoption to internet. X_{imt} is a set of individual control variables including age, age square, race groups, and education levels. P_{mt} is a set of PUMA specific control variables including share of population with a bachelor degree or above. π_m is PUMA fixed effects and φ_{rt} is survey year by region fixed effects to absorb time-invariant confounders at PUMA level and region specific shocks in year t.[18] μ_{imt} is the error term. The main sample includes married women age between 18 and 59.

[18]The four US Census regions are the Northeast, Midwest, South, and West.

Self-reported internet variables from the ACS may have measurement errors and more importantly, are possibility endogenous. It is not possible to control all the confounders in the error term correlated with the main internet treatment variable, so I employ FCC local adoption to broadband data as a second source for broadband. I slightly modify equation (9) to run LPM in equation (10):

$$y_{imt} = \beta_0 + \beta_1 Broadband_{mt} + \gamma_1 X_{imt} + \gamma_2 P_{mt} + \pi_m + \varphi_{rt} + \mu_{imt} \tag{10}$$

Compared to equation (9), equation (10) revises the main variable of interest from $Broadband_{imt}$ to $Broadband_{mt}$. $Broadband_{mt}$ is a continuous variable measures residential fixed high-speed connections over 200 kbps in at least one direction per code connections per 1000 Households.[19] FCC categorizes number of connections per 1,000 households into six groups: zero if no connections, one if connections are greater than zero and lower than 200, two if connections are greater than 200 and lower than 400, three if connections are greater than 400 and lower than 600, four if connections are greater than 600 and lower than 800, five if connections are greater 800. Although there are only six groups in the original FCC census tract data, this variable becomes a continuous variable range from zero to five after aggregated to a PUMA level.

Based on equation (10), to quantify the effects of local access to high-speed broadband, I use number of high-speed broadband providers as $Broadband_{mt}$. Specifically, this variable measures number of providers that can or do provide internet access service at least 10 Mbps downstream and 1 Mbps upstream to at least one residential location. The original FCC data is at a census block level, so I aggregate this variable to a PUMA level using the average number of broadband providers.

FCC reported broadband providers' data may be still correlated with unobserved factors in the error term and it is not possible to capture all the unobserved confounders. This will cause endogeneity and bias the estimates. In light of the previous literature, I use land surface terrain as topography instrumental variables (IVs) for FCC access to internet

[19]FCC census tract data: https://www.fcc.gov/general/form-477-census-tract-data-internet-access-services.

variable Kolko (2012). Construction of broadband incurs high fixed costs especially in areas with unfavorable topography. According to Kolko (2012), flatter terrain constitutes good geography for broadband deployment. I use USDA land surface index as a proxy to steepness of terrain. This index has 21 values. Larger values mean steeper areas. For example, one means flat plains, 11 means plains with high mountains, and 21 means high mountains. A good instrument requires both correlation and exclusion. As noted in Kolko (2012), flatter areas are more viable for broadband construction, so larger land surface topography index areas are less likely to attain broadband. But the steepness of terrain is less likely correlated with the self-employment and work-from-home variables directly. In equation (10), I include PUMA fixed effects, in which absorbs time-invariant variables including the land surface topography index. So, I interact the land surface topography index with year dummies to construct the instruments to allow PUMA fixed effects Chan et al. (2016). Since I use the year dummies interacted with the land surface topography index to form the IVs, year by region fixed effects are not included in the regressions.

5. Results

5.1. Main Results

Figure 7 shows the mean distribution of adoption to broadband from 2013-2017 based on FCC broadband data. It indicates the spread of broadband is not evenly distributed in the U.S. In the central U.S. areas, there are still many places with less than 75% adoption to high-speed broadband. Figure 8 displays the distribution of average high-speed providers from 2014-2016. Table 15 shows the descriptive statistics for all the variables used for regressions and grouping criteria for sub-sample estimation.

Table 16 presents the results estimated by equation (9) using the ACS adoption to high-speed broadband variable as the main treatment variable. This variable is a dummy equal to one if a member has subscription to the internet using high-speed broadband such as cable, fiber optic, or DSL service. To explore MSA and non-MSA areas' heterogeneities, estimation sample is restricted to non-MSA areas in Panel A, and MSA areas in Panel B. In Panel A, column (1) shows the adoption to high-speed broadband increases the probability to be SE by 0.7%. In column (2), the dependent variable is the WFH dummy. It shows adoption to high-speed broadband increases the probability of WFH by 0.82%. I also estimate different combinations of SE and WFH in column (3), (4), and (5). In particular, the dependent variable is WFH and SE in column (3); the dependent variable is WFH but not SE in column (4); and the dependent variable is SE but not WFH in column (5). I find statistically significant positive effects of adoption to high-speed broadband on WFH and SE combinations in all column (3), (4), and (5), though the magnitudes are smaller than column (1) and (2). Results for MSA sample are presented in Panel B. All the estimates are similar to non-MSA estimation in Panel A. This shows for adoption of high-speed broadband, effects are not very different between non-MSA and MSA areas. It may suggest that broadband adoption is more related to the demand side, which is not largely affected by the access of broadband from the supply side as the broadband infrastructure is broadly available in recent years.

In Table 17, I estimate the effects of adoption of high-speed broadband on SE and WFH separated for women with children in Panel A, and women without children in Panel B. In Panel A column (1), adoption to high-speed broadband is associated with 0.76% increase in SE; while in column (2), adoption to high-speed broadband is associated with 0.82% increase in WFH. Significantly positive effects are also obtained in column (3), (4), and (5). Effects on women without children are reported in Panel B, and all the estimates are similar to Panel A. In summary, Table 16 and 17 show that the adoption to high-speed broadband has significant effects to increase SE and WFH, and the effects are similar between non-MSA and MSA areas, and women with or without children.

In Table 18, I estimate equation (1) again using ACS adoption to broadband variable for non-MSA and MSA sample. The main treatment variable is assigned to one if a member of the household has adoption to broadband. Compared to the main high-speed broadband variable in Table 16, adoption to broadband does not restrict the speed of internet. Similar to Table 16, I find statistically positive effects in all the columns in Table 18. For example, Panel A column (1) shows adoption of broadband increases the probability of SE by 1.5%. In Table 19, estimation is separated for women with and without children. I also find significant effects in all columns in both Panel A and Panel B. Results obtained in Table 16 through 19 are estimated by using the ACS self-reported broadband variables, so there may be measurement errors or endogeneity issues. Although I include a comprehensive set of control variables, there may be unobserved factors that may bias the estimation.

Table 20 reports regression results estimated by equation (10). In this Table, I use the FCC reported local adoption of broadband variable at a PUMA level as the main treatment variable. In Panel A column (1), it shows local adoption of high-speed broadband increases by 20% is associated with 1.03% increase in SE. No statistically significant effects are obtained on WFH in column (2). For MSA sample estimates reported in Panel B, I still find significant positive effects of local adoption of high-speed broadband on SE not WFH, but the coefficients are much smaller. In general, Table 20 shows higher adoption of broadband encourages SE but not WFH, and effects are stronger for non-MSA areas. This may suggest local adoption of high-speed broadband has higher effects for non-MSA areas than MSA areas, possibly because MSA areas have high levels of broadband adoption, additional adoption may have lower effects to increase SE. Table 21 presents how local adoption of high-speed broadband affects SE and WFH for women with children and without children. It shows that the effects are stronger for women without children.

Table 16 through 21 focus on how the adoption of broadband affects SE and WFH decisions. Results for the impacts of access to broadband on SE and WFH are shown in Table 22 to Table 25. I use number of local high-speed broadband providers to measure access to broadband, but due to data limitation, the data is consistent from 2014-2016. So, the main estimation sample is restricted to ACS 2014-2016.

Table 22 presents the effects of broadband providers on SE and WFH for non-MSA sample in Panel A and MSA sample in Panel B. For non-MSA sample, more high-speed broadband providers are associated with higher probability of WFH but not SE. For example, in column (4), the point estimate is 0.00175, which shows a one unit increase in broadband providers increases the probability of WFH by 0.18%. MSA sample estimation is reported in Panel B, and I find high-speed broadband providers have statistically significant effects in all column (1) to (5). Table 23 shows the effects of broadband providers on SE and WFH for women with children in Panel A and women without children in Panel B. Most estimates are positive and statistically significant except Panel B column (1) and (5), which show that access to high-speed broadband providers have no significant effects on WFH.

As noted in the previous literature, access to broadband is not exogenous (Kolko, 2012). For example, areas with potentially higher economic growth rate due to unobserved factors may have higher broadband coverage. In light of the existing works, I use USDA topography index interacted with the year dummies as instrumental variables for high-speed broadband providers (Kolko, 2012; Chan et al., 2016). IV results are reported in Table 24 and 25 and their first stage estimates are shown in Table A16. Table 24 presents results for non-MSA sample in Panel A and MSA results are shown in Panel B. In Panel A column (1), a one unit increase in high-speed broadband providers is associated with 0.62% increase in SE. In column (2), I find a one unit increase in providers raises the probability of WFH by 0.65%. Panel B shows similar effects for MSA sample as non-MSA sample in Panel A. Table 25 shows the IV estimation separated for women with children and without children. I find significant effects of broadband on SE for women with children but not women without children. This is consistent with the previous works that women with children are more likely to choose SE due to they are more responsible for household work. The impacts of broadband on WFH are significant for both women with or without children.

Comparing results estimated by broadband adoption versus broadband access, I find in general, both adoption and access to broadband contribute to women's SE or WFH decisions, but the broadband adoption seems to have stronger effects. There are three possible reasons. First, the main data used for analyses are very recent data (2013-2017). Broadband construction started to roll out in the U.S. in early 2000s, but after more than a decade, the access of broadband become available for most areas (see Figure 8). This may indicate now the adoption of broadband is more important than access to broadband for labor market decisions like SE and WFH in the recent years. Second, due to the FCC data availability, numbers of broadband providers' data are only consistent for year 2014-2016 but not 2013 and 2017. So, effects of broadband adoption and broadband access on SE and WFH may not directly comparable. Third, for broadband adoption estimation, I include a rich set of control variables, but cannot offer a clean causal effect identification. Moreover, using land topography as an IV for broadband access is clearly exogenous using early years' broadband data (Kolko, 2012). However, this IV may not be fully functional for more recent broadband data as the construction of broadband was upgraded and became easier even in areas with steep terrain.

5.2. Additional Results

I also test several other outcome variables related to SE and WFH. If some marginal workers choose to SE or WFH from being out of the labor force, the hours worked per week might decrease (Cai et al., 2019). In Table A12, I show how high-speed broadband affects hours worked per week for women living in non-MSA and MSA areas. In Panel A column (2), adoption to high-speed broadband decreases hours worked per week by 0.75 for WFH women. I find larger negative effects for WHF and SE sample in column (3). Table A13 presents the effects of adoption to high-speed broadband on hours worked per week for women with and without children. I find significantly negative effects in Panel A column (3). Finally, in Table A14 and A15, I test whether access to high-speed broadband have effects on women's labor force participation rate. Similar to Dettling (2017), I find the adoption of broadband has significant positive effects on labor force participation for married women.

6. Summary and Concluding Remarks

In this paper, based on the ACS and FCC broadband data, I empirically assess the effects of adoption and access to broadband on SE and WFH for married women. I find both adoption and access to high-speed broadband increase the probabilities of SE and WFH. The positive effects are stronger for women living in non-MSA areas and women with children. This paper contributes to the existing literature that examines how ICT affects the labor market. During the period of ongoing COVID-19, people realize the importance of WFH, which is safer and can slow the spread of virus. Without high-speed internet, WFH would become much more difficult. Based on the ACS data, this paper identifies the positive effects of adoption to internet on SE, WFH, and their combinations. It suggests for the recent years as the broadband infrastructure is widely available, how to increase the local adoption of high-speed internet becomes a more important question. Adoption of broadband is related to household income levels and whether there is a need for that. Therefore, for the policy implications, stipend from the government to lower the cost of high-speed broadband adoption may be a good place-based policy, especially for low income and less developed areas. Consequently, this place-based policy may promote local well-being and economic growth.

REFERENCES

Acemoglu, D., Hemous, D., Barrage, L., Aghion, P., et al. (2019). Climate change, directed innovation, and energy transition: The long-run consequences of the shale gas revolution. In *2019 Meeting Papers*, number 1302. Society for Economic Dynamics.

Acs, Z. J., Anselin, L., and Varga, A. (2002). Patents and innovation counts as measures of regional production of new knowledge. *Research policy*, 31(7):1069–1085.

Agrawal, A., Cockburn, I., and Rosell, C. (2010). Not invented here? innovation in company towns. *Journal of Urban Economics*, 67(1):78–89.

Autor, D., Dorn, D., and Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2):161–78.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American Economic Review*, 103(5):1553–97.

Bai, J. (2009). Panel data models with interactive fixed effects. *Econometrica*, 77(4):1229–1279.

Baird, S., Friedman, J., and Schady, N. (2011). Aggregate income shocks and infant mortality in the developing world. *Review of Economics and statistics*, 93(3):847–856.

Barro, R. J. and Sala-i Martin, X. (1992). Convergence. *Journal of political Economy*, 100(2):223–251.

Bartik, A. W., Currie, J., Greenstone, M., and Knittel, C. R. (2019). The local economic and welfare consequences of hydraulic fracturing. *American Economic Journal: Applied Economics*, 11(4):105–55.

Bartik, T. J. (1991). Who benefits from state and local economic development policies?

Battaglini, M. (2014). A dynamic theory of electoral competition. *Theoretical Economics*, 9(2):515–554.

Belzunegui-Eraso, A. and Erro-Garcés, A. (2020). Teleworking in the context of the covid-19 crisis. *Sustainability*, 12(9):3662.

BenYishay, A. (2013). Childhood shocks and trust in adulthood: Impacts of early-life rainfall in sub-saharan africa. In *Working Paper, Department of Economics, University of New South Wales.*

Besley, T., Persson, T., and Sturm, D. M. (2010). Political competition, policy and growth: theory and evidence from the us. *The Review of Economic Studies*, 77(4):1329–1352.

Bhuller, M., Havnes, T., Leuven, E., and Mogstad, M. (2013). Broadband internet: An information superhighway to sex crime? *Review of Economic studies*, 80(4):1237–1266.

Black, D. A., McKinnish, T. G., and Sanders, S. G. (2005). Tight labor markets and the demand for education: Evidence from the coal boom and bust. *ILR Review*, 59(1):3–16.

Bloom, N. (2014). To raise productivity, let more employees work from home. *Harvard Business Review, January–February*.

Bloom, N., Liang, J., Roberts, J., and Ying, Z. J. (2015). Does working from home work? evidence from a chinese experiment. *The Quarterly Journal of Economics*, 130(1):165–218.

Boden Jr, R. J. (1996). Gender and self-employment selection: An empirical assessment. *The Journal of Socio-Economics*, 25(6):671–682.

Boden Jr, R. J. (1999). Flexible working hours, family responsibilities, and female self-employment: Gender differences in self-employment selection. *American Journal of Economics and Sociology*, 58(1):71–83.

Bozzoli, C., Deaton, A. S., and Quintana-Domeque, C. (2007). Child mortality, income and adult height. Technical report, National Bureau of Economic Research.

Brown, J. (2018). Response of consumer debt to income shocks: The case of energy booms and busts. *Federal Reserve Bank of Kansas City Working Paper No. RWP*, pages 17–05.

Bruce, D. (1999). Do husbands matter? married women entering self-employment. *Small Business Economics*, 13(4):317–329.

Brückner, M. and Ciccone, A. (2011). Rain and the democratic window of opportunity. *Econometrica*, 79(3):923–947.

Brunnschweiler, C. N. and Bulte, E. H. (2008). The resource curse revisited and revised: A tale of paradoxes and red herrings. *Journal of environmental economics and management*, 55(3):248–264.

Budig, M. J. (2006). Intersections on the road to self-employment: Gender, family and occupational class. *Social Forces*, 84(4):2223–2239.

Cai, Z., Stephens, H. M., and Winters, J. V. (2019). Motherhood, migration, and self-employment of college graduates. *Small Business Economics*, 53(3):611–629.

Carbone, G. and Pellegata, A. (2017). To elect or not to elect: leaders, alternation in power and social welfare in sub-saharan africa. *The Journal of Development Studies*, 53(12):1965–1987.

Carlino, G. and Kerr, W. R. (2014). Agglomeration and innovation. Technical report, National Bureau of Economic Research.

Carr, D. (1996). Two paths to self-employment? women's and men's self-employment in the united states, 1980. *Work and occupations*, 23(1):26–53.

Carrillo, B. (2020). Early rainfall shocks and later-life outcomes: Evidence from colombia. *The World Bank Economic Review*, 34(1):179–209.

Chan, J., Ghose, A., and Seamans, R. (2016). The internet and racial hate crime: Offline spillovers from online access. *Mis Quarterly*, 40(2):381–403.

Clain, S. H. (2000). Gender differences in full-time self-employment. *Journal of Economics and Business*, 52(6):499–513.

Conley, K. L., Whitacre, B. E., et al. (2016). Does broadband matter for rural entrepreneurs and creative class employees? *The Review of Regional Studies*, 46(2):171–190.

Cutler, D., Deaton, A., and Lleras-Muney, A. (2006). The determinants of mortality. *Journal of economic perspectives*, 20(3):97–120.

Czarnowske, D. and Stammann, A. (2020). Inference in unbalanced panel data models with interactive fixed effects. *arXiv preprint arXiv:2004.03414*.

Deaton, A. (2007). Height, health, and development. *Proceedings of the national academy of sciences*, 104(33):13232–13237.

Dettling, L. J. (2017). Broadband in the labor market: the impact of residential high-speed internet on married women's labor force participation. *ILR Review*, 70(2):451–482.

Dettling, L. J., Goodman, S., and Smith, J. (2018). Every little bit counts: The impact of high-speed internet on the transition to college. *Review of Economics and Statistics*, 100(2):260–273.

Dingel, J. I. and Neiman, B. (2020). How many jobs can be done at home? *Journal of Public Economics*, 189:104235.

Dube, A., Lester, T. W., and Reich, M. (2010). Minimum wage effects across state borders: Estimates using contiguous counties. *The review of economics and statistics*, 92(4):945–964.

Fairlie, R. W. and Meyer, B. D. (1996). Ethnic and racial self-employment differences and possible explanations. *Journal of human resources*, pages 757–793.

Felstead, A., Jewson, N., Phizacklea, A., and Walters, S. (2001). Working at home: statistical evidence for seven key hypotheses. *Work, Employment and Society*, 15(2):215–231.

Felstead, A., Jewson, N., Phizacklea, A., and Walters, S. (2002). Opportunities to work at home in the context of work-life balance. *Human resource management journal*, 12(1):54–76.

Ferraz, C. and Finan, F. (2009). Motivating politicians: The impacts of monetary incentives on quality and performance. Technical report, National Bureau of Economic Research.

Feyrer, J., Mansur, E. T., and Sacerdote, B. (2017). Geographic dispersion of economic shocks: Evidence from the fracking revolution. *American Economic Review*, 107(4):1313–34.

Fitzgerald, T. (2012). Frackonomics: some economics of hydraulic fracturing. *Case W. Res. L. Rev.*, 63:1337.

Florida, R. (2014). The creative class and economic development. *Economic development quarterly*, 28(3):196–205.

Florida, R., Mellander, C., and Stolarick, K. (2008). Inside the black box of regional development—human capital, the creative class and tolerance. *Journal of economic geography*, 8(5):615–649.

Forman, C. and Zeebroeck, N. v. (2012). From wires to partners: How the internet has fostered r&d collaborations within firms. *Management science*, 58(8):1549–1568.

Gabe, T., Florida, R., and Mellander, C. (2013). The creative class and the crisis. *Cambridge Journal of Regions, Economy and Society*, 6(1):37–53.

Gerring, J., Knutsen, C. H., Skaaning, S.-E., Teorell, J., Coppedge, M., Lindberg, S. I., and Maguire, M. (2015). Electoral democracy and human development. *V-Dem Working Paper*, 9.

Gilje, E. P., Loutskina, E., and Murphy, D. (2020). Drilling and debt. *The Journal of Finance*, 75(3):1287–1325.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2018). Bartik instruments: What, when, why, and how. Technical report, National Bureau of Economic Research.

Gomez, M. (2017). Regife: Stata module to estimate linear models with interactive fixed effects.

Gorodnichenko, Y. and Roland, G. (2011). Individualism, innovation, and long-run growth. *Proceedings of the National Academy of Sciences*, 108(Supplement 4):21316–21319.

Gottlieb, C., Grobovšek, J., Poschke, M., and Saltiel, F. (2021). Working from home in developing countries. *European Economic Review*, 133:103679.

Grofman, B. and Selb, P. (2009). A fully general index of political competition. *Electoral Studies*, 28(2):291–296.

Grubesic, T. H. and Mack, E. A. (2015). *Broadband telecommunications and regional development*. Routledge.

Hall, B. H., Griliches, Z., and Hausman, J. A. (1984). Patents and r&d: Is there a lag? Technical report, National Bureau of Economic Research.

Hamilton, B. H. (2000). Does entrepreneurship pay? an empirical analysis of the returns to self-employment. *Journal of Political economy*, 108(3):604–631.

Han, L. and Winters, J. V. (2020). Industry fluctuations and college major choices: Evidence from an energy boom and bust. *Economics of Education Review*, 77:101996.

Harding, R. (2020). Who is democracy good for? elections, rural bias, and health and education outcomes in sub-saharan africa. *The Journal of Politics*, 82(1):241–254.

Harding, R. and Stasavage, D. (2014). What democracy does (and doesn't do) for basic services: School fees, school inputs, and african elections. *The Journal of Politics*, 76(1):229–245.

Hill, E. J., Ferris, M., and Märtinson, V. (2003). Does it matter where you work? a comparison of how three work venues (traditional office, virtual office, and home office) influence aspects of work and personal/family life. *Journal of Vocational Behavior*, 63(2):220–241.

Hundley, G. (2000). Male/female earnings differences in self-employment: The effects of marriage, children, and the household division of labor. *ILR Review*, 54(1):95–114.

Hyland, M. and Russ, J. (2019). Water as destiny–the long-term impacts of drought in sub-saharan africa. *World Development*, 115:30–45.

Ivus, O. and Boland, M. (2015). The employment and wage impact of broadband deployment in canada. *Canadian Journal of Economics/Revue canadienne d'économique*, 48(5):1803–1830.

Jaffe, A. (1986). Technological opportunity and spillovers of r&d: Evidence from firms' patents, profits, and market value. *American Economic Review*, 76(5):984–1001.

James, A. and Smith, B. (2017). There will be blood: Crime rates in shale-rich us counties. *Journal of Environmental Economics and Management*, 84:125–152.

Jayachandran, S. (2006). Selling labor low: Wage responses to productivity shocks in developing countries. *Journal of political Economy*, 114(3):538–575.

Kazianga, H. (2012). Income risk and household schooling decisions in burkina faso. *World Development*, 40(8):1647–1662.

Kejriwal, M., Li, X., and Totty, E. (2018). Multidimensional skills and the returns to schooling: Evidence from an interactive fixed-effects approach and a linked survey-administrative data set. *Journal of Applied Econometrics*.

Kjelsrud, A., Moene, K. O., and Vandewalle, L. (2020). The political competition over life and death: Evidence from infant mortality in india. Technical report, Graduate Institute of International and Development Studies Working Paper.

Kolko, J. (2012). Broadband and local growth. *Journal of Urban Economics*, 71(1):100–113.

Komarek, T. M. (2018). Crime and natural resource booms: Evidence from unconventional natural gas production. *The Annals of Regional Science*, 61(1):113–137.

Kudamatsu, M. (2012). Has democratization reduced infant mortality in sub-saharan africa? evidence from micro data. *Journal of the European Economic Association*, 10(6):1294–1317.

Kumar, A. (2017). Impact of oil booms and busts on human capital investment in the usa. *Empirical Economics*, 52(3):1089–1114.

Maccini, S. and Yang, D. (2009). Under the weather: Health, schooling, and economic consequences of early-life rainfall. *American Economic Review*, 99(3):1006–26.

Mack, E. A. (2014). Businesses and the need for speed: The impact of broadband speed on business presence. *Telematics and Informatics*, 31(4):617–627.

Mack, E. A., Anselin, L., and Grubesic, T. H. (2011). The importance of broadband provision to knowledge intensive firm location. *Regional Science Policy & Practice*, 3(1):17–35.

Maniloff, P. and Mastromonaco, R. (2017). The local employment impacts of fracking: A national study. *Resource and Energy Economics*, 49:62–85.

Muehlenbachs, L., Spiller, E., and Timmins, C. (2015). The housing market impacts of shale gas development. *American Economic Review*, 105(12):3633–59.

Munasib, A. and Rickman, D. S. (2015). Regional economic impacts of the shale gas and tight oil boom: A synthetic control analysis. *Regional Science and Urban Economics*, 50:1–17.

Parajuli, J. and Haynes, K. (2017). Spatial heterogeneity, broadband, and new firm formation. *Quality Innovation Prosperity*, 21(1):165–185.

Paredes, D., Komarek, T., and Loveridge, S. (2015). Income and employment effects of shale gas extraction windfalls: Evidence from the marcellus region. *Energy Economics*, 47:112–120.

Partridge, M., Rohlin, S. M., and Weinstein, A. L. (2019). Firm formation and survival in the shale boom. *Small Business Economics*, pages 1–22.

Partridge, M. D., Rickman, D. S., Ali, K., and Olfert, M. R. (2008). Lost in space: population growth in the american hinterlands and small cities. *Journal of Economic Geography*, 8(6):727–757.

Patrick, C., Stephens, H., and Weinstein, A. (2016). Where are all the self-employed women? push and pull factors influencing female labor market decisions. *Small Business Economics*, 46(3):365–390.

Pfister, C., Koomen, M., Harhoff, D., and Backes-Gellner, U. (2021). Regional innovation effects of applied research institutions. *Research Policy*, 50(4):104197.

Popp, D., Pless, J., Haščič, I., and Johnstone, N. (2020). Innovation and entrepreneurship in the energy sector. Technical report, National Bureau of Economic Research.

Rickman, D. S., Wang, H., and Winters, J. V. (2017). Is shale development drilling holes in the human capital pipeline? *Energy Economics*, 62:283–290.

Rozell, D. J. and Reaven, S. J. (2012). Water pollution risk associated with natural gas extraction from the marcellus shale. *Risk Analysis: An International Journal*, 32(8):1382–1393.

Ruggles, S., Flood, S., Goeken, R., Grover, J., Meyer, E., Pacas, J., and Sobek, M. (2020). Ipums usa. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN: IPUMS, `https://doi.org/10.18128/D010.V10.0`.

Sarbu, M. (2015). Determinants of work-at-home arrangements for german employees. *Labour*, 29(4):444–469.

Selb, P. (2009). A deeper look at the proportionality—turnout nexus. *Comparative Political Studies*, 42(4):527–548.

Shah, M. and Steinberg, B. M. (2017). Drought of opportunities: Contemporaneous and long-term impacts of rainfall shocks on human capital. *Journal of Political Economy*, 125(2):527–561.

Shakya, S. (2019). The dark side of the shale boom–increased crime among economically-small, relatively rural american states: Generalized synthetic control approach. Technical report, Technical report, West Virginia University.

Simonen, J. and McCann, P. (2008). Firm innovation: The influence of r&d cooperation and the geography of human capital inputs. *Journal of Urban Economics*, 64(1):146–154.

Stigler, G. J. (1972). Economic competition and political competition. *Public Choice*, pages 91–106.

Tolbert, C. M. and Sizer, M. (1996). Us commuting zones and labor market areas: A 1990 update. Technical report.

Torres-Preciado, V. H., Polanco-Gaytán, M., and Tinoco-Zermeño, M. Á. (2014). Technological innovation and regional economic growth in mexico: a spatial perspective. *The Annals of Regional Science*, 52(1):183–200.

Totty, E. (2017). The effect of minimum wages on employment: A factor model approach. *Economic Inquiry*, 55(4):1712–1737.

Tranos, E. and Mack, E. A. (2016). Broadband provision and knowledge-intensive firms: A causal relationship? *Regional Studies*, 50(7):1113–1126.

Van der Ploeg, F. (2011). Natural resources: curse or blessing? *Journal of Economic literature*, 49(2):366–420.

Van der Ploeg, F. and Poelhekke, S. (2010). The pungent smell of “red herrings”: Subsoil assets, rents, volatility and the resource curse. *Journal of Environmental Economics and Management*, 60(1):44–55.

Van Gaasbeck, K. A. (2008). A rising tide: Measuring the economic effects of broadband use across california. *The Social Science Journal*, 45(4):691–699.

Verspagen, B. (2005). Innovation and economic growth. In *The Oxford handbook of innovation*.

Weber, J. G. (2012). The effects of a natural gas boom on employment and income in colorado, texas, and wyoming. *Energy Economics*, 34(5):1580–1588.

Whitacre, B., Gallardo, R., and Strover, S. (2014). Broadband s contribution to economic growth in rural areas: Moving towards a causal relationship. *Telecommunications Policy*, 38(11):1011–1023.

Wittman, D. (1989). Why democracies produce efficient results. *Journal of Political economy*, 97(6):1395–1424.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. MIT press.

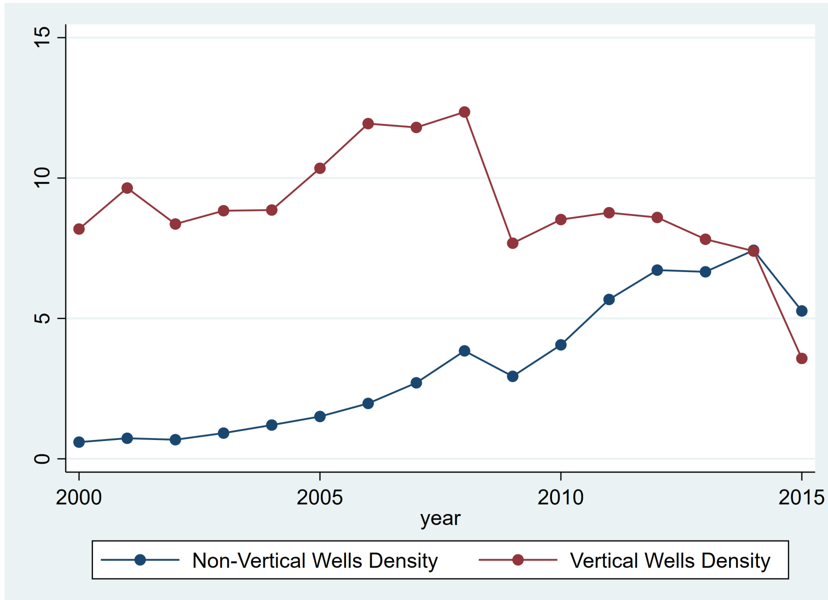
Wullert, K. E. and Williamson, J. B. (2016). Democracy, hybrid regimes, and infant mortality: A cross-national analysis of sub-saharan african nations. *Social Science Quarterly*, 97(5):1058–1069.

Xiao, H., Wu, A., and Kim, J. (2021). Commuting and innovation: Are closer inventors more productive? *Journal of Urban Economics*, 121:103300.

Xu, X., Watts, A., and Reed, M. (2019). Does access to internet promote innovation? a look at the us broadband industry. *Growth and Change*, 50(4):1423–1440.

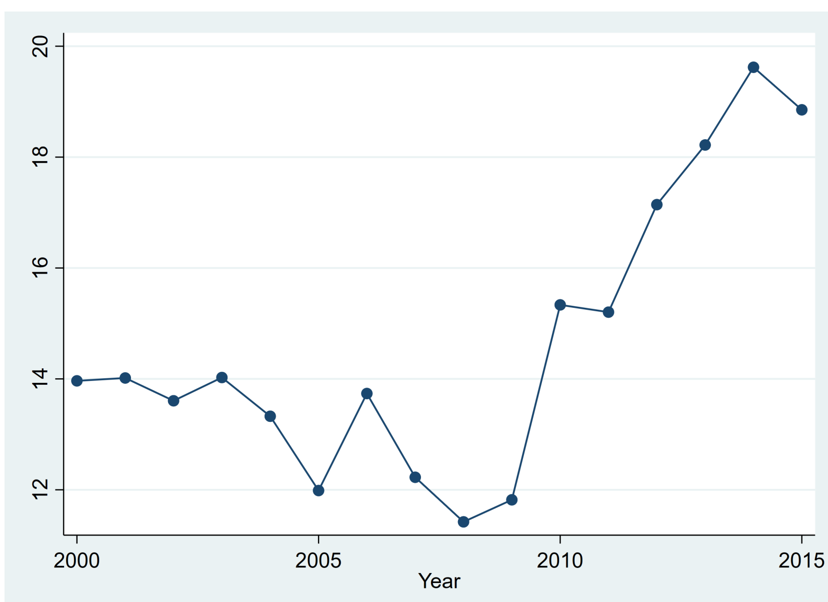
Zhang, M. and Chung, S.-h. (2020). Is air pollution detrimental to regional innovation? evidence from chinese cities. *Growth and Change*, 51(4):1657–1689.

Figure 1: Mean Drilling Well Density by Year, 2000-2015



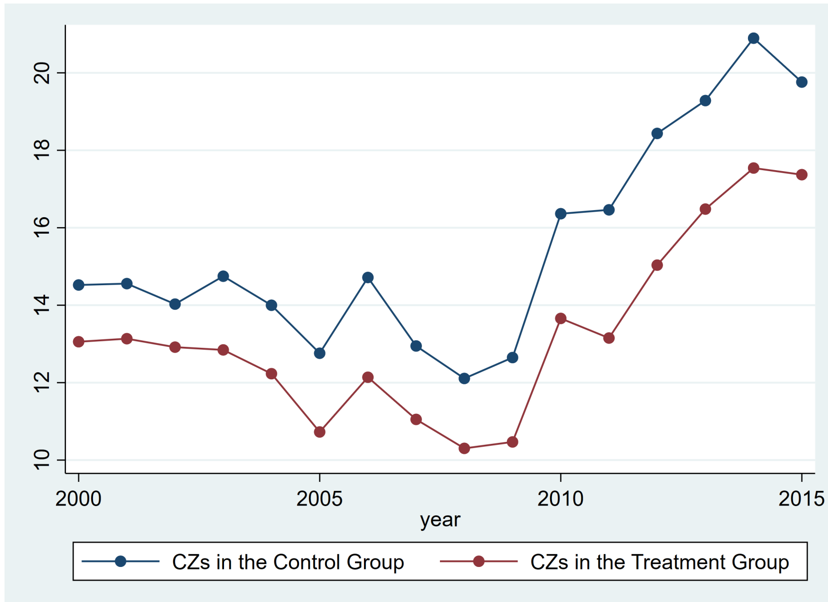
Source: Data is from Enverus.com. Non-Vertical Well Density is calculated as new non-vertical wells divided by total land area in 1000 square miles. Vertical Well Density is calculated as new vertical drilling wells divided by total land area in 1000 square miles.

Figure 2: Mean CZ Patents per 100K Population by Year, 2000-2015



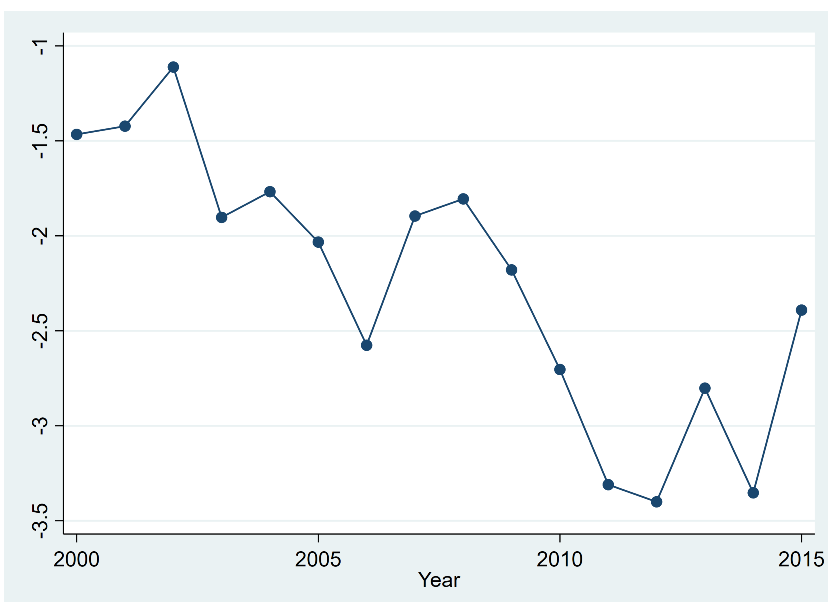
Source: U.S. Patent and Trademark Office (USPTO).

Figure 3: Mean CZ Patents per 100K Population by Groups in Year 2000-2015



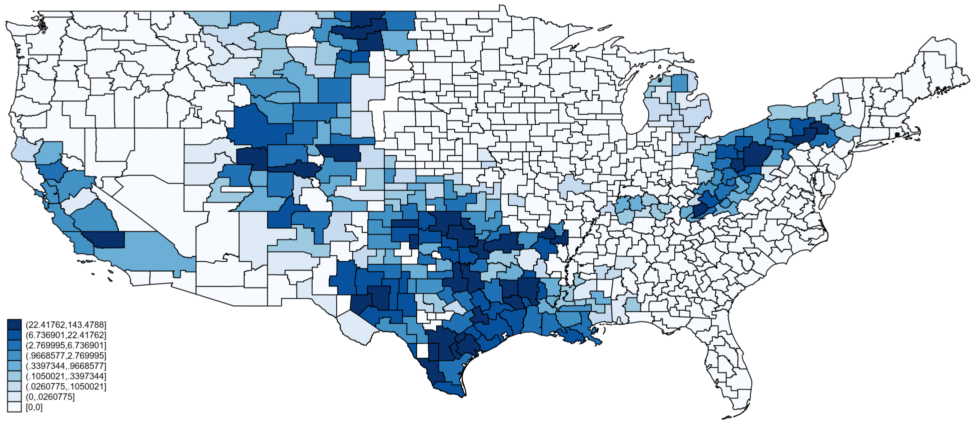
Source: U.S. Patent and Trademark Office (USPTO) and Enverus.com. Control groups include CZs without any non-vertical drilling wells in 2000-2015. Treatment groups include CZs with at least one non-vertical drilling wells in 2000-2015.

Figure 4: Mean CZ Patents per 100K Population Differences by Groups in Year 2000-2015



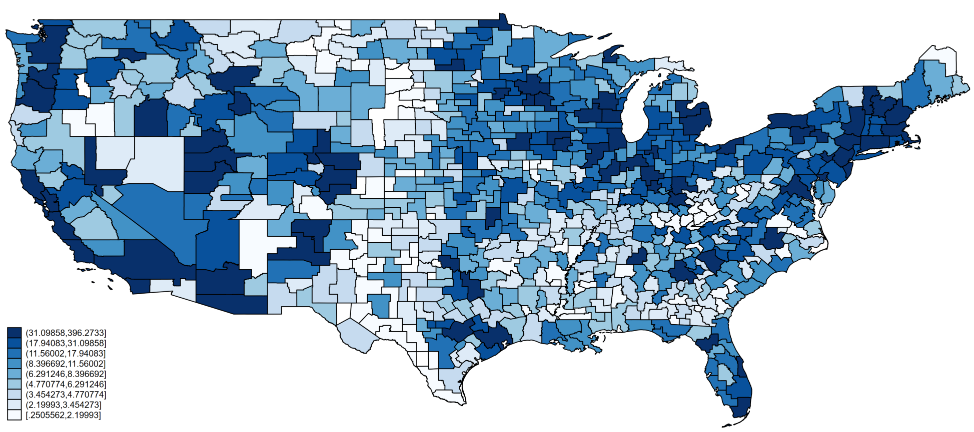
Source: U.S. Patent and Trademark Office (USPTO) and Enverus.com. The figure plots differences in patents per 100K population between CZs in the treatment group and the control group presented in Figure 3.

Figure 5: Mean New Non-Vertical Drilling Well Density Year 2000-2015



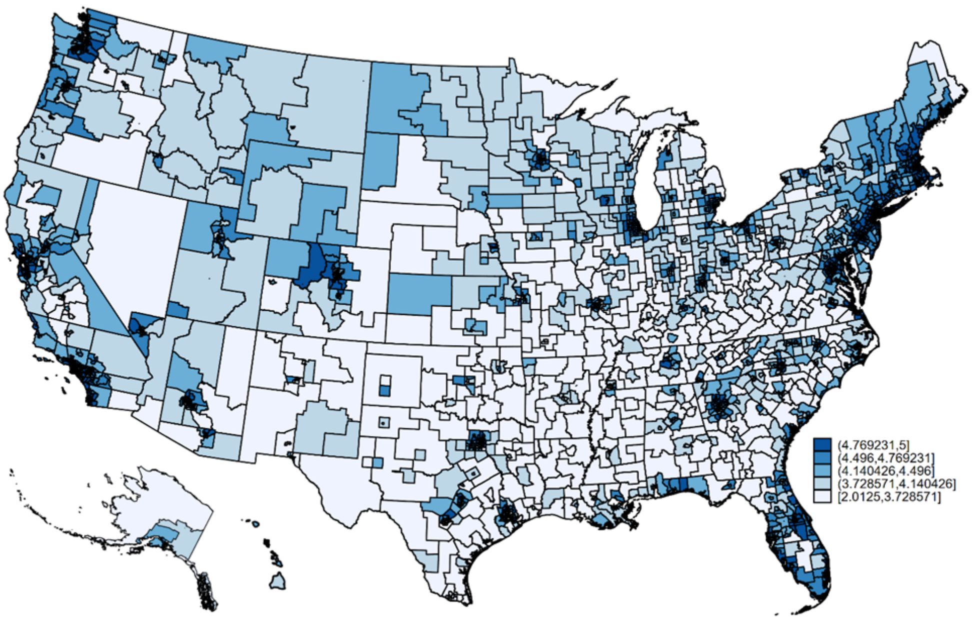
Source: Enverus.com. Non-Vertical Well Density is calculated as new non-vertical wells divided by total land area in 1000 square miles.

Figure 6: Mean Patent per 100K Population Year 2000-2015



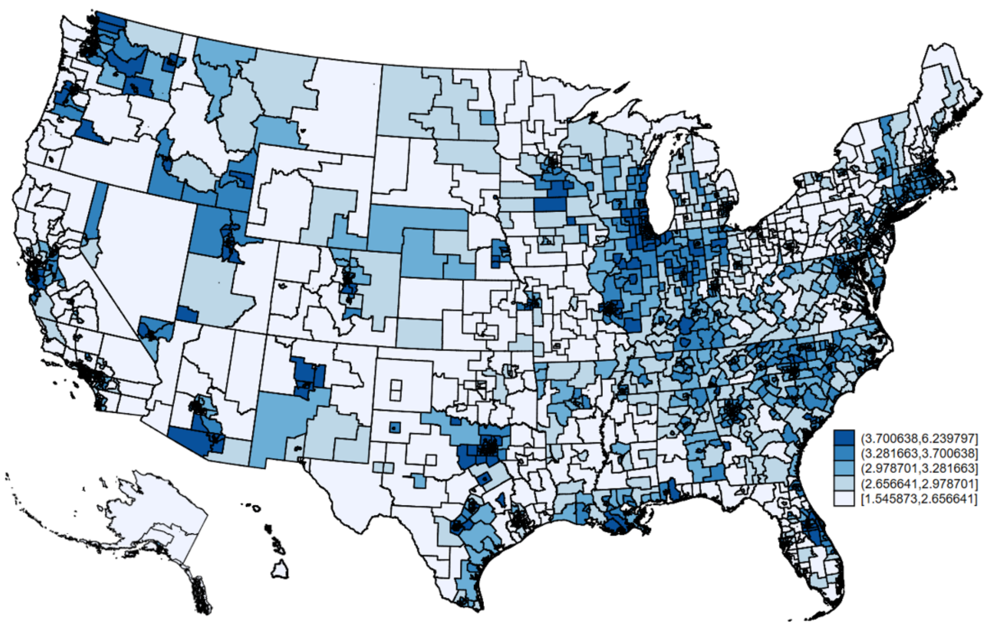
Source: U.S. Patent and Trademark Office (USPTO).

Figure 7: Mean Adoption to High-Speed Broadband 2013-2017



Source: Federal Communications Commission (FCC). FCC categorizes broadband connections per 1,000 households into six groups: zero if no connections, one if connections are greater than zero and lower than 200, two if connections are greater than 200 and lower than 400, three if connections are greater than 400 and lower than 600, four if connections are greater than 600 and lower than 800, five if connections are greater 800.

Figure 8: Mean Numbers of High-Speed Broadband Providers 2014-2016



Source: Federal Communications Commission (FCC). High-speed broadband providers are defined as number of providers that can or do provide internet access service at least 10 Mbps downstream and 1 Mbps upstream to at least one residential location.

Table 1: Means of Variables

	N	Mean	Std. Dev.	Min	Max
Dependent Variable					
Patent per 100K population (Patent Intensity)	11280	14.66	27.08	0.00	655.30
Main Treatment Variables					
Non-vertical wells per 1K square miles (Non-Vertical Well Density)	11280	3.31	18.01	0.00	436.90
Vertical wells per 1K square miles (Vertical Well Density)	11280	8.92	29.91	0.00	608.40
Control Variables					
Population index	11280	1.03	0.09	0.78	1.66
Predicted % share of population age between 18 and 64	11280	60.36	2.67	52.37	74.98
Predicted % share of population with a bachelor's degree or above	11280	14.40	4.80	5.63	39.65
Predicted % share of employed population	11280	46.15	4.60	29.35	66.20
Predicted % share of population in the labor force	11280	49.99	4.25	34.23	68.23
Predicted mean adult age	11280	47.63	2.14	37.90	57.11
Predicted employment outside mining industry index	11280	1.01	0.04	0.93	1.14
Average patent per 100K population 1990-1999	11280	10.27	11.09	0.00	122.90
Urban Hierarchy Variables					
Incremental Dist. to closest MSA with at least 250K Pop.	11280	94.82	95.11	0.00	569.80
Incremental Dist. to closest MSA with at least 500K Pop.	11280	25.63	43.16	0.00	294.90
Incremental Dist. to closest MSA with at least 1500K Pop.	11280	55.53	74.18	0.00	368.90
Incremental Dist. to closest MSA with at least 2000K Pop.	11280	28.48	52.99	0.00	228.20

Notes: The sample includes years 2000-2015. Analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. See Appendix for details on the construction of control variables.

Table 2: Main Full Sample and Border Contiguity Sample Regressions

	(1)	(2)	(3)
Panel A: Full Sample			
Non-Vertical Well Density	-0.0294***	-0.0286***	-0.0221***
	(0.0090)	(0.0081)	(0.0065)
Observations	11,280	11,280	11,280
Panel B: Border Contiguity			
Non-Vertical Well Density	-0.0306***	-0.0278***	-0.0229***
	(0.0095)	(0.0082)	(0.0060)
Observations	6,688	6,688	6,688
CZ FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes
CZ Linear Trend	No	Yes	No

Notes: The sample includes years 2000-2015. In Panel A, analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. In Panel B, analytical sample includes 418 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia based on border contiguity. Specifically, CZs with at least one non-vertical drilling well and their contiguous CZs are included in the sample; other CZs are excluded. Dependent variable is patent intensity, calculated as patents per 100K population. Non-Vertical Well Density is calculated as new non-vertical wells divided by total land area in 1000 square miles for each CZ. All the control variables in Table 1 are included in each column except average patent per 100K population 1990-1999 since it is absorbed by CZ fixed effects. Column (1) reports baseline two-way fixed effects regression; column (2) includes a CZ linear time trend; column (3) uses interactive fixed effects in equation 2. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table 3: First-Difference Regressions

	(1) 1-Year Diff.	(2) 2-Year Diff.	(3) 3-Year Diff.	(4) 4-Year Diff.	(5) 5-Year Diff.
Non-Vertical Well Density	-0.0217***	-0.0238***	-0.0289***	-0.0339***	-0.0292**
	(0.0059)	(0.0059)	(0.0090)	(0.0094)	(0.0123)
Incre. Dis. to MSA 250K	-0.0007	-0.0007	-0.0014	0.0002	-0.0003
	(0.0008)	(0.0018)	(0.0024)	(0.0041)	(0.0046)
Incre. Dis. to MSA 500K	-0.0054*	-0.0117**	-0.0155*	-0.0208*	-0.0262*
	(0.0031)	(0.0059)	(0.0088)	(0.0126)	(0.0151)
Incre. Dis. To MSA 1500K	-0.0014**	-0.0030**	-0.0038**	-0.0057*	-0.0086**
	(0.0007)	(0.0014)	(0.0019)	(0.0032)	(0.0034)
Incre. Dis. To MSA 2000K	-0.0012	-0.0025	-0.0028	-0.0043	-0.0047
	(0.0009)	(0.0019)	(0.0026)	(0.0041)	(0.0045)
Avg. Patent Intensity 90-99	0.0606*	0.1404*	0.1626*	0.2380*	0.2942*
	(0.0338)	(0.0735)	(0.0987)	(0.1388)	(0.1688)
Observations	10,575	5,640	3,525	2,820	2,115
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The sample includes years 2000-2015. Analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. Dependent variable is the first differencing patent intensity. Non-vertical well density is the first differencing new non-vertical well density. In each column, all the control variables and urban hierarchy variables are included in regressions but only certain variables are reported in the Table. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table 4: Instrumental Variables Regressions

	(1)	(2)
Panel A: First Stage Results		
	1.5035***	1.3318***
	(0.29301)	(0.34441)
	[26.33]	[14.95]
Panel B: Second Stage Results		
Non-Vertical Well Density	-0.0220*	-0.0252***
	(0.0112)	(0.0088)
Observations	11,280	11,280
CZ FE	Yes	Yes
Year FE	Yes	Yes
Region*Year FE	Yes	Yes
CZ Linear Trend	No	Yes

Notes: The sample includes years 2000-2015. Analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. Dependent variable is patent intensity, calculated as patents per 100K population. Panel A reports the first stage results and the F-stats are in square brackets. Panel B reports the second stage results, in which regress patent intensity on predicted non-vertical drilling well density. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at CZ-level.

Table 5: Descriptive Statistics

	N	Mean	Std. Error	Min	Max
Main dependent variables:					
Child mortality by age 1	400447	0.0866	0.281	0	1
Child mortality by age 3	400447	0.124	0.33	0	1
Child mortality by age 5	400447	0.131	0.338	0	1
Main treatment variables:					
Political Competition	400447	0.0303	0.0168	0	0.0708
Rainfall Deviation	400447	-0.00976	0.178	-1.852	1.083
Other control variables:					
Runoff	400447	0.0929	0.29	0	1
Incumbent	400447	0.65	0.477	0	1
Girl	400447	0.493	0.5	0	1
Birth order	400447	3.97	2.346	1	10
Oil	395550	66.98	44.82	24.49	195.9
Aluminium	395550	1703	436.4	1140	2640
Bananas	395550	544.1	142.2	373.9	975.9
Cocoa	395550	1601	526.8	903.9	3131
Coffee	395550	110.5	40.92	60.37	273.2
Cotton	395550	66.55	16.4	46.26	154.6
Groundnuts	395550	892.8	214.9	640.8	1724
Logs	395550	226.8	58.44	159.9	390.5
Maize	395550	121.4	37.51	88.22	291.8
Rice	395550	301.3	123.3	172.7	700.2
Tea	395550	217.4	38.84	164.2	346.2
Uranium	395550	23.21	23.48	8.285	99.24
Francophones	400447	0.531	0.499	0	1
Sahel	400447	0.293	0.455	0	1
West Africa	400447	0.591	0.492	0	1
East Afirca	400447	0.145	0.352	0	1
South Africa	400447	0.118	0.323	0	1
Country level regressions:					
GDP capita	301	553.3	409.1	111.9	2962
Health % GDP	167	5.49	2.462	2.491	11.79
Health capita	167	31.03	19.04	5.259	102.6
Gov. health % all health	167	25.03	10.22	5.053	47.4
Gov. health capita	167	7.312	5.007	1.126	28.67
Gov. health % all exp.	167	6.846	3.124	1.77	16.73

Notes: The sample is restricted to rural areas based on the DHS survey data.

Table 6: Effects of Political Competition on Child Mortality by Age 1

	(1)	(2)	(3)	(4)	(5)	(6)
Competition	-0.4859***	-0.3109***	-0.2265***	-0.2895***	-0.2815***	-0.2086***
	(0.0442)	(0.0449)	(0.0470)	(0.0442)	(0.0454)	(0.0477)
Girl		-0.0127***	-0.0127***		-0.0126***	-0.0127***
		(0.0010)	(0.0010)		(0.0010)	(0.0010)
Runoff		-0.0087***	-0.0115***		-0.0081***	-0.0112***
		(0.0026)	(0.0026)		(0.0026)	(0.0026)
Incumbent		-0.0069***	-0.0078***		-0.0059***	-0.0074***
		(0.0014)	(0.0014)		(0.0014)	(0.0014)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age one for all columns. The main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table 7: Effects of Political Competition on Child Mortality by Age 3

	(1)	(2)	(3)	(4)	(5)	(6)
Competition	-0.7308***	-0.3914***	-0.2592***	-0.3674***	-0.3522***	-0.2127***
	(0.0535)	(0.0530)	(0.0552)	(0.0521)	(0.0535)	(0.0560)
Girl		-0.0128***	-0.0128***		-0.0127***	-0.0128***
		(0.0012)	(0.0012)		(0.0012)	(0.0012)
Runoff		-0.0114***	-0.0172***		-0.0108***	-0.0170***
		(0.0030)	(0.0030)		(0.0030)	(0.0030)
Incumbent		-0.0095***	-0.0123***		-0.0080***	-0.0112***
		(0.0016)	(0.0016)		(0.0016)	(0.0016)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table 8: Effects of Political Competition on Child Mortality by Age 5

	(1)	(2)	(3)	(4)	(5)	(6)
Competition	-0.7976***	-0.4247***	-0.2809***	-0.3969***	-0.3879***	-0.2336***
	(0.0557)	(0.0546)	(0.0567)	(0.0538)	(0.0551)	(0.0576)
Girl		-0.0131***	-0.0131***		-0.0130***	-0.0131***
		(0.0012)	(0.0012)		(0.0012)	(0.0012)
Runoff		-0.0106***	-0.0172***		-0.0101***	-0.0171***
		(0.0031)	(0.0031)		(0.0031)	(0.0031)
Incumbent		-0.0098***	-0.0130***		-0.0082***	-0.0119***
		(0.0017)	(0.0016)		(0.0017)	(0.0017)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age five for all columns. The main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table 9: Effects of Rainfall on Child Mortality by Age 1

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall	-0.0057**	-0.0022	-0.0022	0.0004	-0.0010	0.0001
	(0.0028)	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0030)
Girl		-0.0127***	-0.0127***		-0.0126***	-0.0127***
		(0.0010)	(0.0010)		(0.0010)	(0.0010)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age one for all columns. The main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table 10: Effects of Rainfall on Child Mortality by Age 3

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall	-0.0175***	-0.0086**	-0.0080**	-0.0056	-0.0066*	-0.0045
	(0.0034)	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)
Girl		-0.0128***	-0.0128***		-0.0127***	-0.0128***
		(0.0012)	(0.0012)		(0.0012)	(0.0012)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at DHS cluster level.

Table 11: Effects of Rainfall on Child Mortality by Age 5

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall	-0.0194***	-0.0100***	-0.0091**	-0.0069*	-0.0079**	-0.0056
	(0.0035)	(0.0036)	(0.0036)	(0.0036)	(0.0036)	(0.0036)
Girl		-0.0131***	-0.0131***		-0.0130***	-0.0131***
		(0.0012)	(0.0012)		(0.0012)	(0.0012)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age five for all columns. The main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at DHS cluster level.

Table 12: Effects of Political Competition and Rainfall on Child Mortality by Age 1

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original Estimates						
Competition	-0.4851***	-0.3114***	-0.2276***	-0.2917***	-0.2823***	-0.2095***
	(0.0442)	(0.0449)	(0.0470)	(0.0442)	(0.0454)	(0.0478)
Rainfall	0.0064	0.0062	0.0041	0.0115*	0.0091	0.0070
	(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0062)	(0.0062)
Interaction	-0.3605**	-0.2588	-0.2096	-0.3292*	-0.3152*	-0.2355
	(0.1687)	(0.1690)	(0.1695)	(0.1687)	(0.1687)	(0.1692)
Girl		-0.0127***	-0.0127***		-0.0126***	-0.0127***
		(0.0010)	(0.0010)		(0.0010)	(0.0010)
Runoff		-0.0087***	-0.0115***		-0.0081***	-0.0112***
		(0.0026)	(0.0026)		(0.0026)	(0.0026)
Incumbent		-0.0069***	-0.0078***		-0.0059***	-0.0073***
		(0.0014)	(0.0014)		(0.0014)	(0.0014)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes
Panel B: Marginal Effects						
Competition	-0.4814***	-0.3082***	-0.2250***	-0.2882***	-0.2784***	-0.2066***
	(0.0442)	(0.0449)	(0.0470)	(0.0442)	(0.0454)	(0.0478)
Rainfall	-0.0044	-0.0016	-0.0022	0.0016	-0.0004	-0.0001
	(0.0029)	(0.0030)	(0.0030)	(0.0030)	(0.0030)	(0.0030)

Notes: The dependent variable is a dummy one if a child dies before age one for all columns. The first main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. The second main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. "Interaction" refers the interaction term between "Competition" and "Rainfall" variables. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table 13: Effects of Political Competition and Rainfall on Child Mortality by Age 3

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original Estimates						
Competition	-0.7265***	-0.3894***	-0.2599***	-0.3667***	-0.3505***	-0.2136***
	(0.0535)	(0.0530)	(0.0552)	(0.0521)	(0.0535)	(0.0560)
Rainfall	0.0017	0.0007	-0.0031	0.0084	0.0049	0.0011
	(0.0073)	(0.0072)	(0.0072)	(0.0072)	(0.0072)	(0.0072)
Interaction	-0.5775***	-0.2907	-0.1719	-0.4172**	-0.3612*	-0.2059
	(0.1977)	(0.1947)	(0.1953)	(0.1948)	(0.1943)	(0.1948)
Girl		-0.0128***	-0.0128***		-0.0127***	-0.0128***
		(0.0012)	(0.0012)		(0.0012)	(0.0012)
Runoff		-0.0114***	-0.0171***		-0.0108***	-0.0169***
		(0.0030)	(0.0030)		(0.0030)	(0.0030)
Incumbent		-0.0097***	-0.0124***		-0.0081***	-0.0113***
		(0.0016)	(0.0016)		(0.0016)	(0.0016)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes
Panel B: Marginal Effects						
Competition	-0.7204***	-0.3858***	-0.2578***	-0.3623***	-0.3461***	-0.2111***
	(0.0535)	(0.0530)	(0.0552)	(0.0521)	(0.0534)	(0.0560)
Rainfall	-0.0156***	-0.0080**	-0.0082**	-0.0042	-0.0059*	-0.0050
	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0035)	(0.0036)

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The first main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. The second main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. "Interaction" refers the interaction term between "Competition" and "Rainfall" variables. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table 14: Effects of Political Competition and Rainfall on Child Mortality by Age 5

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original Estimates						
Competition	-0.7927***	-0.4222***	-0.2815***	-0.3955***	-0.3856***	-0.2344***
	(0.0557)	(0.0545)	(0.0567)	(0.0538)	(0.0551)	(0.0576)
Rainfall	0.0011	-0.0002	-0.0046	0.0073	0.0035	-0.0009
	(0.0073)	(0.0073)	(0.0073)	(0.0073)	(0.0073)	(0.0073)
Interaction	-0.6116***	-0.3042	-0.1629	-0.4211**	-0.3593*	-0.1792
	(0.2012)	(0.1979)	(0.1983)	(0.1980)	(0.1976)	(0.1980)
Girl		-0.0131***	-0.0131***		-0.0130***	-0.0131***
		(0.0012)	(0.0012)		(0.0012)	(0.0012)
Runoff		-0.0105***	-0.0170***		-0.0101***	-0.0170***
		(0.0031)	(0.0031)		(0.0031)	(0.0031)
Incumbent		-0.0099***	-0.0132***		-0.0084***	-0.0120***
		(0.0017)	(0.0017)		(0.0017)	(0.0017)
Observations	393,573	387,528	387,528	393,573	387,528	387,528
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes
Panel B: Marginal Effects						
Competition	-0.7863***	-0.4184***	-0.2795***	-0.3911***	-0.3812***	-0.2322***
	(0.0557)	(0.0545)	(0.0567)	(0.0538)	(0.0551)	(0.0576)
Rainfall	-0.0173***	-0.0093***	-0.0094***	-0.0054	-0.0072**	-0.0062*
	(0.0035)	(0.0036)	(0.0036)	(0.0036)	(0.0036)	(0.0036)

Notes: The dependent variable is a dummy one if a child dies before age five for all columns. The first main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. The second main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. "Interaction" refers the interaction term between "Competition" and "Rainfall" variables. All sample is restricted to rural areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at DHS cluster level.

Table 15: Descriptive Statistics

Variable	Mean	Std. Error	Min	Max
Self-Employment	0.07	0.26	0	1
Work-From-Home	0.04	0.20	0	1
Work-From-Home and Self-Employment	0.02	0.13	0	1
Work-From-Home not Self-Employment	0.02	0.15	0	1
Self-Employment not Work-From-Home	0.06	0.23	0	1
Hours Worked per Week	27.83	19.03	0	99
Labor Force Participation	0.73	0.45	0	1
High-Speed Broadband (ACS)	0.80	0.40	0	1
Broadband (ACS)	0.94	0.24	0	1
High-Speed Broadband (FCC)	4.24	0.59	1.688	5
High-Speed Broadband Providers	3.17	0.85	1.124	7.476
Surface Topography Index	9.19	6.97	1	21
Age	43.51	10.20	18	59
Age Square	1998.00	865.10	324	3481
Bachelor Share	22.31	11.14	1.396	78.05
No schooling	0.01	0.10	0	1
Nursery School to Grade 4	0.00	0.06	0	1
Grade 5 to 8	0.02	0.14	0	1
Grade 9 to 12, No High School Diploma	0.05	0.21	0	1
High School Diploma	0.22	0.41	0	1
Collge Education, No Bachelor's Degree	0.31	0.46	0	1
Bachelor's Degree or Above	0.39	0.49	0	1
White, Not Hispanic	0.71	0.46	0	1
Black/African American, Not Hispanic	0.06	0.23	0	1
American Indian/Alaska Native, Not Hispanic	0.01	0.08	0	1
Asian, Not Hispanic	0.08	0.26	0	1
Other Race, Not Hispanic	0.02	0.13	0	1
Hispanic	0.14	0.35	0	1
Num. of Children	1.26	1.23	0	9
MSA	0.76	0.43	0	1

Notes: The sample includes years 2013-2017. Analytical sample includes 2,351 ACS Public Use Microdata Area (PUMA) covering the entire U.S. 50 States.

Table 16: Effects of Adoption to High-Speed Broadband on Self-Employment and Work-from-Home, Non-MSA and MSA Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: Non-MSA Sample					
High Speed Broadband	0.00702***	0.00818***	0.00397***	0.00421***	0.00305***
	(0.00102)	(0.00081)	(0.00058)	(0.00048)	(0.00076)
Age	0.00526***	0.00288***	0.00148***	0.00140***	0.00378***
	(0.00026)	(0.00029)	(0.00017)	(0.00019)	(0.00020)
Age Square	-0.00005***	-0.00003***	-0.00002***	-0.00002***	-0.00003***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00031	0.00027	0.00012	0.00016	-0.00043
	(0.00049)	(0.00025)	(0.00016)	(0.00023)	(0.00048)
Observations	542,932	542,932	542,932	542,932	542,932
Panel B: MSA Sample					
High Speed Broadband	0.00689***	0.00937***	0.00357***	0.00580***	0.00332***
	(0.00067)	(0.00078)	(0.00034)	(0.00064)	(0.00064)
Age	0.00634***	0.00407***	0.00148***	0.00259***	0.00486***
	(0.00024)	(0.00024)	(0.00010)	(0.00017)	(0.00023)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00009	0.00028*	0.00005	0.00024***	-0.00014
	(0.00018)	(0.00015)	(0.00012)	(0.00009)	(0.00014)
Observations	1,683,016	1,683,016	1,683,016	1,683,016	1,683,016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). High-Speed Broadband is based on "cihispeed" variable in the ACS, which assigns one when high-speed broadband is under subscription for a household. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table 17: Effects of Adoption to High-Speed Broadband on Self-Employment and Work-from-Home, with and without Children Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: with Children					
High Speed Broadband	0.00759***	0.00817***	0.00330***	0.00487***	0.00429***
	(0.00074)	(0.00068)	(0.00034)	(0.00048)	(0.00066)
Age	0.00593***	0.00324***	0.00113***	0.00211***	0.00480***
	(0.00029)	(0.00021)	(0.00009)	(0.00016)	(0.00027)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00002***	-0.00005***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00027	0.00023	0.00006	0.00017	-0.00033**
	(0.00020)	(0.00016)	(0.00010)	(0.00011)	(0.00016)
Observations	1,436,508	1,436,508	1,436,508	1,436,508	1,436,508
Panel B: without Children					
High Speed Broadband	0.00627***	0.01079***	0.00426***	0.00653***	0.00201**
	(0.00093)	(0.00076)	(0.00043)	(0.00061)	(0.00099)
Age	0.00582***	0.00393***	0.00135***	0.00258***	0.00447***
	(0.00042)	(0.00027)	(0.00016)	(0.00019)	(0.00034)
Age Square	-0.00005***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	0.00019	0.00040*	0.00003	0.00037**	0.00016
	(0.00024)	(0.00023)	(0.00014)	(0.00015)	(0.00019)
Observations	789,440	789,440	789,440	789,440	789,440
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). High-Speed Broadband is based on "cihispeed" variable in the ACS, which assigns one when high-speed broadband is under subscription for a household. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 with children for Panel A, and without children for Panel B. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table 18: Effects of Adoption to Broadband on Self-Employment and Work-from-Home, Non-MSA and MSA Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: Non-MSA Sample					
Broadband	0.01483***	0.01145***	0.00650***	0.00496***	0.00834***
	(0.00154)	(0.00094)	(0.00064)	(0.00053)	(0.00136)
Age	0.00520***	0.00288***	0.00147***	0.00141***	0.00373***
	(0.00026)	(0.00029)	(0.00017)	(0.00019)	(0.00020)
Age Square	-0.00005***	-0.00003***	-0.00001***	-0.00002***	-0.00003***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00030	0.00028	0.00012	0.00016	-0.00042
	(0.00049)	(0.00025)	(0.00016)	(0.00023)	(0.00048)
Observations	542,932	542,932	542,932	542,932	542,932
Panel B: MSA Sample					
Broadband	0.01558***	0.01208***	0.00537***	0.00672***	0.01021***
	(0.00120)	(0.00093)	(0.00059)	(0.00054)	(0.00100)
Age	0.00631***	0.00411***	0.00149***	0.00262***	0.00483***
	(0.00025)	(0.00024)	(0.00010)	(0.00018)	(0.00023)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00009	0.00029*	0.00005	0.00024***	-0.00013
	(0.00018)	(0.00016)	(0.00012)	(0.00009)	(0.00014)
Observations	1,683,016	1,683,016	1,683,016	1,683,016	1,683,016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). Broadband is based on "cinethh" variable in ACS, which assigns one when an individual has access to the internet. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

Table 19: Effects of Adoption to Broadband on Self-Employment and Work-from-Home, with and without Children Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: with Children					
Broadband	0.01622***	0.00973***	0.00495***	0.00478***	0.01126***
	(0.00124)	(0.00078)	(0.00060)	(0.00049)	(0.00110)
Age	0.00590***	0.00329***	0.00114***	0.00215***	0.00476***
	(0.00029)	(0.00021)	(0.00009)	(0.00016)	(0.00027)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00002***	-0.00005***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00026	0.00023	0.00006	0.00017	-0.00032**
	(0.00020)	(0.00016)	(0.00010)	(0.00011)	(0.00016)
Observations	1,436,508	1,436,508	1,436,508	1,436,508	1,436,508
Panel B: without Children					
Broadband	0.01563***	0.01529***	0.00668***	0.00861***	0.00895***
	(0.00158)	(0.00097)	(0.00059)	(0.00057)	(0.00146)
Age	0.00583***	0.00393***	0.00136***	0.00258***	0.00447***
	(0.00042)	(0.00027)	(0.00016)	(0.00019)	(0.00034)
Age Square	-0.00005***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	0.00020	0.00041*	0.00003	0.00038**	0.00017
	(0.00024)	(0.00023)	(0.00014)	(0.00015)	(0.00019)
Observations	789,440	789,440	789,440	789,440	789,440
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). Broadband is based on "cinethh" variable in ACS, which assigns one when an individual has access to the internet. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 with children for Panel A, and without children for Panel B. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table 20: Effects of FCC Adoption to High-Speed Broadband on Self-Employment and Work-from-Home, Non-MSA and MSA Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: Non-MSA Sample					
High Speed Broadband	0.01031***	0.00071	0.00162	-0.00091	0.00869***
	(0.00317)	(0.00311)	(0.00167)	(0.00236)	(0.00289)
Age	0.00536***	0.00300***	0.00154***	0.00146***	0.00382***
	(0.00026)	(0.00029)	(0.00017)	(0.00019)	(0.00020)
Age Square	-0.00005***	-0.00003***	-0.00002***	-0.00002***	-0.00003***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00033	0.00027	0.00011	0.00016	-0.00044
	(0.00050)	(0.00025)	(0.00016)	(0.00023)	(0.00048)
Observations	542,932	542,932	542,932	542,932	542,932
Panel B: MSA Sample					
High Speed Broadband	0.00296**	-0.00148	-0.00005	-0.00143	0.00302**
	(0.00143)	(0.00164)	(0.00072)	(0.00142)	(0.00141)
Age	0.00644***	0.00421***	0.00153***	0.00268***	0.00491***
	(0.00025)	(0.00024)	(0.00010)	(0.00018)	(0.00024)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00003***	-0.00005***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00009	0.00027*	0.00004	0.00023**	-0.00013
	(0.00018)	(0.00016)	(0.00012)	(0.00009)	(0.00014)
Observations	1,683,016	1,683,016	1,683,016	1,683,016	1,683,016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). High-Speed Broadband categorizes broadband connections per 1,000 households into six groups: zero if no connections, one if connections are greater than zero and lower than 200, two if connections are greater than 200 and lower than 400, three if connections are greater than 400 and lower than 600, four if connections are greater than 600 and lower than 800, five if connections are greater 800. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table 21: Effects of FCC Adoption to High-Speed Broadband on Self-Employment and Work-from-Home, with and without Children Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: with Children					
High Speed Broadband	0.00200	-0.00231	-0.00074	-0.00157	0.00274*
	(0.00141)	(0.00180)	(0.00101)	(0.00146)	(0.00149)
Age	0.00608***	0.00340***	0.00120***	0.00220***	0.00488***
	(0.00029)	(0.00021)	(0.00009)	(0.00016)	(0.00027)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00003***	-0.00005***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00027	0.00021	0.00006	0.00016	-0.00032**
	(0.00020)	(0.00016)	(0.00010)	(0.00011)	(0.00015)
Observations	1,436,508	1,436,508	1,436,508	1,436,508	1,436,508
Panel B: without Children					
High Speed Broadband	0.00739***	0.00053	0.00207**	-0.00154	0.00532**
	(0.00253)	(0.00163)	(0.00091)	(0.00144)	(0.00216)
Age	0.00581***	0.00392***	0.00135***	0.00257***	0.00446***
	(0.00042)	(0.00027)	(0.00016)	(0.00019)	(0.00034)
Age Square	-0.00005***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	0.00020	0.00039*	0.00003	0.00036**	0.00017
	(0.00024)	(0.00023)	(0.00014)	(0.00015)	(0.00019)
Observations	789,440	789,440	789,440	789,440	789,440
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). High-Speed Broadband categorizes broadband connections per 1,000 households into six groups: zero if no connections, one if connections are greater than zero and lower than 200, two if connections are greater than 200 and lower than 400, three if connections are greater than 400 and lower than 600, four if connections are greater than 600 and lower than 800, five if connections are greater 800. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 with children for Panel A, and without children for Panel B. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

Table 22: Effects of Broadband Providers on Self-Employment and Work-from-Home, Non-MSA and MSA sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: Non-MSA Sample					
Broadband Providers	0.00087	0.00193*	0.00018	0.00175**	0.00069
	(0.00105)	(0.00105)	(0.00087)	(0.00070)	(0.00105)
Age	0.00526***	0.00288***	0.00147***	0.00141***	0.00379***
	(0.00033)	(0.00039)	(0.00023)	(0.00026)	(0.00031)
Age Square	-0.00005***	-0.00003***	-0.00001***	-0.00002***	-0.00003***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00055	0.00022	0.00010	0.00012	-0.00064
	(0.00055)	(0.00030)	(0.00025)	(0.00025)	(0.00046)
Observations	324,736	324,736	324,736	324,736	324,736
Panel B: MSA Sample					
Broadband Providers	0.00193***	0.00376***	0.00107***	0.00269***	0.00086*
	(0.00064)	(0.00056)	(0.00031)	(0.00041)	(0.00048)
Age	0.00637***	0.00411***	0.00152***	0.00259***	0.00484***
	(0.00024)	(0.00024)	(0.00011)	(0.00019)	(0.00024)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	0.00007	0.00047**	0.00011	0.00036***	-0.00004
	(0.00021)	(0.00019)	(0.00015)	(0.00012)	(0.00011)
Observations	1,004,512	1,004,512	1,004,512	1,004,512	1,004,512
ACS Sample	2014-2016	2014-2016	2014-2016	2014-2016	2014-2016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	No	No	No	No	No

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). Broadband providers are number of providers that can or do provide internet access service at least 10 Mbps downstream and 1 Mbps upstream to at least one residential location. This variable is based on a census block FCC data, and it is aggregated to PUMA level using average values. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

Table 23: Effects of Broadband Providers on Self-Employment and Work-from-Home, with and without Children Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: with Children					
Broadband Providers	0.00178***	0.00312***	0.00061**	0.00250***	0.00117**
	(0.00051)	(0.00053)	(0.00029)	(0.00043)	(0.00046)
Age	0.00592***	0.00334***	0.00129***	0.00205***	0.00463***
	(0.00035)	(0.00023)	(0.00012)	(0.00018)	(0.00035)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00002***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	0.00004	0.00049**	0.00024	0.00025*	-0.00019
	(0.00023)	(0.00020)	(0.00015)	(0.00015)	(0.00017)
Observations	857,008	857,008	857,008	857,008	857,008
Panel B: without Children					
Broadband Providers	0.00182	0.00432***	0.00159***	0.00273***	0.00023
	(0.00142)	(0.00089)	(0.00049)	(0.00059)	(0.00119)
Age	0.00576***	0.00364***	0.00115***	0.00249***	0.00461***
	(0.00055)	(0.00028)	(0.00018)	(0.00021)	(0.00047)
Age Square	-0.00005***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00001)	(0.00000)	(0.00000)	(0.00000)	(0.00001)
Bachelor Share	-0.00011	0.00033	-0.00016	0.00048**	0.00004
	(0.00035)	(0.00025)	(0.00017)	(0.00019)	(0.00028)
Observations	472,240	472,240	472,240	472,240	472,240
ACS Sample	2014-2016	2014-2016	2014-2016	2014-2016	2014-2016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	No	No	No	No	No

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). Broadband providers are number of providers that can or do provide internet access service at least 10 Mbps downstream and 1 Mbps upstream to at least one residential location. This variable is based on a census block FCC data, and it is aggregated to PUMA level using average values. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 with children for Panel A, and without children for Panel B. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

Table 24: Effects of Broadband Providers on Self-Employment and Work-from-Home, Non-MSA and MSA sample, using IV

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: Non-MSA Sample					
Broadband Providers	0.00615*	0.00648***	0.00322*	0.00326**	0.00293
	(0.00352)	(0.00222)	(0.00178)	(0.00155)	(0.00269)
Age	0.00533***	0.00287***	0.00146***	0.00140***	0.00387***
	(0.00033)	(0.00040)	(0.00023)	(0.00026)	(0.00030)
Age Square	-0.00005***	-0.00003***	-0.00001***	-0.00002***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00092	-0.00003	-0.00007	0.00004	-0.00085*
	(0.00056)	(0.00027)	(0.00025)	(0.00025)	(0.00047)
Observations	321,751	321,751	321,751	321,751	321,751
First-Stage F-Stat	[19.81]	[19.81]	[19.81]	[19.81]	[19.81]
Panel B: MSA Sample					
Broadband Providers	0.00234**	0.00661***	0.00183***	0.00478***	0.00051
	(0.00104)	(0.00107)	(0.00040)	(0.00095)	(0.00089)
Age	0.00637***	0.00418***	0.00156***	0.00262***	0.00481***
	(0.00024)	(0.00025)	(0.00011)	(0.00020)	(0.00024)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	0.00001	0.00021	0.00002	0.00019	-0.00001
	(0.00022)	(0.00024)	(0.00016)	(0.00016)	(0.00012)
Observations	989,350	989,350	989,350	989,350	989,350
First-Stage F-Stat	[48.89]	[48.89]	[48.89]	[48.89]	[48.89]
ACS Sample	2014-2016	2014-2016	2014-2016	2014-2016	2014-2016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	No	No	No	No	No

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). Broadband providers are number of providers that can or do provide internet access service at least 10 Mbps downstream and 1 Mbps upstream to at least one residential location. This variable is based on a census block FCC data, as it is aggregated to PUMA level using average values, and it is instrumented by land surface terrain as topography index. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. First-stage f-statistics are reported in square brackets. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table 25: Effects of Broadband Providers on Self-Employment and Work-from-Home, with and without Children Sample, using IV

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: with Children					
Broadband Providers	0.00478***	0.00720***	0.00209***	0.00511***	0.00269**
	(0.00133)	(0.00136)	(0.00063)	(0.00100)	(0.00111)
Age	0.00599***	0.00344***	0.00135***	0.00209***	0.00465***
	(0.00035)	(0.00024)	(0.00012)	(0.00019)	(0.00034)
Age Square	-0.00006***	-0.00004***	-0.00001***	-0.00002***	-0.00004***
	(0.00000)	(0.00000)	(0.00000)	(0.00000)	(0.00000)
Bachelor Share	-0.00023	0.00013	0.00009	0.00004	-0.00033**
	(0.00024)	(0.00025)	(0.00016)	(0.00018)	(0.00016)
Observations	844,755	844,755	844,755	844,755	844,755
First-Stage F-Stat	[44.33]	[44.33]	[44.33]	[44.33]	[44.33]
Panel B: without Children					
Broadband Providers	-0.00072	0.00522***	0.00198**	0.00324***	-0.00270
	(0.00187)	(0.00101)	(0.00084)	(0.00066)	(0.00164)
Age	0.00572***	0.00362***	0.00115***	0.00248***	0.00458***
	(0.00056)	(0.00028)	(0.00018)	(0.00021)	(0.00047)
Age Square	-0.00005***	-0.00004***	-0.00001***	-0.00003***	-0.00004***
	(0.00001)	(0.00000)	(0.00000)	(0.00000)	(0.00001)
Bachelor Share	0.00003	0.00025	-0.00021	0.00046**	0.00024
	(0.00038)	(0.00026)	(0.00018)	(0.00018)	(0.00030)
Observations	466,346	466,346	466,346	466,346	466,346
First-Stage F-Stat	[42.70]	[42.70]	[42.70]	[42.70]	[42.70]
ACS Sample	2014-2016	2014-2016	2014-2016	2014-2016	2014-2016
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	No	No	No	No	No

Notes: The dependent variable is a dummy one for self-employed in column (1); the dependent variable is a dummy one for work-from-home in column (2); The dependent variable is a dummy one for work-from-home and self-employed in column (3); the dependent variable is a dummy one for work-from-home but not self-employed in column (4); the dependent variable is a dummy one for self-employed but not work-from-home in column (5). Broadband providers are number of providers that can or do provide internet access service at least 10 Mbps downstream and 1 Mbps upstream to at least one residential location. This variable is based on census block FCC data, as it is aggregated to PUMA level using average values, and it is instrumented by land surface terrain as topography index. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample restrict to married female age between 18 and 59 with children for Panel A, and without children for Panel B. First-stage f-statistics are reported in square brackets. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

APPENDIX

Chapter 1: Construction of Control Variables

Because the main dependent variable patent intensity and drilling well density are normalized by population and land area respectively, instead of controlling for the actually population from year 2000 to 2015, I create a population index. Population index is calculated as yearly population in 2000-2015 divided by year 2000 population. For example, commuting zone (CZ) code 100 has 10K population in 2000 and 10.1K population in 2001. Population index variable for CZ code 100 in 2000 is 1 and 1.01 in 2001.

Demographic control variables are obtained from the American Community Survey (ACS) via IPUMS (Ruggles et al., 2020). I use the year 2000 5% Census data and 2001-2015 1% ACS data. The smallest identified geographic areas are Public Use Microdata Area (PUMA). I use PUMA to CZ crosswalk files to aggregate PUMA level variables to CZ-level (Autor et al., 2019; Autor and Dorn, 2013). Based on the ACS, I generate percentage share of population age between 18 and 64, percentage share of population with a bachelor's degree or above, percentage share of employed population, percentage share of population in the labor force, and mean adult age as CZ-level demographic control variables (use notation j to represent each variable in equations below). All the variables are constructed as predicted shift-share type variables (Bartik, 1991; Goldsmith-Pinkham et al., 2018).[20] Following equation 11, 12, and 13, I conduct steps below to construct these predicted variables. First, I use year 2000 5% census data to generate initial value of all demographic variables at CZ-level. Second, I use year 2000 census along with 2001-2015 ACS data to generate yearly regional (excluding its own state) growth rate of all variables.[21] Finally, I multiply CZ-level initial value generated in the second step by the national growth rate generated in the third step to form predicted

[20]First, controlling actual values of these variables may not be ideal since these variables may be outcomes of the treatment variable which may cause the "bad control variables" issue. Second, PUMA information is not available in 2001-2004 ACS. Using shift-share type controls can tackle this problem. Survey weights are included in the construction process.

[21]The four US census regions are the Northeast, Midwest, South, and West. Following with footnote 10, since PUMA information is not available in 2001-2004 ACS, I exclude its own state when construct the yearly regional share values. For example, CZ code 100 is located in Tennessee and belongs to census region South. The yearly regional growth rates of CZ code 100 is based on census region South values excluding Tennessee values for each year in 2000-2015.

shift-share type controls.

$$PredictValue^{j}_{c,2001} = ActualValue^{j}_{c,2000} \times (1 + RegionalGrowth^{j}_{c,2001-2000}) \tag{11}$$

$$PredictValue^{j}_{c,2002} = PredictValue^{j}_{c,2001} \times (1 + RegionalGrowth^{j}_{c,2002-2001}) \tag{12}$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$PredictValue^{j}_{c,2015} = PredictValue^{j}_{c,2014} \times (1 + RegionalGrowth^{j}_{c,2015-2014}) \tag{13}$$

Similar to the ACS demographic shift-share type control variables, I construct a variable for predicted employment outside the mining industry index. First, I use year 2000 QCEW data to calculate the initial level of non-mining industry employment for each CZ. Second, I use year 2000 - 2015 QCEW data to generate yearly regional (excluding its own CZ) employment growth rate for non-mining industry employment. Third, I generate predicted non-mining industry employment for each CZ for year 2000-2015 based on the initial level in step one and the growth rates in step two. Finally, I divide each years' predicted non-mining industry employment by year 2000 non-mining industry employment to form an index. The variation of this predicted index variable comes from the initial non-mining industry employment in year 2000 and regional growth rate excluding its own value.

Table A1: Testing Lagging Well Density

Panel A: Include Lags Separately	(1)	(2)	(3)	(4)
Non-Vertical Well Density	-0.0294***			
	(0.0090)			
Non-Vertical Well Density (Yr-1)		-0.0258***		
		(0.0089)		
Non-Vertical Well Density (Yr-2)			-0.0212**	
			(0.0092)	
Non-Vertical Well Density (Yr-3)				-0.0143
				(0.0101)
Observations	11,280	11,280	11,280	11,280
Panel B: Include Lags Simultaneously				
Non-Vertical Well Density	-0.0294***	-0.0317***	-0.0306***	-0.0299***
	(0.0090)	(0.0114)	(0.0095)	(0.0092)
Non-Vertical Well Density (Yr-1)		0.0027	-0.0008	0.0000
		(0.0111)	(0.0078)	(0.0073)
Non-Vertical Well Density (Yr-2)			0.0033	-0.0029
			(0.0124)	(0.0097)
Non-Vertical Well Density (Yr-3)				0.0062
				(0.0139)
Observations	11,280	11,280	11,280	11,280
CZ FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes
CZ Linear Trend	No	No	No	No

Notes: The sample includes years 2000-2015. Analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. Dependent variable is patent intensity, calculated as patents per 100K population. Non-vertical well density is calculated as new non-vertical wells divided by total land area in 1000 square miles for each CZ. 1997-1999 drilling wells data are used for column two to four.Yr-1 refers to one year lag, Yr-2 refers to two-year lag, and Yr-3 refers to three-year lag. All the control variables in Table 1 are included in each column except average patent per 100K population 1990-1999 since it is absorbed by CZ fixed effects. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table A2: Sensitivity Analysis using Sub-Sample

Panel A	(1)	(2)
	Baseline	Exclude Top 1% Wells
Non-Vertical Well Density	-0.0294***	-0.0388***
	(0.0090)	(0.0142)
Observations	11,280	11,168
Panel B		
	Exclude Top 1% Patents	Exclude Top 1% Wells & Patents
Non-Vertical Well Density	-0.0300***	-0.0393***
	(0.0077)	(0.0123)
Observations	11,168	11,056
Panel C		
	Exclude NY	Exclude CA
Non-Vertical Well Density	-0.0252***	-0.0307***
	(0.0077)	(0.0083)
Observations	11,088	10,992
Panel D		
	Exclude NY & CA	Use 2000-2010 Sample
Non-Vertical Well Density	-0.0264***	-0.0222***
	(0.0070)	(0.0084)
Observations	10,800	7,755

Notes: Panel A Column (1) reports baseline two-way fixed model reported in Panel A column (1) of Table 2. Panel A column (2) excludes top 1 percent CZs with highest non-vertical drilling well density in 2000-2015. Panel B column (1) excludes top 1 percent CZs with highest patent intensity in 2000-2015. Panel B column (2) excludes top 1 percent CZs with highest non-vertical drilling wells and top 1 percent CZs with highest patent intensity in 2000-2015. Panel C column (1) excludes all CZs in New York. Panel C Column (2) excludes all CZs in California. Panel D column (1) excludes all CZs in New York and all CZs in California. Panel D column (2) restricts sample to year 2000-2010. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table A3: Testing CZs with MSAs and without MSAs

	(1)	(2)
Panel A: CZs with Metropolitan Areas		
Non-Vertical Well Density	-0.0516**	-0.0389*
	(0.0241)	(0.0226)
Observations	4,928	4,928
Panel B: CZs without Metropolitan Areas		
Non-Vertical Well Density	-0.0234***	-0.0257***
	(0.0082)	(0.0070)
Observations	6,352	6,352
CZ FE	Yes	Yes
Year FE	Yes	Yes
Region*Year FE	Yes	Yes
CZ Linear Trend	No	Yes

Notes: The sample includes years 2000-2015. Dependent variable is patent intensity, calculated as patents per 100K population. Non-vertical well density is calculated as new non-vertical wells divided by total land area in 1000 square miles for each CZ. All the control variables in Table 1 are included in each column except average patent per 100K population 1990-1999 since it is absorbed by CZ fixed effects. For Panel A, Analytical sample includes 308 CZs with metropolitan areas in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. For Panel B, Analytical sample includes 397 CZs without any metropolitan areas in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table A4: Using Vertical Drilling Wells as the Main Treatment Variable

	(1)	(2)
Panel A: Baseline Regressions		
Vertical Well Density	0.001	0.0006
	(0.0048)	(0.0040)
Observations	11,280	11,280
Panel B: Border Contiguity		
Vertical Well Density	0.0039	0.001
	(0.0048)	(0.0042)
Observations	6,688	6,688
CZ FE	Yes	Yes
Year FE	Yes	Yes
Region*Year FE	Yes	Yes
CZ Linear Trend	No	Yes

Notes: The sample includes years 2000-2015. Dependent variable is patent intensity, calculated as patents per 100K population. In Panel A, analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. In Panel B, analytical sample includes 418 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia based on border contiguity. Vertical well density is calculated as number of new vertical wells divided by total land area in 1000 square miles for each CZ. All the control variables in Table 1 are included in each column except average patent per 100K population 1990-1999 since it is absorbed by CZ fixed effects. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table A5: Testing Mechanisms through the Labor Market

	(1) Creative Class	(2) Service Class	(3) Working Class
Non-Vertical Well Density	-0.0029**	-0.0049	0.0078***
	(0.0013)	(0.0037)	(0.0029)
Observations	8,460	8,460	8,460
CZ FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes
Weight	Population 2000	Population 2000	Population 2000
ACS	2000 & 05-15	2000 & 05-15	2000 & 05-15

Notes: The sample includes years 2000 and year 2005-2015. Analytical sample includes 705 CZs in 48 lower U.S. states excluding Alaska, Hawaii, and District of Columbia. Dependent variable is share of the creative class workers in column (1), share of the service class workers in column (2), and share of the working class workers in column (3). Non-vertical well density is calculated as new non-vertical wells divided by total land area in 1000 square miles for each CZ. Creative class dummy = 1 if person reports an occupation of computer and mathematical; architecture and engineering; life, physical and social science; education, training and library; arts, design, entertainment, sports and media; management; business and financial operations; legal or health care practitioners and technical; = 0 otherwise. Service class dummy = 1 if person reports an occupation of health care support; food preparation and food-service related; building and grounds cleaning and maintenance; personal care and service; sales and related; office and administrative support; community and social services or protective services; =0 otherwise. Working class dummy = 1 if person reports an occupation of construction and extraction; installation, maintenance and repair; production or transportation and material moving; =0 otherwise. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at CZ-level.

Table A6: Effects of Political Competition on Child Mortality by Age 3 for Urban Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Competition	-0.3961***	-0.3772***	-0.2711***	-0.3120***	-0.3415***	-0.2552***
	(0.0751)	(0.0764)	(0.0818)	(0.0767)	(0.0780)	(0.0834)
Girl		-0.0168***	-0.0167***		-0.0167***	-0.0167***
		(0.0019)	(0.0019)		(0.0019)	(0.0019)
Runoff		-0.0115**	-0.0156***		-0.0092*	-0.0141***
		(0.0048)	(0.0048)		(0.0049)	(0.0049)
Incumbent		-0.0081***	-0.0104***		-0.0064***	-0.0094***
		(0.0024)	(0.0024)		(0.0025)	(0.0025)
Observations	131,151	128,331	128,331	131,151	128,331	128,331
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. All sample is restricted to urban areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table A7: Effects of Rainfall on Child Mortality by Age 3 for Urban Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Rainfall	-0.0031	0.0032	0.0042	0.0049	0.0040	0.0070
	(0.0049)	(0.0051)	(0.0052)	(0.0052)	(0.0052)	(0.0052)
Girl		-0.0168***	-0.0167***		-0.0167***	-0.0167***
		(0.0019)	(0.0019)		(0.0019)	(0.0019)
Observations	131,151	128,331	128,331	131,151	128,331	128,331
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. All sample is restricted to urban areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table A8: Effects of Political Competition and Rainfall on Child Mortality by Age 3 for Urban Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original Estimates						
Competition	-0.3985***	-0.3813***	-0.2715***	-0.3225***	-0.3478***	-0.2565***
	(0.0755)	(0.0766)	(0.0819)	(0.0772)	(0.0783)	(0.0835)
Rainfall	0.0056	0.0055	0.0019	0.0098	0.0087	0.0043
	(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0112)	(0.0112)
Interaction	-0.2237	-0.0217	0.0969	-0.0952	-0.0969	0.1108
	(0.3075)	(0.3074)	(0.3090)	(0.3076)	(0.3081)	(0.3092)
Girl		-0.0168***	-0.0167***		-0.0167***	-0.0167***
		(0.0019)	(0.0019)		(0.0019)	(0.0019)
Runoff		-0.0117**	-0.0159***		-0.0094*	-0.0144***
		(0.0048)	(0.0048)		(0.0049)	(0.0049)
Incumbent		-0.0079***	-0.0102***		-0.0062**	-0.0091***
		(0.0024)	(0.0024)		(0.0025)	(0.0025)
Observations	131,151	128,331	128,331	131,151	128,331	128,331
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	No	Yes	Yes	No	Yes	Yes
Region trend	No	No	Yes	No	No	Yes
Birth year FE	No	No	No	Yes	Yes	Yes
Panel B: Marginal Effects						
Competition	-0.3931***	-0.3807***	-0.2740***	-0.3202***	-0.3454***	-0.2593***
	(0.0753)	(0.0766)	(0.0819)	(0.0771)	(0.0784)	(0.0836)
Rainfall	-0.0013	0.0048	0.0048	0.0069	0.0058	0.0076
	(0.0049)	(0.0052)	(0.0052)	(0.0052)	(0.0053)	(0.0053)

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The first main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. The second main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. "Interaction" refers the interaction term between "Competition" and "Rainfall" variables. All sample is restricted to urban areas based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table A9: Effects of Political Competition and Rainfall on Child Mortality by Age 3 for Girls' Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original Estimates						
Competition	-0.1825**	-0.1239			-0.1828**	-0.1249
	(0.0897)	(0.0910)			(0.0897)	(0.0910)
Rainfall			-0.0113**	-0.0066	-0.0015	0.0054
			(0.0058)	(0.0058)	(0.0117)	(0.0118)
Interaction					-0.3302	-0.4115
					(0.3309)	(0.3309)
Observations	158,408	158,408	158,408	158,408	158,408	158,408
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Region trend	Yes	Yes	Yes	Yes	Yes	Yes
Birth year FE	No	Yes	No	Yes	No	Yes
Panel B: Marginal Effects						
Competition					-0.1788**	-0.1200
					(0.0897)	(0.0911)
Rainfall					-0.0113*	-0.0068
					(0.0058)	(0.0058)

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The first main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. The second main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. "Interaction" refers the interaction term between "Competition" and "Rainfall" variables. All sample is restricted to rural areas and girls based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at DHS cluster level.

Table A10: Effects of Political Competition and Rainfall on Child Mortality by Age 3 for Boys' Sample

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Original Estimates						
Competition	-0.2670***	-0.2271**			-0.2665***	-0.2271**
	(0.0931)	(0.0946)			(0.0931)	(0.0947)
Rainfall			-0.0066	-0.0048	-0.0111	-0.0088
			(0.0057)	(0.0058)	(0.0121)	(0.0122)
Interaction					0.1211	0.0959
					(0.3237)	(0.3250)
Observations	164,543	164,543	164,543	164,543	164,543	164,543
Mother FE	Yes	Yes	Yes	Yes	Yes	Yes
Control variables	Yes	Yes	Yes	Yes	Yes	Yes
Region trend	Yes	Yes	Yes	Yes	Yes	Yes
Birth year FE	No	Yes	No	Yes	No	Yes
Panel B: Marginal Effects						
Competition					-0.2680***	-0.2283**
					(0.0931)	(0.0946)
Rainfall					-0.0075	-0.0059
					(0.0058)	(0.0059)

Notes: The dependent variable is a dummy one if a child dies before age three for all columns. The first main independent variable "Competition" is defined as 100 minus the share of votes of the presidential election winner in percentage. The second main independent variable "Rainfall" is defined as rainfall at a DHS cluster level in logarithm minus its long-term average in logarithm. "Interaction" refers the interaction term between "Competition" and "Rainfall" variables. All sample is restricted to rural areas and boys based on the DHS survey. Control variables include girl dummy, birth order, runoff dummy, incumbent dummy, and international prices. Regional trend includes a linear time trend for French speaking regions, Sahel regions, West Africa, East Africa, and South Africa. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at DHS cluster level.

Table A11: Effects of Political Competition on GDP per Capita and Health Expenditure

	(1) GDP capita	(2) Health % GDP	(3) Health capita	(4) Gov. health % all health	(5) Gov. health capita	(6) Gov. health % all exp.
Competition	1,421.1	16.9**	307.6***	-101.8**	70.0**	-15.2
	(1,253.0)	(8.4)	(114.5)	(51.1)	(28.2)	(15.5)
Runoff	69.2	1.0***	7.3	-1.7	-0.3	-0.3
	(63.6)	(0.3)	(4.7)	(2.1)	(1.1)	(0.6)
Incumbent	60.4	0.3	8.3***	-0.9	1.5**	-0.1
	(41.4)	(0.2)	(2.9)	(1.3)	(0.7)	(0.4)
N	301	167	167	167	167	167
Country FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	No	No	No	No	No
Sample	1988-2013	2000-2013	2000-2013	2000-2013	2000-2013	2000-2013

Notes: All variables are at a country level. In column (1), the dependent variable is GDP per capita (current US$). In column (2), the dependent variable is current health expenditure (% of GDP). In column (3), the dependent variable is current health expenditure per capita (current US$). In column (4), the dependent variable is domestic general government health expenditure (% of current health expenditure). In column (5), the dependent variable is domestic general government health expenditure per capita (current US$). In column (6), the dependent variable is domestic general government health expenditure (% of general government expenditure).

Table A12: Effects of Adoption to High-Speed Broadband on Working Hours per Week, Non-MSA and MSA Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: Non-MSA Sample					
High Speed Broadband	0.14414	-0.75007**	-1.69483**	0.26409	0.30789
	(0.32152)	(0.35749)	(0.69574)	(0.49037)	(0.31503)
Age	1.31759***	1.03438***	1.13511***	0.88767***	1.29678***
	(0.11972)	(0.14803)	(0.28470)	(0.19123)	(0.13841)
Age Square	-0.01348***	-0.01029***	-0.01137***	-0.00847***	-0.01325***
	(0.00135)	(0.00180)	(0.00338)	(0.00225)	(0.00155)
Bachelor Share	-0.15918	-0.06447	-0.28768	0.10736	-0.16095
	(0.11600)	(0.14279)	(0.21410)	(0.17056)	(0.13400)
Observations	39,507	18,892	9,341	9,545	30,161
Panel B: MSA Sample					
High Speed Broadband	-0.05397	0.00332	-0.75802***	0.28651	-0.09256
	(0.17648)	(0.20014)	(0.28053)	(0.24523)	(0.21747)
Age	0.69584***	0.41673***	0.25096*	0.55445***	0.78946***
	(0.08092)	(0.06719)	(0.12935)	(0.08550)	(0.09166)
Age Square	-0.00603***	-0.00352***	-0.00112	-0.00495***	-0.00705***
	(0.00093)	(0.00078)	(0.00148)	(0.00103)	(0.00105)
Bachelor Share	-0.01218	0.05523	0.12869	-0.01155	-0.04525
	(0.04202)	(0.03912)	(0.09488)	(0.05312)	(0.04104)
Observations	125,829	72,804	28,876	43,904	96,931
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is average working hours per week. In column (1), the sample is restricted to self-employed; in column (2), the sample is restricted to work-from-home; in column (3), the sample is restricted to work-from-home and self-employed; in column (4), the sample is restricted to work-from-home but not self-employed; in column (5), the sample is restricted to self-employed but not work-from-home. High-Speed Broadband is based on "cihispeed" variable in the ACS, which assigns one when high-speed broadband is under subscription for a household. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample is based on married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

Table A13: Effects of Adoption to High-Speed Broadband on Working Hours per Week, with and without Children Sample

	(1) **SE**	(2) **WFH**	(3) **WFH and SE**	(4) **WFH not SE**	(5) **SE not WFH**
Panel A: with Children					
High Speed Broadband	0.13420	-0.22220	-1.14489***	0.29954	0.28145
	(0.21502)	(0.22240)	(0.39751)	(0.24269)	(0.25636)
Age	1.18158***	0.89737***	0.85266***	0.87535***	1.26645***
	(0.08174)	(0.07648)	(0.14631)	(0.08883)	(0.09689)
Age Square	-0.01076***	-0.00860***	-0.00765***	-0.00822***	-0.01151***
	(0.00094)	(0.00088)	(0.00164)	(0.00107)	(0.00111)
Bachelor Share	-0.04496	-0.00021	0.03160	-0.05724	-0.06680
	(0.04384)	(0.05020)	(0.09178)	(0.05611)	(0.04748)
Observations	105,427	60,188	24,892	35,213	80,467
Panel B: without Children					
High Speed Broadband	0.11837	-0.12794	-0.70966	0.55030*	0.02017
	(0.25229)	(0.22902)	(0.51689)	(0.31058)	(0.27391)
Age	1.20537***	1.13414***	0.79882***	1.28078***	1.28528***
	(0.08211)	(0.08430)	(0.20047)	(0.11226)	(0.08382)
Age Square	-0.01448***	-0.01341***	-0.00959***	-0.01466***	-0.01560***
	(0.00097)	(0.00098)	(0.00228)	(0.00130)	(0.00098)
Bachelor Share	-0.02266	0.12977*	0.21583	0.06015	-0.07008
	(0.07705)	(0.06591)	(0.15485)	(0.10011)	(0.08554)
Observations	59,907	31,475	13,085	18,125	46,619
PUMA FE	Yes	Yes	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The dependent variable is average working hours per week. In column (1), the sample is restricted to self-employed; in column (2), the sample is restricted to work-from-home; in column (3), the sample is restricted to work-from-home and self-employed; in column (4), the sample is restricted to work-from-home but not self-employed; in column (5), the sample is restricted to self-employed but not work-from-home. High-Speed Broadband is based on "cihispeed" variable in the ACS, which assigns one when high-speed broadband is under subscription for a household. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample is based on married female age between 18 and 59 with children for Panel A, and without children for Panel B. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table A14: Effects of Different Measures of Broadband on Labor Force Participation, Non-MSA and MSA sample

	(1) **LFP**	(2) **LFP**	(3) **LFP**
Panel A: Non-MSA Sample			
High Speed Broadband	0.03096***		
	(0.00242)		
Broadband		0.07391***	
		(0.00561)	
FCC Broadband			-0.00460
			(0.00723)
Age	0.02437***	0.02404***	0.02483***
	(0.00096)	(0.00095)	(0.00095)
Age Square	-0.00030***	-0.00029***	-0.00030***
	(0.00001)	(0.00001)	(0.00001)
Bachelor Share	-0.00086	-0.00081	-0.00088
	(0.00073)	(0.00073)	(0.00073)
Observations	542,932	542,932	542,932
Panel B: MSA Sample			
High Speed Broadband	0.03367***		
	(0.00154)		
Broadband		0.07525***	
		(0.00314)	
FCC Broadband			-0.00617**
			(0.00296)
Age	0.01678***	0.01665***	0.01728***
	(0.00069)	(0.00070)	(0.00070)
Age Square	-0.00020***	-0.00019***	-0.00020***
	(0.00001)	(0.00001)	(0.00001)
Bachelor Share	-0.00097***	-0.00093***	-0.00100***
	(0.00033)	(0.00034)	(0.00033)
Observations	1,683,016	1,683,016	1,683,016
PUMA FE	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes

Notes: The dependent variable is a dummy equals one if in the labor force. In column (1), the main treatment variable is High-Speed Broadband, which is based on "cihispeed" variable in the ACS, which assigns one when high-speed broadband is under subscription for a household. In column (2), the main treatment variable is Broadband, which is based on "cinethh" variable in ACS, which assigns one when an individual has access to the internet. In column (3), the main treatment variable is FCC adoption of high-speed broadband, which is defined in Table 6. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample is based on married female age between 18 and 59 and reside in non-MSA areas for Panel A, and MSA areas for Panel B. * p<0.1, ** p<0.05, and *** p<0.01 based on clustered standard errors at a state level.

Table A15: Effects of Different Measures of Broadband on Labor Force Participation, with and without Children Sample

	(1) LFP	(2) LFP	(3) LFP
Panel A: with Children			
High Speed Broadband	0.02740***		
	(0.00200)		
Broadband		0.06906***	
		(0.00406)	
FCC Broadband			-0.00763**
			(0.00358)
Age	0.02944***	0.02919***	0.02997***
	(0.00067)	(0.00067)	(0.00066)
Age Square	-0.00032***	-0.00031***	-0.00032***
	(0.00001)	(0.00001)	(0.00001)
Bachelor Share	-0.00081*	-0.00077*	-0.00085*
	(0.00043)	(0.00043)	(0.00043)
Observations	1,436,508	1,436,508	1,436,508
Panel B: without Children			
High Speed Broadband	0.04231***		
	(0.00182)		
Broadband		0.08574***	
		(0.00303)	
FCC Broadband			-0.00397
			(0.00508)
Age	0.02445***	0.02449***	0.02439***
	(0.00103)	(0.00102)	(0.00103)
Age Square	-0.00034***	-0.00034***	-0.00034***
	(0.00001)	(0.00001)	(0.00001)
Bachelor Share	-0.00147***	-0.00140***	-0.00150***
	(0.00040)	(0.00041)	(0.00039)
Observations	789,440	789,440	789,440
PUMA FE	Yes	Yes	Yes
Region*Year FE	Yes	Yes	Yes

Notes: The dependent variable is a dummy equals one if in the labor force. In column (1), the main treatment variable is High-Speed Broadband, which is based on "cihispeed" variable in the ACS, which assigns one when high-speed broadband is under subscription for a household. In column (2), the main treatment variable is Broadband, which is based on "cinethh" variable in ACS, which assigns one when an individual has access to the internet. In column (3), the main treatment variable is FCC adoption of high-speed broadband, which is defined in Table 6. Bachelor share is the share of population with a bachelor degree at PUMA level. Additional controls not reported in the table include dummies for education levels, dummies for field of study, and dummies for race groups. All sample is based on married female age between 18 and 59 with children for Panel A, and without children for Panel B. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

Table A16: First Stage Estimates of Table 24 and 25

	(1) Table 24, Panel A	(2) Table 24, Panel B	(3) Table 25, Panel A	(4) Table 25, Panel B
Broadband Providers				
Topography*Year 2014	-0.0414***	-0.0544***	-0.0523***	-0.0509***
	(0.0074)	(0.0056)	(0.0056)	(0.0055)
Topography*Year 2015	-0.0331***	-0.0506***	-0.0479***	-0.0456***
	(0.0072)	(0.0057)	(0.0060)	(0.0058)
F - Statistics	[19.81]	[48.89]	[44.33]	[42.70]

Notes: This table reports first-stage estimates of Table 10 and 11. Year 2016 is in the exclusion group when construct the IVs. F-statistics is reported in square brackets. * $p<0.1$, ** $p<0.05$, and *** $p<0.01$ based on clustered standard errors at a state level.

VITA

Luyi Han

Candidate for the Degree of

Doctor of Philosophy

Thesis: THREE ESSAYS ON REGIONAL AND DEVELOPMENT ECONOMICS

Major Field: Economics

Biographical:

Education:

Completed the requirements for the Doctor of Philosophy in Economics at Oklahoma State University, Stillwater, Oklahoma in May, 2021.

Completed the requirements for the Master of Science in Economics at Oklahoma State University, Stillwater, Oklahoma in August, 2019.

Completed the requirements for the Bachelor of Arts in Economics at University of Minnesota - Duluth, Duluth, Minnesota in May, 2016.

Experience:

Graduate Student Instructor, Introduction to Microeconomics
Summer 2019
Graduate Student Instructor, Money and Banking
Fall 2018 - Spring 2021

Graduate Teaching Assistant
August 2016 - May 2021

Graduate and Professional Student Government Association (GPSGA) Representative Spring 2018 - Spring 2019

Professional Memberships: Agricultural & Applied Economics Association (AAEA); American Economic Association (AEA); Southern Regional Science Association (SRSA); Regional Science Association International (RSAI)