10-12-2017

The Importance of Transportation, Broadband, and Intellectual Infrastructure for Entrepreneurship

James E. Prieger
Pepperdine University, james.prieger@pepperdine.edu

Heng Lu

Habi Zhang

The Importance of Transportation, Broadband, and Intellectual Infrastructure for Entrepreneurship

Heng Lu and Habi Zhang

under the advising of

James E. Prieger
Pepperdine University School of Public Policy[1]

October 12, 2017

This empirical study uses a unique panel dataset to investigate the link between regional entrepreneurship and infrastructure. This topic is vital for understanding the factors that facilitate entrepreneurship, yet it receives scant scholarly attention. It is of particular value to policy makers because entrepreneurship is crucial for economic growth. We therefore examine how broadband infrastructure (internet connectivity), intellectual infrastructure (human capital), and transportation infrastructure (roads, bridges, and intermodal facilities) affect the establishment of new businesses in the United States. We primarily focus on broadband infrastructure, which is the least explored of these factors in the literature. We find that all kinds of infrastructure help entrepreneurship, but especially intellectual and broadband infrastructure. The importance of infrastructure—particularly broadband—for entrepreneurship varies among industries, but is more important in innovative industries. When transportation and broadband infrastructure are both increasing, they provide an additional benefit to entrepreneurship. The impact of broadband on the startup rate is greater in rural areas and when unemployment is higher. These results may help policymakers understand which regional factors facilitate entrepreneurship.

[1] We thank the Charles Koch Foundation for funding the Pepperdine School of Public Policy Research Fellowships, the program under which this research was performed. We also thank Josh Amberg and Josh Arnold for editing assistance.

Introduction

Why is entrepreneurship important? The literature suggests that entrepreneurship is a major contributing factor to economic growth. Acs and Armington (2003) point out that entrepreneurship is positively correlated with long term regional employment growth. Acs, et al. (2009) show that small firms and entrepreneurship are crucial to economic growth, employment, and competitiveness. Wennekers & Thurik (1999) argue that economic agents (small firms, entrepreneurs, etc.) play a primary role in "linking the institutions at the micro level to the economic outcome at the macro level". The zeitgeist of the "soaring nineties" was fueled by a revival of entrepreneurship (Wennekers & Thurik, 1999). However, various measurements[2] all converge at one conclusion - that entrepreneurship has been declining in the United States during the period we study: 2000 to 2012. The Kauffman Foundation suggests that in the United States, the birth rate of startups plummeted more than 30 percent during the Great Recession. An unambiguous downward trend in the startup rate is reflected clearly in our dataset as well. Knowing the factors that encourage entrepreneurship is important for policymakers because economic growth often moves with entrepreneurship. Therefore, we aim to explore and quantify which factors promote entrepreneurship. We investigate how broadband networks, which we call *communications infrastructure*, human capital, which we term *intellectual infrastructure,* and networks of roads, bridges, and intermodal facilities (*transportation infrastructure*) affect the formation of new business establishments in the United States.

Acs & Szerb (2007) recommend that "middle-income countries should focus on increasing human capital, upgrading technology availability and promoting enterprise development". Our unique dataset addresses the effect of various kinds of infrastructure on entrepreneurship at the county level. Drawing theoretical inspiration from the literature review of Sternberg (2009), we explore the regional dimension of entrepreneurship with a focus on infrastructure as a driving force. We have three different sets of infrastructure in our data set, namely, digital infrastructure (i.e., broadband), transportation infrastructure (railroads, highways, and bridges) and intellectual

[2] Among the three measures of The Kauffman Index (an index measuring startup activity in the United States), Startup Density is the ratio of the number of new employer businesses divided by the total population (in 100,000s), new businesses are defined as employer firms less than one year old employing at least one person besides the owner. All industries are included on this measure. The national trends of Startup Density of the period from 1977 to 2012 provided by The Kauffman Foundation is in line with the time trend from our data set.

infrastructure (education levels and the number of inventors in a county). In addition, we include basic demographic and economic variables (e.g., how rural the county is and the unemployment rate). We are most interested in broadband infrastructure due to the lack of literature on this subject.. As Audretsch, Heger, & Veith (2014) point out, “The impact of broadband as a ‘‘new’’ infrastructure on economic growth has drawn particular attention in policy-making for more than 15 years”. We now have enough years of data on broadband to assess its importance as a driver of new business creation. In the next section we present and discuss several hypotheses about entrepreneurship. In section III we describe the data. The empirical exploration is in section IV, which contains the regression results. A final section concludes.

Entrepreneurship and infrastructure

Using our unique dataset, we wish to test the following hypotheses:

> ***H1***: All three types of infrastructure (broadband, intellectual, and transportation) facilitate entrepreneurship.

The more broadband providers competing to offer service and the higher the quality (speed) of the service, the higher the startup rate we expect to see. Audretsch et al. (2015) suggest that although startup activity is positively linked to infrastructure in general, broadband is more conducive to startup activity than are highways and railroads. Theories suggest that broadband “should serve to reduce barriers to startup in that it facilitates connectivity, interaction and the exchange of knowledge and ideas that potentially could fuel entrepreneurial ventures” (Audretsch et al., 2014). As one type of physical infrastructure, broadband is expected to reduce production cost, lower entrance barriers, and increase information accessibility. Digital infrastructure is typically composed of General Purpose Technologies such as broadband, which is greatly useful to businesses and can be employed to increase productivity (Bresnahan and Trajtenberg, 1995). Empirical evidence is also available, for example, Squicciarini (2007) suggests broadband usage within a firm is positively associated with expanding firms’ sales such as to export markets. Hagsten & Kotnik (2016) suggest there is a significant and positive relationship between the capacity to use information and communications technology (ICT)[3] and the exportation of small- and medium-sized firms in Europe. Dalla Pellegrina et al. (2016)

[3] For example, one of the measures they use is the proportion of employees with broadband access.

suggest banks tend to offer better lending conditions to small entrepreneurs who use ICT more extensively when conducting their business activities than to those that do not.

A large literature links human capital and entrepreneurship. This developed from earlier research linking individual education or training and economic value (e.g., Mincer, 1958; Schultz, 1961; Becker, 1964). Mincer first associated the unequal distribution of personal income with the investment in human capital. Schultz contended that “such investment in human capital (e.g., ‘mature students attending school’, ‘workers acquiring on-the-job training’) accounts for most of the impressive rise in the real earnings per worker”. Becker, in his book *Human Capital*, makes an analogy between investment in education and investment in equipment. Built upon these seminal theories, extensive research has explicitly investigated the connection of entrepreneurial alertness, activities, and success with human capital or prior knowledge (Ardichvili, Cardozo, & Ray, 2000; Alvarez & Barney, 2007). For example, one meta-analysis (Unger, Rauch Frese, & Rosenbusch, 2011) “integrates results from three decades of human capital research in entrepreneurship” and found “a significant but small relationship between human capital and success”. Alvarez & Barney (2007) suggest human capital is critical to entrepreneurship for both discovery and creativity. Besides the direct link between human capital and entrepreneurship, several peripheral questions have been explored. Bruns, Holland, Shepherd, & Wiklund (2008) reveal that human capital characteristics are conducive to financial aid: they found, “that the similarity between the loan officers’ human capital and the applicants’ human capital was a significant indicator of loan approval.” Shepherd, Ettenson, & Crouch (2000) identify educational capability as belonging to “the second tier of importance” to venture capitalists’ decision-making in determining whether a venture is profitable. Because of the importance of human capital for entrepreneurship, we include it in our dataset as intellectual infrastructure (i.e., education levels and the number of inventors in a county) and test its effect.

Infrastructure investment, especially physical infrastructure (e.g., roads, highways), has been found to “provide a significant return to manufacturing firms and augments productivity growth.” (Morrison & Schwartz 1996). Aschauer (1989), in an empirical paper, points out that “a ‘core’ infrastructure of streets, highways, airports, mass transit, sewers, water systems, etc. has most explanatory power for productivity.” Other studies have demonstrated that transportation infrastructure plays a major role in economic growth. (Canning & Pedroni 2008; D´emurger, 2000; Boopen, 2006; Liu & Hu, 2010; Banistera & Berechmanb, 2001). However, as Audretsch et al. (2015) point out, “virtually no study to date has considered the impact of (physical) infrastructure on entrepreneurship in the form of startup activity;" we evaluate what role transportation infrastructure plays in entrepreneurship, to fill the gap in the entrepreneurship literature.

H2: The importance of infrastructure—particularly broadband—for entrepreneurship varies among industries.

We expect that each type of infrastructure will vary in importance among industries. For example, transportation infrastructure should play a more important role in manufacturing, while as Audretsch et al. (2015) suggest, broadband will be particularly conducive to new business formation in high tech industries.

H3: There are synergies among transportation and broadband infrastructure in creating new businesses.

The alternative hypothesis is that transportation and broadband infrastructure are independent or negatively-related factors for new business formation.

As Armington & Acs (2002) point out in their study, there are spillovers among different types of infrastructure because entrepreneurship is a regional economic activity. We expect to find complementarities, in the sense that the marginal effect of broadband on new business formation increases with the availability and quality of transportation infrastructure, and vice versa. One reason might be the importance of logistics to product delivery, and the importance of digital infrastructure to logistics. Alternatively, the marginal effects of one type of infrastructure might be independent of other types, or even negatively related. The latter may happen if there is a substitutive effect between two types of infrastructure when both compete for the impact on startup rate.

H4: The impact of broadband on the startup rate is greater in rural areas.

Armington & Acs (2002) point out that there is regional disparity in rates of new firm formation, and we will investigate how this relates to and interacts with broadband infrastructure. If broadband helps rural areas overcome their traditional disadvantage of greater distance to input and output markets, by the broadband "death of distance" phenomenon (Cairncross, 1997), then the marginal impacts of broadband on the startup rate might be greater in rural areas. Conversely, if there are agglomeration economies or other entrepreneurial advantages to urban areas, then broadband might have a greater impact there.

When finding employment is difficult, an individual is more likely to start his own business. On the other hand, Sutaria & Hicks (2004) note that high regional unemployment and its resulting drag on disposable income in the area may reduce demand for local goods and services, and therefore reduce the incentive for entrepreneurs to begin local businesses. Thus the impact of the unemployment rate on entrepreneurship is indeterminate. Given our focus on infrastructure, we are also interested in how broadband availability changes the relationship between the unemployment rate and new business formation.

H5: The impact of broadband on the startup rate is greater when unemployment is higher.

If the primary effect of unemployment on new business formation is to push jobless workers into entrepreneurship, then an individual may be more likely to start his own business when broadband is available, because it reduces startup costs. If so, then the facilitative nature of broadband for entrepreneurship would be even more important in areas with depressed local economies.

Description of the data

This section discusses in detail the unique dataset we constructed from various sources.

A. Definition and Measures of Entrepreneurship

Over the past several decades, the literature has developed many definitions of entrepreneurship. Henderson, Low, & Weiler (2007) thought “entrepreneurs could decide when to be innovative, what innovation to adopt and how far to push the innovative changes on the firm”. Entrepreneurs are identified as a unique economic player with two simultaneous roles: decision maker (corporate manager) and risk bearer (corporate owner). However, Low (2008) aggregates three broad yet distinct attributes of entrepreneurship: Ownership or operation of a firm, risk and uncertainty bearing, and innovation or the reallocation of resources. She believed innovation was the most representative part of entrepreneurship. In this project, we adopt Low's definition of entrepreneurs and regard innovation as a crucial part of entrepreneurial activities—one which promotes the long run growth of regional entrepreneurship.

Many studies provide different measurements of entrepreneurship. There is no one measure that is clearly superior to others (Low, 2009). "The choice of certain measures of entrepreneurship is likely to influence the research results" and each measurement has its own strengths and weaknesses (Gartner and Shane, 1995). The three most commonly used measures of entrepreneurship are: 1) self employment, 2) the number of establishments, and 3) dynamic data on births of establishments or firms.

The self-employment rate is the most widely used measure of entrepreneurship in economic development research. (Iversen et al., 2008). It is easy to access and relatively cheap to use. However, it is recognized as an inefficient measure since it captures all types of small business activity without differentiation (Acs et al., 2008). which makes it This makes it hard to distinguish the quality of entrepreneurial activity (Munn, 2008).

The number of establishments is another commonly used measure of entrepreneurship. It is widely available and easy to compute across time and space. Also, it is relatively stable across time compared to self-employment (Gartner and Shane, 1995). The number of establishments is a good longitudinal measure of past entrepreneurship (Gartner and Shane, 1995; Low, 2008). However, it fails to capture the innovation and risk attributes of entrepreneurship, making it a weak indicator for current entrepreneurship development (Low, 2008).

Dynamic data include establishment (or firm) flows over a period, generally a year, and include births, deaths, churn, and even survival of employer establishments (Low, 2008). Dynamic data as a flow measure captures changes over a period of time (Iversen et al., 2008), and is less related to the stock of establishments (Low, 2008). The largest disadvantage of dynamic data is the relatively high cost compared to stock data. Also, many dynamic databases are unavailable to researchers without special access granted by the Census Bureau.

This project will use *establishment births per member of the labor force* to measure entrepreneurial activities. Data on establishment births are from the Statistics of U.S.

Businesses (SUSB) dynamic database. SUSB is “an annual series that provides national and subnational data on the distribution of economic data by enterprise size and industry”.[4] Establishment births are identified as “establishments that have zero employment in the first quarter of the initial year and positive employment in the first quarter of the subsequent year.” Establishment counts are actual, with no censoring or infusion of statistical noise. In SUSB, “an establishment is a single physical location where business is conducted or where services or industrial operations are performed.” SUSB data is the only publicly available county-level measure of new business formation in the United States. The data captures innovation and dynamic micro data at the county and 2-digit NAICS level which can be refined for industry-specific analysis. This is the most finely disaggregated level for establishment births available to the public.

The time trends of establishment births and establishment births per member of the labor force are in Figure 1. The spikes in years ending in 1 and 6 are a consequence of the greater coverage of establishments in the dynamic data during the quinquennial Economic Censuses. Time fixed effects are employed in all regressions to remove these spurious aspects of the trends.

B. Broadband Infrastructure Data

Broadband infrastructure data come from two sources: Federal Communications Commission (FCC) and the National Telecommunications and Information Administration (NTIA). The National Broadband Map (NBM) was the creation of NTIA in collaboration with the FCC. The broadband provider data available from 2000 to 2008 (which we term “the early years” for this variable) come from the FCC’s Form 477. The FCC 477 program provides us with two variables to measure broadband accessibility: the Broadband Provider and the Residential Fixed Access per 1000 Households. The original count in the source data is at the ZIP code level and includes all types of providers, terrestrial or satellite, fixed or mobile, offering speed greater than 200 kbps. The provenance of these data is described in Prieger and Connelly (2013).[5] The count of providers were aggregated across all ZIP codes (ZCTAs) in the county, using population-weighted allocation factors. The time trend for average provider counts is shown in Figure 2.

---

[4] See https://www.census.gov/programs-surveys/susb/about.html.

[5] In the public version of the data the counts from one to three providers are grouped (censored). We use the uncensored data, under permission from the FCC (as described in Prieger and Connolly (2013)).

The discontinuity in the trend is because of a minor definitional change in 2005 (which has minimal effect).[6]

Later waves of the FCC data include the residential fixed broadband connections per 1000 households (available from 2009 to 2012). Only fixed (non-mobile) lines are included in these data (including fixed terrestrial wireless connections). The data on residential fixed connections commingle the accessibility of broadband to households with household demand for the services. Since the source data are quantized, the regressor is a categorical variable. Category 0 is for fewer than 200 connections per 1000 households, category 1 is for between 200 and 400 connections, and so on up to category 4, which is for counties with more than 800 connections per 1000 households. Figure 2 shows that subscription to residential broadband increased steadily during these years.

The broadband quality and speed data from 2011 to 2012 are extracted from the NBM. Maximum advertised download speed categories and count of providers in different broadband types are recorded and weighted by population over Census blocks. The block-level provider counts are aggregated to county level averages using population-weighted allocation factors. Figure 2 shows that broadband providers offered 25 mbps speed relatively rarely during these years, but that the percentage increase from 2011 to 2012 was substantial.

C. Intellectual Infrastructure Data

Intellectual infrastructure data come from two sources. Educational attainment data come from U.S. Census Bureau, American Community Survey, from 2000 to 2015, estimated on the county level. Data on the stock of inventors come from USPTO and take into account the depreciation rate.[7]

Educational attainment in the population is measured by three variables: 1. *Education Bachelor*: Fraction of adult population with at least a four-year college degree. 2. *Education Higher Degree:* Fraction of adult population with a master's degree, professional degree, or PhD. 3. *Education PhD*: Fraction of adult population with a PhD. In this version of the paper we use the

[6] Before 2005, providers were not required to submit information to the FCC if they (i.e., the holding company) had fewer than 250 lines in the state.
[7] There are two depreciation rates: 10% p.a. and 25% p.a. in our data set. The 25% depreciation rate is calculated following the method presented by Pakes and Schankerman (1984).

second of these. Figure 3 shows that the fraction of adults with a higher degree has grown steadily over time.

The stock of inventors is defined as Log Inventor Stock per Capita, which measures the stock of individuals named as inventors on patents who live in the county.[8] This stock is depreciated at 10% or 25% per annum starting from the year the patent was granted (using utility patents[9] from 1976-2016). Inventors are counted once per patent, and an inventor on patents with multiple (n > 1) inventors gets only a 1/n share to add to the local count. Figure 3 shows that the inventor stock per capita declined from 2002 to 2008 and then rose slightly. This count of local inventors is intended to measure both the pool of potential entrepreneurs and the knowledge infrastructure in the area to help entrepreneurs accomplish their goals.

D. Transportation Infrastructure Data

Transportation infrastructure data came from the following sources: *Railway Mileage*[10] comes from the National Rail Network 1:100,000, Bureau of Transportation Statistics; *Highway Miles*[11] comes from National Highway Planning Network, Federal Highway Administration (FHWA), Washington D.C.; *Bridge Quality*[12] comes from the National Bridge Inventory database from the U.S. Dept. of Transportation, Federal Highway Administration; *Count of Intermodal Facilities*[13] comes from the Research and Innovative Technology Administration's Bureau of Transportation Statistics (RITA/BTS). Missing years of *highway miles* and *highway lane miles* data were

[8] The authors' calculations of this variable use data from the U.S. Patent Inventor Database (Lai, et al, 2011).

[9] A utility patent is issued in the U.S. for the invention of a new and useful process, machine, manufacture, or composition of matter, or a new and useful improvement thereof. Approximately 90% of the patents issued by the USPTO in recent years have been utility patents, which are also referred to as "patents for invention".

[10] Only trackage from Class I and Class II railroads (and not abandoned) is counted. Class I railroads are the major national companies (including Amtrak). Class II railroads are the regionals. Class III railroads are locals (and are not included in the mileage totals).

[11] Highway miles includes only roads that are part of the National Highway System OR in functional classes 1 (Rural Principal Arterial - Interstate), 2 (Rural Principal Arterial - Other), 11 (Urban Principal Arterial - Interstate), 12 (Urban Principal Arterial-Other Freeways & Expressways), or 14 (Urban Principal Arterial - Other), as designated in NHPN database.

[12] Bridge quality is measured in two ways: Percentage of bridges in acceptable condition out of all bridges in the county, expressed as a fraction (0-1 scale), weighted and unweighted; Or the average bridge quality for a county.

[13] An intermodal facility supports two or more modes of transportation. Possible types of facilities counted are: air & truck, port & truck, rail & port, rail & truck, truck & truck, truck - air - rail, truck - port - air, truck - port - rail - air, and truck - port -rail.

interpolated and extrapolated to fill in all years from 1999-2015.[14] Highway data are normalized in logarithmic per capita and per square mile form to adjust for skewness.[15] *Bridge Quality* is expressed as the percentage of the bridges in the county that are in "acceptable condition."[16] The bridge data is weighted by average daily traffic. *Railway Mileage* is normalized in logarithmic per capita and per square mile form to adjust the skewness. *Intermodal Facilities* is available only for 2002 and were never updated in any other year under observation.[17] Nevertheless, the number of facilities per capita in each county varies over time due to population growth. Intermodal facilities data are normalized in logarithmic per capita form to adjust for skewness. Figure 4 shows that there is little change over time in these measures. By far the greatest variation in these measures is between counties, not over time. The implications will be relevant for our fixed effects estimations below.

E. Summary Statistics

Table 1 displays the summary statistics in our dataset. All regressors are specific to the year and county but common across industries, so the summary statistics are virtually identical between the full sample and the innovative industries subsample—except for the dependent variable. Still, for establishment births, both samples share an almost identical spread and range. The majority of regressors are logged for skewness.

Figure 1 describes the time trend of establishment births within the period from 2000 to 2012. It displays a downward trend, despite the three spikes in 2001, 2006, and 2011, which are due to more complete enumeration of establishments during the quinquennial Economic Censuses. In general, Figure 2 depicts a steady, upward time trend for all four broadband variables; there is a rather sharp increase in the counts of providers offering high speed broadband (25 mbps) from 2011 to 2012; this discontinuity reflects the definitional change of the FCC's Form 477. Figure 3

[14] Except when too many years were missing, to be conservative. For example, if data for 2001 were missing for a location, then data for 1999-2002 were not extrapolated from the later available data. Similarly, if data for 2014 were missing, then values for 2012-2015 were not extrapolated from the earlier available data. Interpolation and extrapolation was linear except where extrapolated values would have been negative; in such cases exponential extrapolation was used.

[15] A mile of a highway or rail network is less useful if a) it must serve more people, or b) it is in a larger county (since it connects a smaller proportion of land out of all land in the county). Thus we normalize these regressors by dividing by both population and land area.

[16] Only bridges that are part of the Base Highway Network are included, to focus on the most important part of the road and bridge network.

[17] The National Transportation Atlas Database (NTAD) provides the latest information on highways, bridges, railroads and intermodal facilities at county scales.

shows a steady but minor upward trend of the fraction of adults with a higher degree and no clear trend for the local stock of inventors, which reached its lowest point in 2008. Finally, all four physical infrastructure variables remain almost constant over time.

Econometric model and results

A. Econometric specifications

Our econometric estimations of the importance of infrastructure for entrepreneurship employ pooled OLS, random effects, and fixed effects regressions using the panel data. In our panel data, the unit of observation is an industry (or group of industries) in a particular county. Our methodology is straightforward, except regarding the treatment of broadband regressors. As described above, the measures of broadband infrastructure and usage change over time. Instead of running separate regressions for the different time periods, which would give a confusingly large number of estimates, we combine all broadband regressors into a single regression spanning the entire period of the panel.

Let X_{jit} be a broadband measure available during the subset of years T_j, and let $1(L)$ be an indicator function taking value 1 if logical statement L is true and 0 if not. Then (ignoring the other regressors for the moment) our three main broadband variables enter the regression specification as

$$E(Y_{it}|X) = \alpha_t + \beta_1 1(t \in T_1) \times X_{1it} + \beta_2 1(t \in T_2) \times X_{2it} + \beta_3 1(t \in T_3) \times X_{3it} \ldots$$

This formulation allows the marginal effect of X_{jit}, which is $\gamma_j = \beta_j 1\left(t \in T_j\right)$, to switch between zero during years when the variable is irrelevant and β_{j} in years X_{jit} is available. It is easily estimable by defining new regressors $Z_{jit} = 1\left(t \in T_j\right) \times X_{jit}$ and using Z_{j} instead of X_{j} for the broadband regressors.

This formulation of the regression line requires that year fixed effects α_{t} be employed, since the mean values of the effective regressors Z_j change greatly between years in T_j and other years. Furthermore, as shown above, the different mean level of the establishment births every five years due solely to the measurement effect of the Economic Census also necessitates year fixed effects to wash out the (meaningless, for the present investigation) undercounting of establishments in the other years. Therefore, we include year fixed effects in all estimations, which effectively removes the aggregate secular trend from each regressor and the dependent variable. Note that year fixed effects also removes the impact of macroeconomic forces (such as the Great Recession) on entrepreneurship.

In many investigations of panel data, researchers employ unit-specific fixed effects to prevent bias from unobserved unit-specific factors that are both determinants of Y and correlated with the regressors. However, fixed effect estimators require that there be sufficient variation over time within the unit of observation to identify the coefficients after the data are demeaned for the within estimator. Our transportation infrastructure variables, however, vary little over time within a county. Since these regressors are important for our investigation, most of our estimations do not include unit fixed effects (i.e., "full" fixed effects).

Our framework for the various pooled, random effects, and fixed effects estimations is the following, where *i* indexes the county, *n* is for the industry group, *s* is for the state, and *t* is for the year:

$$Y_{int} = \alpha_t + I_n + S_s + \beta' X_{nt} + u_{in} + \varepsilon_{int}$$

The dependent variable is the log of establishment births per member of the labor force in county *n* and year *t*. In the rest of the expression, the α_t terms compose a set of time dummy variables as discussed above, I_n captures systematic differences among industries in rates of establishment births, S_s represents the state fixed effects that account for unobserved state-level factors affecting entrepreneurship (not included in all estimations), and ε is the idiosyncratic error term. Note that we have included no industry-specific regressors. The remaining term, u_{in}, will be treated either as a random effect to account for unit-specific unobserved heterogeneity and clustering or as a fixed effect (i.e., a random variable to be conditioned on), depending on the regression specification. In some estimations, α_t and S_s are replaced with state-year fixed effects, which account for all unobserved factors affecting new business formation within each state each year.

B. Results from pooled estimations

In this section we discuss our main set of results.

1. Results for all industries

The first set of estimations, shown in Table 2, are from pooled OLS models, where the only fixed effects included are those for time and industry.Identification in these models is based on variation within industry groups across time (but recalling that the data are detrended by the year fixed effects) and county. Estimation 1 in the first column of Table 2 is from data encompassing all industry groups available in SUSB. In this section we focus on the significance of the coefficients; we defer further exploration of the magnitude of the impacts of regressors until later.

There is strong evidence for hypothesis H1. All the broadband regressors have positive, highly significant coefficients. Where there are more broadband providers, whether measured at the ZIP code level at the 200 kbps standard in the early years or at the block level at the 25 mbps standard in the last years, there is a higher birth rate of new establishments. The log-log specification of these regressors implies that the coefficients are elasticities. Similarly, except for the highest category, the greater the number of residential broadband connections per household in a given county, the higher the establishment birth rate. These coefficients are relative to the omitted category for fewer than 20 connections per hundred households. The coefficients are monotonic, increasing until the highest category (for more than 80 connections per hundred households), which is lower than the previous category (between 60 and 80 connections per hundred households). The nonmonotonicity for the highest category may reflect that once there are 60 connections per hundred households, broadband is usually fully available within the area (since the household subscription rate for fixed broadband was about 60% during this time); demand for broadband higher than this average level may be driven by household factors that have little to do with entrepreneurship and new business formation. Similarly, intellectual infrastructure matters. The fraction of adults with higher degrees and the local stock of inventors per capita are positively and significantly associated with the establishment birth rate.

For transportation infrastructure, the evidence is mixed. Intermodal transportation facilities have the largest apparent impact, both in terms of the magnitude of the coefficient and the statistical significance. This regressor is the only one of the transportation infrastructure variables that remains significant in virtually every estimation. On the other hand, railroad track mileage never has a positive and significant association with the establishment birth rate. This presents a puzzle, since rail accounts for about 40% of the mile-tonnage of movement of freight in the nation[18] and it was found to be important for entrepreneurship in Germany (Audretsch et al, 2015). Highwayshave a positive and significant association with entrepreneurship. Bridge quality does not have a significant impact in this regression (but does in some later specifications). Thus, despite the mixed evidence for transportation infrastructure, the weight of the evidence is in favor of *H1*.

Regarding the other control variables, the higher the rural population the lower the establishment birth rate, and the higher the unemployment rate the lower the birth rate. The latter finding is not obvious; Armington & Acs (2002) suggest that "higher unemployment may deter start-ups in some sectors and increase them in others." The estimation explains about 39% of the variation in establishment births (adjusted R2 = 0.391); clearly there is significant local and temporal variation in new business formation that is driven by factors other than the included regressors. The patterns found in this estimation carry through in broad form to most subsequent estimations as well.

2. Results for the subsample of innovative industries

To more closely study entrepreneurship, we turn to a subsample of three industries that other researchers have found to be innovative compared to other industries.[19] Estimation 2 is similar to Estimation 1 except that the sample is restricted to these three industries. The sign and significance level of each coefficient is the same in Estimation 2 as before, except that the coefficient for railroad track mileage is now significant at the 5% level (but still negative). However, in the case of each regressor, the magnitude of the coefficient is larger for the innovative industry sample. Thus, the impacts of infrastructure on entrepreneurship in the all-industry sample were apparently diluted by relatively less-innovative industries. All remaining estimations use the subset of innovative industries.

To investigate hypothesis *H2*, we allow the impacts of broadband infrastructure on establishment births to vary by industry in Estimation 3 (also in Table 2). The number of broadband providers is interacted with industry dummy variables so that the main coefficients

[18] Statistic is for 2010. See the Federal Railroad Administration's statistics at https://www.fra.dot.gov/Page/P0362.
[19] For example, Low (Table 4.3, 2009) finds that the manufacturing industry is relatively high tech, high churn, and high patenting; the information industry is high tech and high churn; and the professional, scientific, and technical services industry is high skill, high tech, and high churn, all of which characteristics she identifies with innovative industries.

for the broadband variables are for the manufacturing industry and the increments to the coefficients for the other two industries are reported at the bottom of the table. Compared to the baseline impact of broadband for manufacturing births, the number of broadband providers has a higher elasticity for the information industry (whether the provider counts during the early or last years are examined). Given the intrinsic importance of moving information across space to the information industry, these results make sense. For the ProfSciTech services industry, the elasticity is higher in the early years but lower in the last years (again, relative to manufacturing). Whether the importance of high-speed broadband has become more important for manufacturing in recent years or less important for ProfSciTech services cannot be determined. It may also be the case that while ProfSciTech services are heavy users of broadband, relatively high-speed broadband (25 mbps or higher) is more important for manufacturing users' needs than for ProfSciTech services.

Examination of hypothesis H3 requires interacting the broadband and transportation infrastructure regressors, as shown in Estimation 4 in Table 3. Because the coefficients for residential broadband connections were nearly monotonic and to cut down on the number of interactions, that regressor is treated as a continuous variable rather than as a categorical variable in the estimations in Table 3. Here, if X is a broadband measure with coefficient β and γ_k is the coefficient on the interaction of the broadband and a transportation variable Z_k, then the marginal effect of X on log establishment births is $\beta+\Sigma_k\gamma_k Z_k$. Several of the coefficients on the interactions are not significant; here we focus only on those that are significant at the conventional 5% level. The positive coefficient on the interaction of the broadband provider count in the early years and intermodal transportation facilities, 0.011, implies that the positive marginal effect on establishment births of each of these variables individually is even greater when larger amounts of the other variable are present. For example, when intermodal facilities are at their first quartile, the elasticity of establishment births with respect to the early provider count is 0.210, while it is 0.225 when facilities are at their third quartile.[20]The same positive association holds for residential broadband connections and intermodal facilities. Doing the same thought experiment for broadband connections, the elasticity rises from 0.212 to 0.263. In other words, these results provide evidence for *H3* that there are synergies in entrepreneurship

[20] The marginal effect is a function of all transportation variables interacted with the broadband measure; the calculation is averaged over the actual values of the other transportation variables in the sample. Since many counties have no intermodal facilities, its quartiles used in the calculation are for the distribution conditional on having at least one facility.

between these particular types of communications infrastructure and transportation infrastructure.

The opposite finding holds for the interaction between broadband providers in the later years and bridge quality: the coefficient is negative (-0.066), implying that the marginal effect of the provider count decreases when bridges are higher quality. The elasticity with respect to the number of providers is 0.044 at the first quartile of bridge quality and 0.028 at the third quartile. It is unclear why this single measure of transportation infrastructure would lower the impact of broadband when the others do the opposite. To investigate this further, we investigated the net effect of all the transportation variables together on the marginal effect of the log broadband providers in the later years. Figure 5 shows how the calculated elasticity changes when calculated at different percentiles of the transportation variables. While the marginal effect is decreasing over much of the range, a horizontal line would fit inside the confidence band, indicating that the impact of this broadband regressor is not estimated precisely in this regression. We conclude that any evidence against *H3* from this= interaction coefficient is weak at best.

Turning to hypothesis *H4*, we find in Estimation 5 that the importance of broadband infrastructure--as measured by connections per household in earlier years, and by broadband providers in later years--for new business creation is higher the more rural the county is. Here, if δ is the coefficient on the interaction term $X \times RuralPop$, then the marginal effect of broadband measure X on log establishment births is $\beta + \delta \times RuralPop$**.** The estimated coefficients are relatively small because the regressor for rural population (*RuralPop*) is measured in percentage points, but they are highly statistically significant. For fixed connections, the marginal effect is 0.187 when *RuralPop* is at its first quartile and 0.283 when *RuralPop* is at its third quartile. For the count of providers in the latter years, the elasticity is 0.025 when *RuralPop* is at its first quartile and 0.047 when it is at its third quartile. Relative to the magnitude of the coefficients, the marginal effects are highly sensitive to the rurality of the county.

We test the final hypothesis, H5, in Estimation 6 by interacting the unemployment rate (UErate) with the broadband variables (shown in Table 3). Here, if X is a broadband measure with coefficient β and ζ is the coefficient on the interaction of the variables, then the marginal effect of X on log establishment births is $\beta + \zeta \times UErate$. The results indicate that connections per household and broadband providers in the latter years have higher positive marginal effects on establishment births when the unemployment rate is higher.[21] For fixed connections, the marginal effect is 0.080 when unemployment is at its first quartile and 0.180 when it is at its third quartile. For the count of providers in the latter years, the elasticity is 0.015 when unemployment is at its first quartile and 0.031 when it is at its third quartile. Thus the impact of broadband measured with these two variables varies greatly (in percentage terms, at least) with the jobless rate. The coefficient on the unemployment interaction for the provider count in earlier years is also positive, but it is not significant.

C. Results from random and fixed effects estimations

To investigate whether omitted variables of various sorts are creating bias in the regression estimates, we turn now to specifications that account for additional heterogeneity. Estimation 7 in Table 4 repeats Estimation 3 for innovative industries but with the addition of state fixed effects. Comparison of the two estimations shows that the differences are slight. The broadband and intellectual infrastructure regressors retain the same level of significance but are a bit smaller. There are some greater differences with the transportation infrastructure variables. In particular, the difficult-to-explain negative coefficient for railroad track mileage is no longer significant. The only other major difference is that unemployment coefficient loses significance. In Estimation 8, the state fixed effects are allowed to vary by year. The estimates are highly similar to those from Estimation 7, except that the broadband coefficients increase a small amount. The fact that estimates didn't change much when controlling for state fixed effects suggests that the state specific attributes, which were constant over time and unobserved in our dataset, didn't contribute much to the startup rate. For example, demographic characteristics seem to have little impact on entrepreneurship.

Estimation 9 in Table 4 is the first to account for unobserved heterogeneity at the county-industry level, with random effects. The addition of the random effects attenuates the broadband coefficients but changes little else (in particular, the significance levels do not change much). Estimation 10 adds county-industry fixed effects. The broadband coefficients are attenuated further, and their significance level drops in some cases (as is typical when relying only on variation within the unit of observation to identify the coefficients). One change is made to the specification, since the broadband provider count in later years was found to best enter in

[21] When interpreting this result it is important to recall that the year fixed effects already remove national trends in unemployment, and so this finding is not directly a consequence of the great recession or its aftermath.

quadratic terms. The impact of this regressor on establishment births is still positive on average, with an average elasticity of .046 (s.e.=.022, $p = 0.029$). The transportation regressors lose all significance, which is to be expected since they vary little over time within a county.

D. The magnitude of the impacts on establishment births

In much of the discussion above we focused on the statistical significance of the coefficients to indicate which factors affected entrepreneurship. In this section we turn to the different question of how much these factors matter for new establishment births. While the magnitudes of many of the coefficients can be interpreted as elasticities (when the regressor is in logs) or semi-elasticities (when the regressor is in levels) of establishment births per member of the labor force, here we present results in terms of the number of new establishments created across the nation. For each of the statistically significant coefficients in Table 4, we consider a discrete change in the associated regressor and its predicted impact on establishment births. The particular change in the regressor varies.

For the number of broadband providers in the early years (see the first lines of Table 5), the first change considered is from 1 provider to 2 providers (on average in the ZIP codes in the county). These figures are converted to the log form of the regressor and the change in the predicted value of the dependent variable is calculated.[22] Since the dependent variable is logged, this change is approximately the percentage change in the number of establishment births per member of the labor force, or, equivalently, in the number of births alone. The percentage change is converted to an absolute number of births first by applying it to either the median number of establishment births (which is two per industry-county in the innovative industries subsample) and aggregating across industries and counties to arrive at a national figure. For example, for the change in broadband providers of moving from monopoly to duopoly on average in the counties' ZIP codes, the prediction is 1,799 new establishment births nationally (in a single year).[23] To gain perspective on this number, recognize that about 1,300 establishments would represent 1% of new establishment births in the innovative industries on average during our sample. The number of additional births is also calculated a second way, by applying the estimated percentage change to the average number of establishment births

[22] The predicted values are calculated holding all other regressors at their actual values except the regressor under consideration, and then averaging the individual predictions across the sample.
[23] The standard errors for the predictions are shown in parentheses in the table. All are small enough that the predictions are significant at the 1% level.

(which is 14.2 per industry-county in the innovative industries subsample) and aggregating. By this method, increase from broadband monopoly to duopoly creates an additional 12,801 establishment births per year. The next lines of the table show the impact of a second possible change in broadband providers: from 2 to 3 providers. Given the log form of this regressor, the impact on establishment births is not as large: 1,052 births (calculated at median births) or 7,488 births (calculated at mean births).

For the impact of fixed broadband connections per household, the first change considered is from category 0 (less than 20%) to category 1 (between 20% and 40%). The impact on the number of additional establishment births is large: 10,311 (at the median) or 73,383 (at the mean). Moving from the first category to the third (60% to 80%, the second most common category after category 2) creates about 72% as many new births.

The number of broadband providers in the last years has a roughly similar impact on new business formation as in the early years (although the two measures have differing speed thresholds). An increase of one additional provider from the median (0.07)[24] leads to 1,768 (at the median) or 12,583 (at the mean) new establishments. Increasing by another additional provider in the second change creates about 24% as many births.

Intellectual capital is similarly important. Increasing the proportion of people with higher degrees or the local stock of inventors from quartile to quartile creates about 3,000 (at median births) to 21,000 (at mean births) new establishments. For transportation infrastructure, increasing from no intermodal facilities (the median) to one facility in the county (or, equivalently, 8.2 facilities per million people) is associated with the creation of 3,150 to 22,422 establishments. The largest impacts of all come from the rural proportion of the population. Hypothetically switching the nation from completely urban to half rural is associated with the loss of 19,404 to 138,105 establishment births. Switching from completely urban to completely rural would double those figures.

E. Robustness checks

[24] That is to say, 25 mbps broadband was relatively scarce during 2011 to 2012.

In addition to the various specifications estimated above, we tried alternate versions of several of the variables. These included alternative measures of broadband in later years (where, with the advent of the National Broadband Map many new measures become available) including other speed thresholds. Alternative measures of intellectual capital, such as the fraction of the population with PhDs and the inventor stock recalculated with 10% depreciation, generally yielded similar results. The transportation measures can be defined in terms of per capita, per square mile, or both, with little change in significance levels in any particular regression. Bridge quality can be measured different ways and weighted differently to aggregate to the county level. Few of these alternative forms of the regressors appear to make any significant difference in any particular regression.

Discussion and conclusions

As Audretsch et al. (2015) points out, there is an omission in the entrepreneurship literature, which has yet to link infrastructure to new firm formation. This study, therefore, is aimed at shedding light on the overlooked areas and on some of the most ignored factors.

Using a comprehensive data set, this study plays a pioneering role in the investigation of the relationship between infrastructures of various sorts, broadband infrastructure in particular, and new firm formation in the United States. This study can claim several advantages over previous studies, including the unique construction of the panel data that provides us with a large base of observations (N = 740,533 for the full sample and 117,002 for the innovative industries subsample), and a more representative measurement of key variables than previous studies. In addition, our dependent variable *Establishment Births per Member of the Labor Force* captures the innovative aspect of entrepreneurship that Low (2008) deems as crucial, and it addresses startup success in addition to startup activity. Our measures of broadband infrastructure and of transportation infrastructure take into account not only the availability of different types of infrastructure, but the quality as well. Our two measures of intellectual infrastructure (educational attainment and the stock of inventors in a given region) cohere with the human capital theory from which we draw our inspiration. Many of these measures correlate closely with our dependent variable to help account for startup success.

Our findings are strongly supportive of hypothesis 1, except for transportation infrastructure. Here, the evidence is mixed: in the case of intermodal transportation facilities and highway track mileage it is positive and significant, but for bridge quality it is negative but insignificant; these findings run contrary to previous results from Germany in the case of railroad track mileage (negative but insignificant), so this matter warrants further investigation. Hypothesis 2 is also supported by our findings. By interacting broadband provider regressors with two innovative industry dummy variables, we demonstrate that broadband infrastructure is more conducive to entrepreneurship in innovative industries than entrepreneurship in general. Though we do not have a set of unified evidence to support hypothesis 3, there is also no unequivocal evidence to reject it at this point. It appears that whether there is a complimentary or substitutive effect between transportation and broadband infrastructure in new business formation may be case-specific. Hypothesis 4 is corroborated by the results, and so is hypothesis 5: the impacts of broadband on startup rates are greater in rural areas and where the is more unemployment.

Our findings have important implications for policy makers. It is clear that all three types of infrastructure facilitate entrepreneurship. The positive impact on new firm formation of the three kinds of infrastructure is large. We predict that there will be 1,799 new establishment births per year (against the median number) or 12,801 establishment births per year (against the mean number), when the broadband providers move from monopoly to duopoly on average in the counties' ZIP codes, for instance. We have also demonstrated that broadband infrastructure has an incremental positive impact on establishment births in rural areas or when unemployment rate is higher. Therefore, we recommend that policymakers should consider policies that promote infrastructure (perhaps by removing regulatory barriers to investment), particularly in rural areas, and sensible private investment in certain types of infrastructure ought to be encouraged.

References

Acs, Z. J., & Armington, C. (2003). *Endogenous growth and entrepreneurial activity in cities.* Washington, DC: Center for Economic Studies, Bureau of the Census.

Acs, Z. J., & Szerb, L. (2007). Entrepreneurship, economic growth and public policy. *Small Business Economics,* 28(2-3), 109-122.

Acs, Z. J., Parsons, W., & Tracy, S. (2008). *High-impact firms: Gazelles revisited.* Washington, DC: Center for Economic Studies, Bureau of the Census.

Alvarez, S. A., & Barney, J. B. (2007). Discovery and creation: Alternative theories of entrepreneurial action. *Strategic Entrepreneurship Journal*, 1(1-2), 11-26.

Ardichvili, A., & Cardozo, R. N. (2000). A model of the entrepreneurial opportunity recognition process. *Journal of Enterprising Culture*, *8*(02), 103-119.
Armington, C., & Acs, Z. J. (2002). The determinants of regional variation in new firm formation. *Regional Studies*, 36(1), 33-45.

Aschauer, D. A. (1989). Is public expenditure productive? *Journal of Monetary Economics*, *23*(2), 177-200.

Audretsch, D. B. (2009). *Entrepreneurship, growth and public policy*. AL REPORT, 18.

Audretsch, D. B., Heger, D., & Veith, T. (2015). Infrastructure and entrepreneurship. *Small Business Economics*, 44(2), 219-230.

Banister, D., & Berechman, Y. (2001). Transport investment and the promotion of economic growth. *Journal of Transport Geography*, 9(3), 209-218.

Boopen, S. (2006). Transport infrastructure and economic growth: evidence from Africa using dynamic panel estimates. *The Empirical Economics letters*, *5*(1), 37-52.

Bruns, V., Holland, D. V., Shepherd, D. A., & Wiklund, J. (2008). The role of human capital in loan officers' decision policies. *Entrepreneurship Theory and Practice*, *32*(3), 485-506.

Bresnahan, T. F. and M. Trajtenberg (1995). "General purpose technologies: 'Engines of growth'?" *Journal of Econometrics*, vol. 65, pp. 83–108.

Cairncross, F. (1997). *The Death of Distance*. Harvard Business School Press. Boston, MA.

Canning, D., & Pedroni, P. (2008). Infrastructure, long-run economic growth and causality tests for cointegrated panels. *The Manchester School*, 76(5), 504-527.

Demurger, S. (2001). Infrastructure development and economic growth: an explanation for regional disparities in China?. *Journal of Comparative Economics*, *29*(1), 95-117.

Gartner, W. B., & Shane, S. A. (1995). Measuring entrepreneurship over time. *Journal of Business Venturing*, 10(4), 283-301.

Hagsten, E., & Kotnik, P. (2014). ICT as facilitator of internationalisation in small-and medium-sized firms. *Small Business Economics*, 1-16.

Henderson, J., Low, S., & Weiler, S. (2007). The drivers of regional entrepreneurship in rural and metro areas. *Entrepreneurship and Local Economic Development*, 81-102.

Iversen, J., Jorgensen, R., & Malchow-Moller, N. (2008). 3 Measuring entrepreneurship. *Foundations and Trends in Entrepreneurship*, 4(1), 17-17.

Lai, R., D'Amour, A., Yu, A., Sun, Y., & Fleming, L. (2011), Disambiguation and Co-authorship

Networks of the U.S. Patent Inventor Database (1975 - 2010", hdl:1902.1/15705, Harvard Dataverse, V5; invpat.zip.

Liu, S. L., & Hu, A. G. (2010). Transport Infrastructure and Economic Growth: Perspective from China's Regional Disparities [J]. *China Industrial Economics*, 4, 14-23.

Low, S. A. (2009). *Defining and measuring entrepreneurship for regional research: A new approach* (Doctoral dissertation). University of Illinois at Urbana-Champaign.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4), 281-302.
Morrison, C. J., & Schwartz, A. E. (1996). Public infrastructure, private input demand, and economic performance in New England manufacturing. *Journal of Business & Economic Statistics*, 14(1), 91-101.

Pakes, A., & Schankerman, M. (1984). The rate of obsolescence of patents, research gestation lags, and the private rate of return to research resources. In R&D, patents, and productivity (pp. 73-88). University of Chicago Press.

Prieger, J. E., & Connolly, M. (2013). A basic analysis of entry and exit in the US broadband market, 2005-2008. *Review of Network Economics*, 12(3), 229-270.

Schultz, T. W. (1961). Investment in human capital. *The American Economic Review*, 51(1), 1-17.

Shepherd, D. A., Ettenson, R., & Crouch, A. (2000). New venture strategy and profitability: A venture capitalist's assessment. *Journal of Business Venturing*, *15*(5), 449-467.

Squicciarini, M. (2017). Entrepreneurship, innovation and enterprise dynamics. *Small Business Economics* 48: 273. doi:10.1007/s11187-016-9784-z

Sternberg, R. (2009). Regional dimensions of entrepreneurship. *Foundations and Trends in Entrepreneurship*, 5(4), 211-340.

Sutaria, V., & Hicks, D. A. (2004). New firm formation: Dynamics and determinants. *The Annals of Regional Science*, 38(2), 241-262.

Unger, J. M., Rauch, A., Frese, M., & Rosenbusch, N. (2011). Human capital and entrepreneurial success: A meta-analytical review. *Journal of Business Venturing*, *26*(3), 341-358.

Wennekers, S., & Thurik, R. (1999). Linking entrepreneurship and economic growth. *Small Business Economic*s, 13(1), 27-56.

Tables

Table 1: Summary statistics of the data

Variable	Mean	Std. Dev.	Min	Max
Log(Establishments Births/Labor Force)				
Full sample	-3.72	3.36	-8.58	3.19
Innovative industries subsample	-3.76	3.24	-8.58	3.19
Log BB providers (200 kbps), 2001-2008	0.73	1.47	-12.60	3.00
Residential broadband connections/ household, 2009-2012	0.63	1.07	0.00	4.00
Log BB providers (25 mbps), 2011-2012	-0.73	2.55	-10.78	1.35
Education: % with a higher degree	0.06	0.04	0.00	0.44
Log inventor stock per capita (25% depreciation p.a.)	-8.58	1.92	-19.21	-1.37
Log intermodal transportation facilities per capita	-13.60	2.02	-14.86	-7.23
Log railroad track mileage per capita/ per square mile	-15.82	4.48	-23.56	-7.85
Log highway lane miles per capita/ per square mile	-12.59	4.33	-39.58	-7.23
% BHN bridges in acceptable condition	0.80	0.21	0.00	1.00
% rural population	58.61	30.57	0.00	100.00
Unemployment rate	6.38	2.80	1.12	28.86

Notes: *N* = 740,533 for the full sample and 117,002 for the innovative industries subsample. An observation in the regressions pertains to a county-year-industry triple; all regressors are county-year specific and so the summary statistics are virtually identical between the two estimation samples except for the dependent variable.

Table 2: Pooled OLS regressions

Y = log(Establishments Births/Labor Force)	Estimation 1 All Industries	Estimation 2 Innovative Inds.	Estimation 3 Innovative Inds.
Log BB providers (200 kbps), 2001-2008	0.137 (0.008)***	0.175 (0.013)***	0.144 (0.015)***
Residential BB connections/ HH, 2009-2012: 0.2 to 0.4	0.388 (0.087)***	0.597 (0.143)***	0.598 (0.143)***
Residential BB connections/ HH, 2009-2012: 0.4 to 0.6	0.653 (0.087)***	0.874 (0.141)***	0.875 (0.141)***
Residential BB connections/ HH, 2009-2012: 0.6 to 0.8	0.780 (0.089)***	1.043 (0.144)***	1.044 (0.144)***
Residential BB connections/ HH, 2009-2012: ≥ 0.8	0.687 (0.100)***	0.783 (0.163)***	0.783 (0.163)***
Log BB providers (25 mbps), 2011-2012	0.016 (0.003)***	0.035 (0.006)***	0.042 (0.008)***
Education: Higher Degree	6.984 (0.555)***	12.357 (0.845)***	12.359 (0.845)***
Log Inventor stock per cap. (25% depr.)	0.170 (0.008)***	0.236 (0.011)***	0.236 (0.011)***
Log intermodal transportn. facilities per cap.	0.033 (0.007)***	0.041 (0.011)***	0.041 (0.011)***
Log RR track mi. per cap. per sq. mi	-0.004 (0.003)	-0.010 (0.005)**	-0.010 (0.005)**
Log highway lane miles per cap. per sq. mi	0.015 (0.004)***	0.015 (0.005)***	0.015 (0.005)***
% BHN bridges in acceptable condition	-0.031 (0.070)	-0.064 (0.100)	-0.064 (0.099)
% rural population	-0.017 (0.001)***	-0.022 (0.001)***	-0.022 (0.001)***
Unemployment rate	0.056 (0.006)***	0.088 (0.008)***	0.088 (0.008)***
(Log BB providers, 200 kbps) × (Information industry)			0.037 (0.016)**
(Log BB providers, 200 kbps) × (Prof., Sci., & Tech Svcs ind.)			0.056 (0.016)***
(Log BB providers, 25 mbps) × (Information industry)			0.017 (0.008)**
(Log BB providers, 25 mbps) × (Prof., Sci., & Tech Svcs ind.)			-0.037 (0.008)***
SER	2.63	2.73	2.73
R^2	0.391	0.290	0.291
Adjusted R2	0.391	0.290	0.290
N	740,533	117,002	117,002

* p<0.1; ** p<0.05; *** p<0.01. All estimations include year and industry fixed effects. S.e.'s account for clustering at the county-industry level. The "innovative industries" are manufacturing, information, and professional, scientific, and technical services. The excluded category for *Residential BB connections/HH* is "zero to 0.2". The coefficients for the interacted terms in Estimation 3 are the increments to the base slope for the BB variable, where the base slope is for the manufacturing industry.

Table 3: OLS regressions with broadband interactions

Y = log(Establishments Births/Labor Force)	Estimation 4 BB × Transport. interactions	Estimation 5 BB × Rural interactions	Estimation 6 BB × UE interactions
Log BB providers (200 kbps), 2001-2008	0.315 (0.094)***	0.181 (0.035)***	0.163 (0.024)***
Residential BB connections/HH, 2009-2012	0.126 (0.100)	0.122 (0.030)***	-0.043 (0.045)
Log BB providers (25 mbps), 2011-2012	0.121 (0.039)***	0.010 (0.011)	-0.005 (0.013)
(Log BB providers, 200 kbps) × (log IM trans. facilities/cap.)	0.011 (0.006)**		
(Log BB providers, 200 kbps) × (log RR track mi/cap/sqmi)	0.001 (0.002)		
(Log BB providers, 200 kbps) × (Log hiway lane mi/cap/sqmi)	0.001 (0.001)		
(Log BB providers, 200 kbps) × (% acceptable cond'n bridge)	0.067 (0.035)*		
(Res. BB connections/HH) × (log IM trans. facilities/cap.)	0.010 (0.005)**		
(Res. BB connections/HH) × (log RR track mi/cap/sqmi)	-0.005 (0.003)*		
(Res. BB conns/HH) × (Log highway lane mi/cap/sqmi)	-0.004 (0.004)		
(Res. BB connections/HH) × (% acceptable cond'n bridge)	0.089 (0.061)		
(log BB providers, 25 mbps) × (log IM trans. facilities/cap.)	0.003 (0.002)		
(Log BB providers, 25 mbps) × (log RR track mi/cap/sqmi)	-0.000 (0.001)		
(Log BB providers, 25 mbps) × (log hiway lane mi/cap/sqmi)	-0.000 (0.001)		
(Log BB providers, 25 mbps) × (% acceptable cond'n bridge)	-0.066 (0.021)***		
(Log BB providers, 200 kbps) × (% rural population)		-0.0001 (0.0004)	
(Res. BB connections/HH) × (% rural population)		0.0018 (0.0004)***	
(Log BB providers, 25 mbps) × (% rural population)		0.0004 (0.0001)***	
(Log BB providers, 200 kbps) × (UE rate)			0.003 (0.005)
(Res. BB connections/HH) × (UE rate)			0.028 (0.005)***
(Log BB providers, 25 mbps) × (UE rate)			0.005 (0.001)***
SER	2.94	2.73	2.73
R^2	0.178	0.290	0.290
Adjusted R2	0.178	0.290	0.290

* *p*<0.1; ** *p*<0.05; *** *p*<0.01. *N* = 117,002.

All estimations include year and industry fixed effects. S.e.'s account for clustering at the county-industry level.

Table 4: Fixed and random effects regressions

Y = log(Establishments Births/Labor Force)	Estimation 7 OLS: state FE	Estimation 8 OLS: state-year FE	Estimation 9 Random effects	Estimation 10 Fixed effects
Log BB providers (200 kbps), 2001-2008	0.130 (0.012)***	0.144 (0.013)***	0.055 (0.010)***	0.028 (0.018)
Log BB providers (200 kbps), squared				0.003 (0.002)*
Residential BB connections/HH, 2009-2012: 0.2 to 0.4	0.571 (0.141)***	0.573 (0.144)***	0.345 (0.120)***	0.229 (0.096)**
Residential BB connections/HH, 2009-2012: 0.4 to 0.6	0.849 (0.139)***	0.863 (0.144)***	0.392 (0.118)***	0.213 (0.095)**
Residential BB connections/HH, 2009-2012: 0.6 to 0.8	0.961 (0.141)***	0.983 (0.147)***	0.422 (0.120)***	0.258 (0.099)***
Residential BB connections/HH, 2009-2012: more than 0.8	0.691 (0.158)***	0.730 (0.166)***	0.294 (0.131)**	0.222 (0.118)*
Log BB providers (25 mbps), 2011-2012	0.025 (0.005)***	0.036 (0.006)***	0.013 (0.005)**	0.054 (0.025)**
Log BB providers (25 mbps), squared				0.005 (0.002)**
Education: Higher Degree	10.717 (0.831)***	10.673 (0.846)***	11.874 (0.601)***	3.515 (1.376)**
Log Inventor stock per cap. (25% depr.)	0.208 (0.012)***	0.208 (0.012)***	0.162 (0.009)***	0.008 (0.011)
Log intermodal transportn. facilities per cap.	0.056 (0.011)***	0.056 (0.011)***	0.068 (0.008)***	-0.491 (0.309)
Log RR track mi. per cap. per sq. mi	-0.002 (0.005)	-0.002 (0.005)	0.001 (0.003)	-0.005 (0.007)
Log highway lane miles per cap. per sq. mi	0.010 (0.005)*	0.010 (0.005)*	0.012 (0.004)***	0.022 (0.035)
% BHN bridges in acceptable condition	0.197 (0.103)*	0.197 (0.104)*	0.196 (0.071)***	0.093 (0.086)
% rural population	-0.022 (0.001)***	-0.022 (0.001)***	-0.022 (0.001)***	0.001 (0.002)
Unemployment rate	0.010 (0.009)	0.014 (0.011)	0.004 (0.008)	-0.001 (0.007)
Year-state fixed effects	no	yes	yes	no
County-industry fixed effects	no	no	no	yes
SER	2.70	2.70	2.46	2.44
R^2	0.306	0.311	.	0.013†
Adjusted R^2	0.306	0.307	.	-0.072†

* p<0.1; ** p<0.05; *** p<0.01. † Statistic calculated from the demeaned data; not comparable to similar statistic from other regressions.

N = 117,002. All estimations include year, state, and industry fixed effects. S.e.'s account for clustering at the county-industry level. The random and full fixed effects models take county-industry to be the unit of observation.

Table 5: Impact on national establishment births of various discrete changes in regressors

Regressor that changes	Description of the discrete changes	establish't births, calculated at median births (s.e.)	establish't births, calculated at mean births (s.e.)
BB providers (200 kbps), 2001-2008			
	1st change: from 1 provider to 2 providers.	1,799 (160)	12,801 (1,141)
	2nd change: from 2 provider to 3 providers.	1,052 (94)	7,488 (667)
Residential BB connections/HH, 2009-2012			
	1st change: From category 0 (< 0.2) to category 1 (0.2 to 0.4).	10,311 (2,596)	73,383 (18,479)
	2nd change: from category 1 (0.2 to 0.4) to category 3 (0.6 to 0.8)	7,380 (1,179)	52,529 (8,388)
BB providers (25 mbps), 2011-2012			
	1st change: from median # providers (0.07) to one more (1.07)	1,768 (314)	12,583 (2,232)
	2nd change: from 1.07 providers to 2.07 providers	427 (76)	3,037 (539)
Higher education degrees			
	1st change: from 1st quartile to median	2,309 (183)	16,437 (1,303)
	2nd change: from median to 3rd quartile	4,474 (355)	31,846 (2,524)
Inventor stock per capita			
	1st change: from 1st quartile to median	3,052 (173)	21,725 (1,228)
	2nd change: from median to 3rd quartile	2,877 (163)	20,473 (1,157)
Intermodal transportation facilities per capita			
	Change: from the median (no facilities) to 3rd quartile (1 facility or 8.2 facilities per cap.)	3,150 (599)	22,422 (4,265)
% Rural population			
	1st change: from completely urban (0) to half rural (50)	-19,404 (787)	-138,105 (5,601)
	2nd change: from half rural (50) to completely rural (100)	-19,404 (787)	-138,105 (5,601)

Notes: S.e.'s account for clustering at the county-industry level.

Figures

Figure 1: Time trend of establishment births

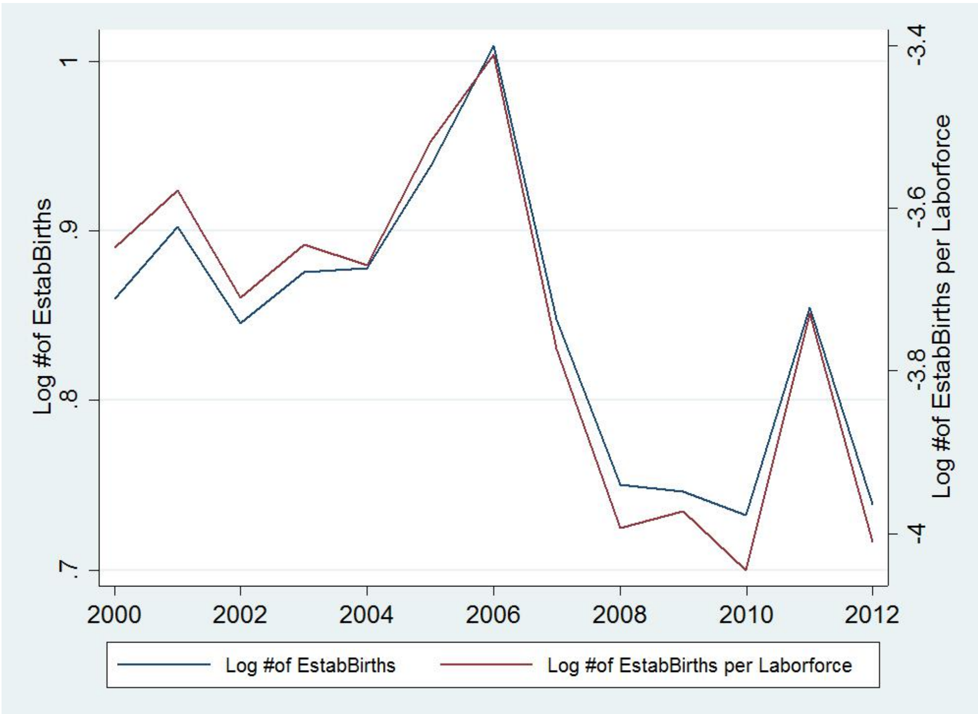

Figure 2: Time trend of broadband infrastructure

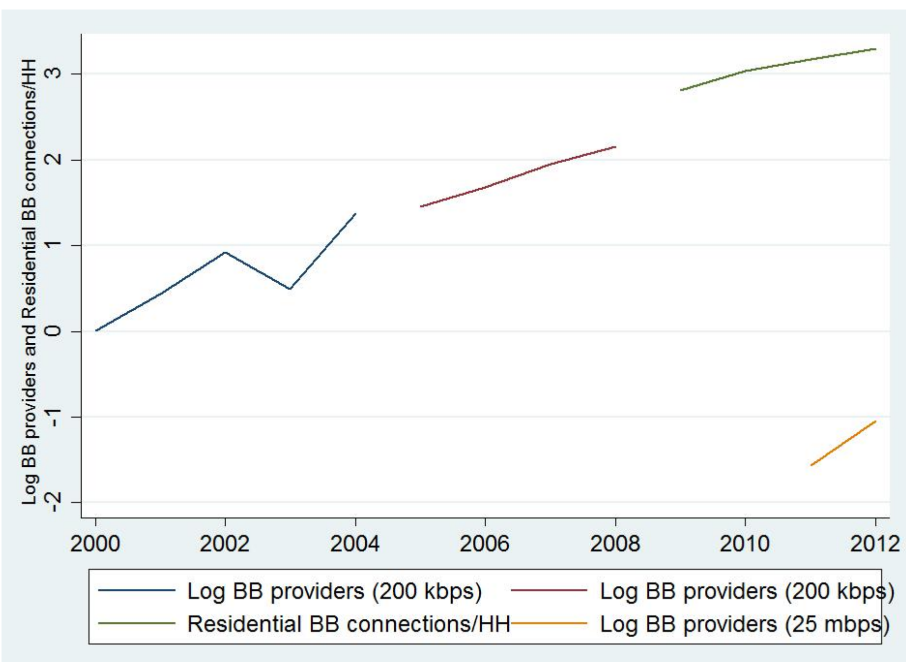

Figure 3: Time trend of intellectual infrastructure

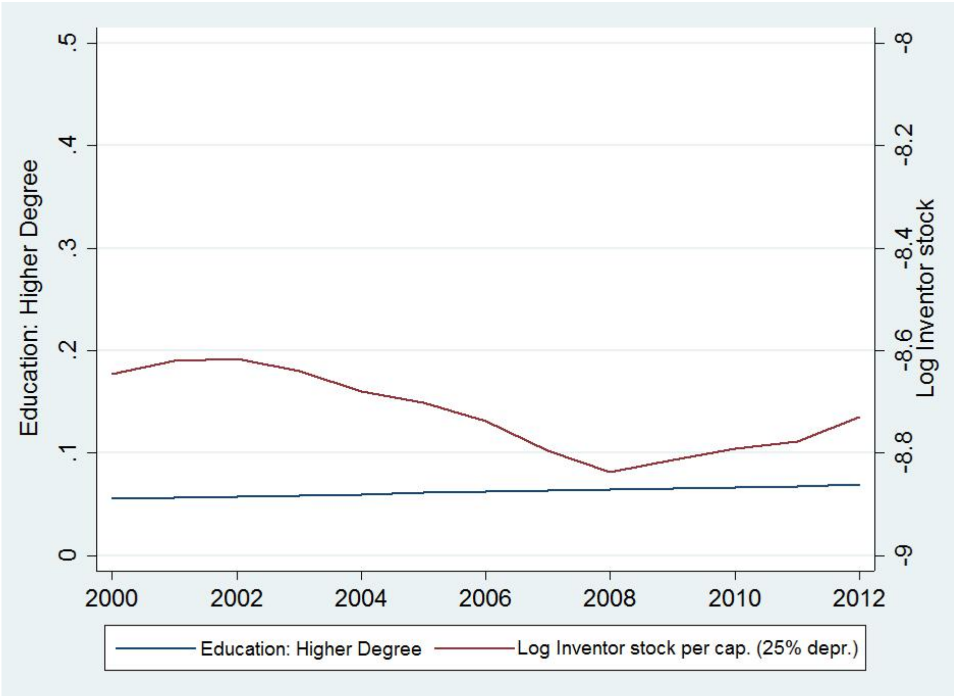

Figure 4: Time trend of transportation infrastructure

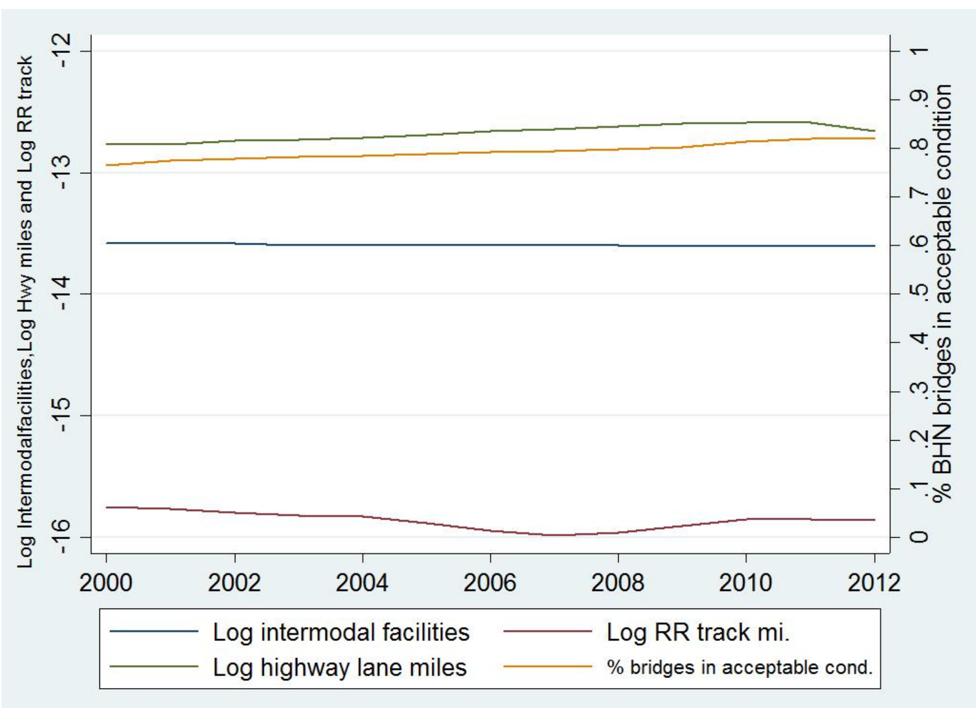

Figure 5: How the marginal effect of broadband providers changes with transportation infrastructure