UNIVERSIDAD EAFIT

Vigilada Mineducación

Documentos de trabajo

Finanzas, Economía y Gobierno

N° 01-22

2022

Quality of Communications Infrastructure, Local Structural Transformation, and Inequality

Camilo Acosta, Luis Baldomero-Quintana

Quality of Communications Infrastructure, Local Structural Transformation, and Inequality

Camilo Acosta [*], Luis Baldomero-Quintana [†‡]

November 2, 2022

[LATEST VERSION HERE]

Abstract

We analyze the causal impact of improvements in the quality of communication infrastructure on the structural transformation of US counties. Our treatment is the quality of communication infrastructure in a county, measured by the average Internet speed offered to businesses. We use as an instrumental variable the spatial structure of ARPANET, a network funded by the Department of Defense that is considered the precursor of the Internet, and whose location we determine using historical government documents. We show that faster Internet stimulates short-run growth and increases the shares of employment and GDP in high-skilled services, while negatively affecting sectors such as retail, accommodation, and food services. Two mechanisms explain our results. First, input-output linkages since industries that buy more ICT inputs increase their weight on the local economy. Second, a rise in high-skilled workers in ICT-intensive occupations, which is consistent with the Rybczynski theorem of the Heckscher-Ohlin-Vanek model and with the presence of capital-skill complementarities. Lastly, we find that better Internet increases earnings inequality within U.S. counties. Such finding has implications for Internet subsidies across the country.

JEL Codes: F16, H54, L86, N92, R12

Keywords: communication costs, Internet, infrastructure, local structural transformation, Heckscher-Ohlin-Vanek model, history of technology

[*]School of Finance, Economics, and Government, Universidad EAFIT, cacosta7@eafit.edu.co

[†]Department of Economics, William & Mary, labaldomeroqui@wm.edu

[‡]This project was funded by William & Mary Arts & Sciences Faculty Grants Fund 2021. We are thankful to Vint Cerf and Bob Kahn, (creators of the TCP/IP protocol and considered among the *fathers of the Internet*), for sharing their knowledge about the ARPANET network. We are grateful for the detailed comments from John Lopresti and Astrid Marinoni. This paper benefited from conversations with Barthelemy Bonadio, Kerem Coşar, Mario Crucini, Brian Feld, Marco Gonzalez-Navarro, Xian Jiang, Andrei Levchenko, Frédéric Robert-Nicoud, and William Strange. Eli Rothleder, Yiwen Sun, Maia Tindall, and Lauryn Walker provided extraordinary research assistance. Lauryn Walker obtained coding and econometrics training and worked for this project with funding from the Garrett Fund of the Economics Department at William & Mary. All errors are our own.

You can see the computer age everywhere but in the productivity statistics
- Robert M. Solow

1 Introduction

Communications infrastructure differs from other types of infrastructure like roads or railroads. First, it facilitates the transmission of ideas. Second, it allows individuals to consume digital services produced in any location. Third, communication infrastructure allows a region to reduce communication costs with any location while roads or railroads reduce trade costs only between locations connected by such physical infrastructures.[1] The importance of communication infrastructure is such that U.S. lawmakers recently allocated 10 billion dollars of federal funds to improve Internet access (The White House, 2022). Nevertheless, it is unclear to what extent the quality of communication infrastructure impacts local economic outcomes, and whether it mattered only for early stages of the Internet in the 1990s or if its impact is still relevant today.

In this paper, we provide causal evidence that modern improvements in the quality of communication infrastructure induce local structural transformation in U.S. counties. To obtain causal estimates, we use cross-sectional regressions at the county level in 2018 in which the quality of Internet offered to businesses is our main treatment, and our outcomes of interest are the counties' short-run growth in GDP, payroll and employment, as well as employment and GDP sectoral shares. Our first set of findings shows that faster Internet offered to businesses positively impact short-run GDP and employment growth. Our second set of results shows that better provision of Internet to firms shifts the local economic structure towards high-skilled services, and away from other services such as retail, accommodation services, restaurants, and some finance sub-sectors. We also find important heterogeneous impacts of better Internet on employment and GDP shares within some specific sectors, such as finance, health, support services, retail and wholesale. Lastly, we observe that faster Internet increases local wage inequality.

Our findings suggest that quality improvements in communications infrastructure induce local structural transformation. In particular, faster Internet offered to businesses favors industries that use intensively information and communication technology (ICT) inputs, and that hire intensively workers that use ICT more. Hence, our results are consistent with the Rybczynski theorem from the Hecksher-Ohlin-Vanek (HOV) model (Rybczynski, 1955). We also find that high-skilled workers sort into counties with lower communication costs, which leads to the observed output growth in sectors that hire these workers intensively.[2] Most tests of the HOV model have specific assumptions

[1] Consider regions i, j, and k. Regions i and j are adjacent while region k is far from both. Let trade costs between two locations be $_{od} = t_o t_{od} t_d; \forall o; d \in \{i; j; k\}$ where $o \neq d$. A new road connecting i and j reduces trade costs $_{ij}$ mainly through reductions in t_{ij}. Faster Internet in i reduces trade costs with both j and k via lower t_i.

[2] The Rybczynski theorem shows that higher endowments of one factor lead to a more than proportional ex-

about technology, production, exports, nontraded sector, etc. and several papers attempted to relax them (Harrigan, 1997; Davis and Weinstein, 2001; Trefler, 1995). Our approach does not require functional form assumptions on production, exports, or technology, since we find that Internet quality is a shifter of one factor endowment. Moreover, we include both tradable and non-tradable services in our analysis. Our findings are also in line with the presence of capital-skill complementarities for the case of ICT equipment.

Obtaining causal estimates of the impact of telecommunications infrastructure on local economic outcomes is challenging. On one hand, counties with better amenities or higher productivity attract more college-educated workers and productive establishments (Glaeser et al., 2004), thus Internet Service Providers (ISPs) may offer better service in these locations. On the other hand, counties with low competition among ISPs may provide lower quality Internet, which may disincentivize establishments that need high-quality Internet to locate in these areas.

We circumvent these issues by using an instrumental variable approach. Our instrument is the distance of counties' centroids to the lines connecting ARPANET nodes,[3] a military research network that was the backbone of Internet in its initial stage.[4] These lines represent the telecommunications equipment that connected ARPANET nodes. To build our instrument, we digitized ARPANET maps using historical government reports. We also talked with Bob Kahn and Vint Cerf, two of the ***fathers of the Internet***, who worked on the design of ARPANET.

Historical government reports help us document that the ARPANET structure satisfies the IV assumptions. First, the history of Internet supports the relevance of our instrument. From 1969 to the early 1980s, ARPANET's physical infrastructure (both the nodes and the lines connecting them) was the backbone of the Internet. Due to path dependence of infrastructure (Duranton et al., 2014; Duranton, 2015), the old Internet backbone is a good predictor of the modern Internet backbone, whose location is not public.[5] Physical closeness to the modern Internet backbone allows ISPs to provide higher quality Internet at a lower cost.[6] Our data supports the idea that locations closer to ARPANET lines have faster Internet offered to firms today.

pansion of the output in the sector which uses such factor intensively, and a decline of the output of the sector that does not use such factor intensively (Rybczynski, 1955).

[3]In computer networks, a node is a connection point in a network that is a processing device with an assigned address, as a router, computer terminal, peripheral device, or mobile device, (Encyclopedia.com, 2022).

[4]A backbone of a network is a high speed network that connects low-speed local networks. A national backbone of the Internet refers to the high speed network infrastructure that connects local networks in cities, companies, universities, among other, across the country and the globe.

[5]There are two reasons for this. First, the information belongs to private companies who are top tier networks that maintain the Internet backbone equipment and provide service to last-mile Internet service providers (Xfinity, Comcast, etc.). Second, the Internet backbone specific location is of national security interest. (Wall, 2021).

[6]The modern Internet backbone requires underground fiber optic infrastructure protected by a external layer called raceway or conduit. Installing fiber network in general (for a city or for a national backbone) requires underground construction. But the closer a county is to existing fiber network, the lower the underground construction costs since the amount of underground construction is lower.

The instrument plausibly satisfies the exclusion restriction since the Department of Defense (DoD) decided the location of ARPANET considering military technology needs, contracting relationships between the DoD and academics, and characteristics of computer science departments. Commercial companies did not influence ARPANET. Moreover, we exclude counties with ARPANET nodes from our sample and our instrument only uses the ARPANET lines. We expect our instrument to satisfy the exogeneity assumption since current productivity and amenities of counties are unlikely to be related to whether ARPANET lines are routed through these counties. This is because a straight line is the least cost way to connect two locations with physical infrastructure (roads, cables, etc.). Our data confirms that the presence of an ARPANET line in a county is uncorrelated with its share of high-skilled services in the 1970s, which also supports the exogeneity assumption. Lastly, our results are robust to excluding from our sample counties from six major metropolitan areas, including three California tech hubs (Boston, DC, Los Angeles, NYC, San Francisco, and San Jose).

Our instrument is related to Forman et al. (2012) and Jiang (2022). Forman et al. (2012) uses ARPANET nodes as an instrument for counties' private investment on Internet to analyze how such investment impacted wages of U.S. counties between 1995 and 2000. Their instrument is weak because nodes were scarce and investment is lumpy. We adjust their instrument in different ways: we use the lines that connected the nodes as an IV, we drop the counties with these nodes, and we also use a different treatment, the quality of Internet provision to businesses rather than Internet investment. Our instrument is valid and can be used for other cross-sectional analyses. Similarly, Jiang (2022) uses NSFNET nodes from the late 1980s, combined with the privatization of Internet in 1995, to analyze how access to Internet affects the spatial structure of manufacturing firms. Informed by history, we chose ARPANET lines instead of NSFNET nodes in our empirical framework. Historical sources suggest that NSFNET nodes' location are correlated with counties' productivity in the 1970s due to the incentives that the National Science Foundation (NSF) gave to universities who managed the nodes, while ARPANET nodes are not since their location was driven by military interests (Abbate, 2000; DARPA, 1981; Hauben et al., 1998; Leiner et al., 1997). Nevertheless, our approach does not contradict the identification strategy of Jiang (2022).[7]

We analyze two mechanisms: industry linkages, and sorting of high-skilled workers with ICT-related occupations. First, better Internet benefits those sectors with strong dependence on ICT-inputs and workers who use them, while the impact on the other sectors is zero or statistically insignificant. Second, we find that faster Internet in a county increases the number of workers in occupations

[7] Correlation between the location of NSFNET nodes and county productivity does not represent an identification issue for Jiang (2022). First, her variables of interests are outcomes for manufacturing firms, whose location determinants are likely to be unrelated to ICTs before the mid-1990s. Second, her framework requires changes between two spatial equilibria in the internal organization of these firms. Her empirical framework satisfies these assumptions. If researchers were to use her identification strategy to analyze the internal organization of services firms or county level outcomes, some adjustments would be necessary to address identification issues.

related to ICT, such as management, business and nance, o ce and administrative support, engineering, and computer sciences. Interestingly, the sectors who have a greater abundance of these workers are the same sectors that are favoured more by faster Internet speeds.In our heterogeneity analysis, we show that this sorting leads to an increase in earnings inequality, it increases the earnings of high-skilled workers and it does not a ect low-skilled workers wages.

Our work has implications for infrastructure policy: policies to reduce the digital gap via subsidies for Internet access can induce changes in the regional economic structure. Many nations use public funds to subsidize Internet access in rural or isolated areas: Canada (Government of Canada, 2019), the U.S. (The White House, 2022), Germany (European Commission, 2022), and the U.K. (Hutton, 2022). Our results indicate that these subsidies may favor high-skilled workers within the favoured region.

Our ndings relate to three international trade topics. First, our results are in line with the Rybczynski theorem predictions. We nd that counties with faster Internet have more college-educated workers and a larger share of workers who are engineers, managers, and computer science professionals. Such counties have higher economic activity in industries that employ intensively workers with college education or in the aforementioned occupations. Di erently, sectors that employ these professionals less intensively have smaller sectoral shares (e.g., wholesale, branch banks, and food services). Second, our empirical results suggest the presence of capital-skill complementarities as the use of better Internet is likely to induce rms to purchase new ICT capital, since the use of faster Internet requires newer devices. ICT capital is a complement for high-skilled workers who might become more productive in their tasks. Third, our results show the impact of globalization on local inequality: high-skilled workers bene t more in a county with lower communications costs.

Literature review . We relate to empirical work exploring the impacts of infrastructure on local outcomes (Cosar et al., 2022; Duranton et al., 2014; Duranton, 2015; Gertler et al., 2022). First, Cosar et al. (2022) show that highway upgrades increase trade ows and manufacturing employment. Second, Duranton et al. (2014) nd that cities with more highways in the US specialize in the production of heavy goods, while Duranton (2015) documents that major roads induce Colombian cities to trade light goods. Third, Gertler et al. (2022) document that road maintenance rises local welfare in Indonesia. We complement these studies about road infrastructure by analyzing how telecommunications infrastructure quality impacts industry composition and inequality.

Our study relates to classical studies on HOV model tests using aggregate data (Leontief, 1953; Leamer, 1980; Deardor , 1984; Bowen et al., 1986; Tre er, 1995; Harrigan, 1997; Davis and Weinstein, 2001). Our approach di ers in four ways. First, we focus on communication costs that a ect regional trade with other regions and the rest of the world. Second, our context varies from cross-national studies since products between counties are likely to be competing goods, while this is not necessarily the case for trade between developing and developed nations. For example, a craft beer from the rural county of Hanover, VA, competes with a beer produced in the City of Richmond, VA. Third, we include traded and non-traded services sectors. To some extent, we could argue that many non-traded service sectors can be traded between regions within a country, since people travel to a nearby locations to acquire them. Fourth, we complement previous work by not including restrictive assumptions on production, technology, or exports in our regional analysis.[8] Our work also di ers from studies that test the HOV model predictions based on trade liberalization in developing nations (Atolia, 2007; Esquivel and Rodr guez-Lopez, 2003) since we analyze a reduction in communication costs in U.S counties.

Furthermore, our work relates to studies on how both technological change and trade increase the use of capital that complements high-skilled workers, thus impacting inequality. This work includes structural models (Burstein et al., 2013; Parro, 2013), reduced-form analysis of trade liberalization in developing nations (Verhoogen, 2008), and historical evidence from the early 1900s in the U.S. (Goldin and Katz, 1998). We add to the previous literature by providing causal estimates of how ICT-related capital impacts local inequality in U.S. regions.

The closest study to our work is Michaels (2008). He nds that the construction of the US Interstate Highway System increased the demand for skilled labor in skilled-abundant counties. His results are consistent with some predictions of the HOV model. We complement his work by studying a continuous measure of communication infrastructure, rather than a single road infrastructure project. Contrarily from us, Michaels (2008) does not nd conclusive evidence that highways changed the local industrial composition. Moreover, we nd that better infrastructure quality has heterogeneous e ects within the same broad sector, as are the cases of nance and health services.

We relate to the economic geography literature that explores how infrastructure a ects specialization at the regional level using structural models (Fajgelbaum and Redding, 2022; Sotelo, 2020; Baldomero-Quintana, 2022). Other studies consider how infrastructure a ects welfare and aggregated outcomes leaving the sectoral impacts aside (Alder, 2016; Allen and Arkolakis, 2022; Allen and Atkin, 2016; Asturias et al., 2019; Bonadio, 2016; Donaldson, 2018; Donaldson and Hornbeck, 2016; Faber, 2014). We complement this literature by providing causal evidence that the quality of infrastructure impacts local specialization from a purely empirical standpoint, without making functional form assumptions on production, preferences, or trade costs.

[8] Several tests of the HOV model have attempted to relax many strong theoretical assumptions, but they still require some, such as constant return of scale, non-joint production, or exogenous prices as in Harrigan (1997); or similar production functions across nations, exports level is expressed as a proportion of the purchasing country's GDP, and factor price equalization as in Davis and Weinstein (2001).

We relate to work in spatial economics on the effects of communication costs and ICT. This includes work on ICT and agglomeration (Charlot and Duranton, 2006; Gaspar and Glaeser, 1998; Glaeser and Ponzetto, 2007; Malecki, 2002; Lin, 2011; Zook, 2002); Internet and real estate prices (Ahlfeldt et al., 2017; Ford et al., 2005a; Dietzel, 2016; Beracha and Wintoki, 2013); and communication and innovation (Carlino et al., 2007; Kantor and Whalley, 2019; Rosenblat and Mobius, 2004). We complement this literature by focusing on Internet and local economic structure.

Recent work has explored how communication costs impact firms' entry, performance and structure (Arfi and Hikkerova, 2021; Acosta and Lyngemark, 2020; Beem, 2022; Forman and Van Zeebroeck, 2012; DeStefano et al., 2022; Marinoni and Roche, 2022), and trade flows (Fink et al., 2005; Allen, 2014; Freund and Weinhold, 2004; Blum and Goldfarb, 2006; Breinlich and Criscuolo, 2011; Juhász and Steinwender, 2018; Steinwender, 2018; Cristea, 2011). We complement these studies in two ways. First, we document empirically that the quality of communications infrastructure (intensive margin), and not only access (extensive margin), impacts local economic outcomes. Second, we study regional economic specialization, while previous work focuses on trade or firm outcomes.

We share similarities with Forman et al. (2012) and Jiang (2022), but our research questions differ substantially. Forman et al. (2012) focuses on the impacts of private Internet investment on local labor markets in the late 1990s, finding no impacts. Our results differ since we analyze a period of study when Internet is a highly adopted technology that experienced major changes since the 2000s.[9] Our study is related to Jiang (2022), who finds that the spatial organization of manufacturing firms changes when they accessed Internet after 1995. Our results complement hers since we find that counties with better Internet have higher manufacturing and service activity in specific sub-sectors.

The rest of the paper proceeds as follows. In Section 2, we document the history of Internet. Such historical context informs our empirical strategy. In Section 3, we describe our data and present descriptive statistics. Section 4 presents our empirical model, including our identification strategy. In Section 5 we present the results of the paper, together with the two mechanisms proposed that can explain such results. Section 6 concludes.

2 Context: The History of Internet

In this section we document the origins of the Internet. We provide information on ARPANET, the first network to operate as the Internet backbone. We also present facts about the National Sciences

[9] Internet speeds increased exponentially since the 2000s. Connections used in the 1990s were based on dial-up with download and upload speeds of at most 56 Kpbs. The average speeds in US counties in 2018 were between 2 and 671.1 Mbps for download, and between 1.1 and 615.7 Mpbs for upload. If we focus on the lower bounds, speeds were 17 to 35 times faster in 2018 relative to the 1990s. Moreover, social media and modern search engines changed markets, including housing, retail and labor markets (Dietzel, 2016; Beracha and Wintoki, 2013; Ford et al., 2005b; Oestmann and Bennöhr, 2015; Cavallo, 2018; Dingel and Neiman, 2020).

Figure 1: ARPANET Connections in 1979 and 2018 Internet Speeds

Panel A: Download Internet speeds o ered to businesses

Panel B: Upload Internet speeds o ered to businesses

Note: these gures show the mean download and upload speeds o ered by Internet service providers to businesses in every U.S. continental county (i.e., di erent from the Internet service o ered to households). Blue counties belong to the top quintile, while yellow counties are in the lowest quintile. The gures also display the ARPANET nodes and lines (which represent the connections between nodes) for April 1979. Source: FCC and Cerf and Khan (1990).

Foundation network, NSFNET, which took over the functions of ARPANET in the late 1980s. Our historical context supports the idea that the location of ARPANET nodes and connecting lines is exogenous to county productivity while NSFNET nodes are not. In addition, we document that ARPANET network was the rst backbone. Thus, it is a good predictor of Internet speeds in the U.S. as Figure 1 shows.

Origin of ARPANET . ARPANET was a research project nanced and developed by DARPA, a research o ce of the DoD. DARPA's objective was to create a network that connected military computers. Two factors drove the nancing of the project. First, the DoD had a strong interest in nancing research after the Soviet Union launched its rst satellite, the Sputnik. Second, the DoD wanted to lower administrative costs by allowing its computers to share data. In the early 1960s, the DoD had thousands of computers operating autonomously without any connection, which generated large costs since all data les and software had to be reprogrammed in every device. In addition, military personnel had to be trained to use devices from di erent manufacturers. These costs doubled the DoD's budget for software creation and maintenance (DARPA, 1981).

DARPA hired a reputable computer science researcher to build a computer network: Dr. J.C.R. Licklider. He proposed a novel idea at the time: use of computers to improve human communication (Licklider and Taylor, 1968). In the 1960s, computing rms and most U.S. academics focused on improving the speed of computers to complete tasks, a concept known as batch processing (Hardy, 1980). This situation incentivized Dr. Licklider to shift all research computer research contracts away from private companies and towards academic departments interested mainly in computer network research. Thus, DARPA established DoD contracting relationships with some computer research departments at U.S. universities. Consequently, commercial interests played no role in the design and development of ARPANET (DARPA, 1981).

The main idea behind ARPANET and the Internet was to connect independent local networks (DARPA, 1981; Leiner et al., 1997): if one local network was lost, the system continued working. ARPANET demanded novel technological resources: a speci c software to de ne a protocol (set of rules for data transmission) and equipment to guarantee the stability and delay of a network (the time for a signal to traverse a network). During the late 1960s, prior to the rst ARPANET connection in 1969, DARPA researchers focused on solving these technical issues. Notably, the existing telephone infrastructure was insu cient for ARPANET. [10] The rst four ARPANET nodes were connected in 1969. These nodes were research centers with DoD contracts and were chosen due to speci c technical expertise.[11] The network was declared fully operational in 1971. Figure

[10] In 1966-1967, an MIT laboratory and the military contractor SDC in Santa Monica, CA, connected two computers using the existing telephone circuit technology. The experiment proved that existing telephone infrastructure was insu cient to establish a good quality network (Hauben et al., 1998; Leiner et al., 1997).

[11] These four nodes correspond to the University of California|Los Angeles, University of California{Santa Bar-

A-1 presents the original description of the network as described by Vint Cerf and Bob Kahn (two ARPANET researchers that are part of the list of the fathers of the Internet) in 1969.

In the 1970s, ARPANET required DoD's resources and well-coordinated frontier engineering expertise (DARPA, 1981; Hauben et al., 1998; Leiner et al., 1997). Academic researchers solved novel engineering issues frequently. Researchers working on ARPANET collaborated intensively even if they belonged to di erent universities and their engineering documentation was open (Leiner et al., 1997).[12] Thus, the location of ARPANET nodes also depended on the work style of computer science departments: uncooperative computer science departments were less likely to be a part of ARPANET.

To summarize, three factors in uenced the location of ARPANET nodes. First, in the 1970s only universities with DoD contracts were connected to ARPANET (Abbate, 2000). Second, universities with a research agenda focusing on networks were more likely to participate in ARPANET, while computer science departments that focused on batch processing{predominant across U.S. universities{were unlikely to participate in the network (Hauben et al., 1998). Third, collaborative computer science departments were more likely to join the project. Thus, we conclude that the location of ARPANET nodes in the 1970s was driven by factors unique to national defense and computer science departments (DARPA, 1981).

From ARPANET to NSFNET . In the early 1980s, only nodes related to national defense computing research or operations were connected to the Internet. The situation was di erent in the late 1980s: the backbone was transferred from military to civilian control, and Internet users included universities, private companies and federal agencies (Abbate, 2000). Five events in uenced this change. First, DoD created a separate military computing network in 1983, MILNET. Afterward, ARPANET did not transmit military data, although some military research centers were connected to the network for scienti c purposes. Second, NSF founded NSFNET, a national network of regional academic networks. Third, DARPA and NSF collaborated to connect ARPANET and NSFNET. Fourth, NSF encouraged NSFNET regional networks to provide Internet service to commercial companies. Fifth, ARPANET transferred the role as backbone of Internet to NSFNET in 1988, and was decommissioned in 1990 (Abbate, 2000; Leiner et al., 1997).

During the 1980s ARPANET opened its network to more academic centers.[13] DARPA wanted more academics to use the technology, and transfered the backbone responsibilities to another

bara, Stanford Research Institute, and the University of Utah.

[12] ARPANET's technical notes about engineering network design were key for the network's success. The notes were open to the public even if the DoD nanced the network (Leiner et al., 1997; Hauben et al., 1998). Collaboration and open documentation helped ARPANET researchers to develop fast solutions for engineering issues.

[13] The Internet is de facto a network of local networks. It is divided in three tiers. The upper tier is considered a backbone because it reaches all local networks.

group (DARPA, 1981). In 1981, a consortium of universities requested a grant from the NSF to create CSNET, an academic network implemented between 1981 and 1984. In 1981, DARPA signed a cooperative agreement with CSNET to share ARPANET's infrastructure (McKenzie and Walden, 1991; Leiner et al., 1997). To connect to CSNET, NSF agreed to provide resources, but each university or research center would independently manage their internal network.

In 1985 NSF started building NSFNET, a national network that connected regional networks created by universities[14]. NSFNET was built onto the CSNET infrastructure. The new network created a two-tier system: NSFNET regional networks, and a national backbone network connecting them. The construction of NSFNET took three years (Abbate, 2000). In 1986, NSF and DARPA made an agreement to guarantee that ARPANET and the regional NSF networks had interoperability (Leiner et al., 1997). Later, in 1987, the NSF reached an agreement to use ARPANET resources to connect regional networks, in exchange for sharing ARPANET's operating costs (Abbate, 2000). The NSFNET national network started functions in 1988.

Determinants of the location of NSFNET. The NSFNET regional networks (nodes) were subject to commercial interests due to the incentives imposed by the NSF during its creation. First, the regional networks had to be financially autonomous within three years of receiving funding. Second, NSF encouraged regional networks to find commercial clients (Leiner et al., 1997). Although Federal law prohibited commercial companies to use the NSF national backbone (U.S. Congress, 1992), private companies could connect to regional networks (Leiner et al., 1997; McKenzie and Walden, 1991). This scheme was implemented as the NSF had the objective that private firms would create national networks (Leiner et al., 1997). Although commercial networks emerged, no national private backbone appeared until the privatization of NSFNET in 1995.

The NSF financing scheme implied that only universities (or groups of institutions) would create a regional network or node in a region if they had economies of scale. A university could reach economies of scale if it had productive firms nearby, the institution was large, or the college had other institutions nearby. Such factors are plausibly correlated with the productivity of regions. Thus, the presence of NSFNET nodes and local productivity could be correlated.

NSFNET as Internet backbone and its privatization. In the late 1980s, the growth of Internet users made ARPANET physical infrastructure obsolete as Internet backbone. Thus, DARPA planned the decommissioning of the network and NSFNET overtook ARPANET as the Internet backbone (McKenzie and Walden, 1991; Abbate, 2000). The protocol designed by Vint Cerf and

[14] The regional networks were BARRNET for the San Francisco Area, MIDNET in the Midwest, WESNET in the Rocky Mountain region, USAN for those research centers connected to the National Center for Atmospheric Research, NORTHWESTNET in the Northwest, NYSERNET in the New York state area, SESQUINET in Texas, SURANET in the Southeast. NSFNET also connected large computer centers in San Diego, University of Illinois, and Pittsburgh.

Robert Kahn for ARPANET made the transition smooth (TCP/IP). Between 1988 and 1989, DARPA sites transferred their host connections from ARPANET to NSFNET, and in February 1990 ARPANET was officially decommissioned (Abbate, 2000). Independent commercial networks began to grow in the 1980s and 1990s due to the NSFNET regional network' services provided to private firms (Abbate, 2000). In 1995, the NSF transferred the NSF national backbone to private companies. Harris and Gerich (1996) documents the technical details of this transition.

The history of Internet allows us to reach three conclusions, which will be helpful for our empirical strategy. First, ARPANET was not influenced by commercial interests, rather by military ones. Second, the location of nodes was driven by DoD contracts, and the research agenda and work philosophies of computer science departments in the 1960s and 1970s. Third, NSFNET regional networks were influenced by commercial interests. Thus, NSFNET nodes are likely to be correlated with local productivity, while ARPANET nodes and the lines connecting them are not.

3 Data

Our data comes from four main sources: the US Census Bureau, the Bureau of Economic Analysis (BEA), the Federal Communications Commission (FCC), and historical maps created by the DARPA office at the DoD. In this subsection we describe in more detail each of these data sources. Afterward, we show some descriptive statistics for the main variables of interest.

Internet data. To capture current the quality of the Internet offered to businesses, we use data from the FCC for 2014 and 2018.[15] These data come from FCC's Form 477, which must be filled by all broadband providers twice a year to report the list of census blocks in which they offer a particular technology (Federal Communications Commission, 2019). Therefore, the data contain for each census block, a list of all Internet providers, together with the offered technology of transmission to firms (e.g., cable model, xDSL, fiber, fixed wireless, etc.), the maximum advertised download and upload speeds/bandwidth (for consumers) and the maximum contractual download and upload speeds/bandwidth (for businesses) for each technology.[16] From these data, we keep only those providers and technologies offered to businesses, which we use to compute the mean and

[15] We use this period for three reasons. First, the FCC started publishing these data in December 2014, hence we cannot obtain Internet speed measures at the county level from the FCC before this year. Second, when we downloaded the data on early September, 2022, some Internet technologies were missing in the 2019 FCC files. Our hypotheses are that there were either changes in their methodology or pending updates in the data. Since we did not find any documented reason for these missing data, we do not use data from 2019. Third, we avoid using data for 2020 as the COVID-19 pandemic substantially altered the demand and supply of Internet due to the boom of working from home. In addition, the local economic structure was also impacted by the pandemic, since some services sector were negatively impacted and manufacturing industries experienced a boom in 2020-2021, while the opposite occurred in 2021-2022. Thus, we chose 2018 as our base year.

[16] Download speed corresponds to the speed used to download data from a server to a computer in the form of text, files, audio, images, videos, etc. On the other hand, upload speed refers to how fast can information be sent (in the form of text, audio, images, etc.) from a computer to another device connected to the Internet.

the median download and upload speeds for each county.

Current Economic Activity. We use employment counts and aggregate payroll by county and sector from the Census Bureau's 2014 and 2018 County Business Patterns (CBP). The CBP includes the number of establishments and employment for every county during the second week of March, as well as the annual payroll. We compute the average wage within counties as the ratio between annual payroll and total employment. The CBP data also contain information by NAICS (North America Industry Classification System) sector, up to 6-digits. However, due to confidentiality restrictions, a data point is only published if it contains at least three establishments. Thus, to minimize the number of missing values arising from these restrictions, we use sectoral data at the 2- and 3-digit NAICS code.[17] For each sector, we compute the share of employment and payroll in each sector within each county. Such shares represent a proxy that measures the level of industrial specialization or composition of the county. We complement these data with information on each county's GDP in 2018 (total and by 2-digit NAICS code) from the BEA.

ARPANET. For our instrumental variable, we digitized images of decommissioned DoD documents that contained ARPANET maps. In 1990, Vint Cerf and Robert Kahn (Internet pioneers) collected these maps and published them in the Journal of the Association for Computing Machinery (Cerf and Khan, 1990), and are only available in the physical version of the journal. These maps include the abbreviated name and the location of the nodes, and the lines connecting them, which represent the infrastructure connecting ARPANET nodes. As the name of the nodes is not available in the journal (only their abbreviation), we used DoD historical reports and other historical sources to find the exact address of some of the nodes (Network Information Center, SRI International, 1978; DARPA, 1981; Hauben et al., 1998). We also talked with Vint Cerf and Bob Kahn, who shared their knowledge about the ARPANET network, including the names and location of those nodes whose information was not in historical government reports. Figure 2 shows the original and the digitized maps for 1979. Using the exact location of the nodes and their connecting lines, we compute the minimum distance between each county's centroid and one of the lines, together with an indicator variable that equals 1 if a county contains a node (and 0 otherwise), and another that equals 1 if a line is routed through it.

Other variables. We also gather different geographic and economic characteristics of each county. First, we include dummy variables for counties on the Canadian or Mexican border, or along a coast (oceanic or Great Lakes). Second, we compute the average elevation in each county, its land and water area, as measures of physical geography affecting the structure of ARPANET and modern Internet provision. Third, we compute a measure of market access as the distance of each county to the centroid of the nearest Metropolitan Statistical Area as in Michaels (2008). Finally, we

[17] The full list of 2-digit NAICS sector is included in Table A-1 in the Appendix.

Figure 2: Structure of ARPANET as of 1979

Panel A: Original Map

Panel B: Digitized Map

Note: Panel A displays a scanned image of the ARPA network as of April 1979. The map was created by Vint Cerf and Bob Khan, Internet pioneers. Panel B displays our digitized version of the same map. Source. Cerf and Khan (1990) compute population density in 2018 and population growth between 2014 and 2018, similar to the growth rates computed above. These variables are our control variables.

Descriptive statistics. In Table 1, we present the average and median values for some of the variables of interest. Regarding local sectoral structure, around half of employment and total payroll in the average county is generated in the *Other services* category, which includes *Wholesale and Retail Trade*, *Administration and Support*, *Construction*, *Accommodation and Food Services*, and similar. A third of the counties' employment on average is in *High-skilled services*, which includes *Information Services*, *Management*, *Professional Services*, *Educational Services*, and similar. *Manufacturing* accounts for approximately 15% of the employment on average. For all variables, the mean and the median share of sectoral employment are quite similar. Figures A-3 and A-4 show the spatial distribution of these sectoral shares across the U.S. For the case of Internet quality provision, the average download and upload speeds offered to firms in an US county are 97 and 86 megabits per second (Mbps), respectively. Finally, in 1979 only 1% of counties had an ARPANET node, while 14% had a line going through; and the average county was 192km away from a line.

4 Empirical Model and Identification

In this section we discuss the empirical framework we use to estimate the effects of quality of Internet provision on current local economic structure. Moreover, we examine the use of ARPANET lines as an instrumental variable, and we asses the assumptions we need to interpret our results as causal. To measure the effects of Internet quality on modern county short-run economic growth and local structural transformation outcomes, we estimate the following model:

$$Y_c = \alpha + \beta \log(\text{InternetQuality}_c) + X_c'\gamma + \varepsilon_c \quad (1)$$

where Y_c is a local economic outcome in county c in 2018, such as the 4-year growth rate of real GDP or employment, or the share of employment in a particular sector in 2018. We use these two years due to the FCC data issues highlighted in the previous section. Moreover, InternetQuality_c is a measure of Internet speed offered to businesses (average download or upload speeds) in county c in 2014 (for growth regressions) or in 2018 (for levels); and X_c' is a group of geographical and economic characteristics; ε_c considers the unobserved factors that impact local economic outcomes, such as productivity in county c.[18] In all our specifications, we exclude from our sample any county that had at least one ARPANET node.

The group of control variables includes several geographic characteristics, including dummy vari-

[18] In all the estimations we use Spatial Heteroskedasticity and Autocorrelation Consistent (SHAC) standard errors as proposed by Conley (1999) with a ratio of 28.55km, which correspond to the ratio of a circle that would cover the surface of the median metropolitan statistical area in the U.S.

Table 1: Descriptive Statistics

Category	Average across Counties	Mean	Median
	Share agriculture & mining employees	0.02	0.00
	Share manufacturing employees	0.15	0.12
	Share high-skilled services employees	0.30	0.29
	Share other services employees	0.53	0.52
	Share agriculture & mining payroll	0.04	0.00
	Share manufacturing payroll	0.19	0.16
	Share high-skilled services payroll	0.33	0.32
	Share other services payroll	0.44	0.42
Current Economic Activity	Total GDP (millions)	109.7	106.2
	Total Employment	40,116	6,557
	Total Payroll (millions)	2,178	248
	Average Wage (thousands)	39.72	37.83
	Growth GDP (4 year)	4.64%	3.95%
	Growth Employment (4 year)	2.99%	3.04%
	Growth Payroll (4 year)	7.54%	7.70%
	Growth Wage (4 year)	4.30%	4.22%
	Mean download speed o ered to rms in 2019 (Mbps)	96.44	38.87
Internet Speeds	Mean upload speed o ered to rms in 2019 (Mbps)	85.92	27.86
	Mean download speed o ered to rms in 2014 (Mbps)	49.72	21.53
	Mean upload speed o ered to rms in 2014 (Mbps)	36.73	6.13
	Distance to ARPANET lines in 1979 (km)	191.82	119.57
ARPANET	Had node in 1979	0.01	0.00
	Had connection line in 1979	0.14	0.00

Notes: This table shows the mean and the median of our variables of interest averaging across all counties in continental US. Variables regarding current economic activity and Internet speeds correspond to 2018. Speeds are measured in Megabits per second (Mbps). High-skilled services include information; nance and insurance; real estate services; professional, scienti c and technical services; management of companies and enterprises; educational services; and health care and social assistance. Other services include utilities; construction; wholesale trade; retail trade; transportation and warehousing; administrative and support and waste management and remediation services; arts, entertainment and recreation; accommodation and food services; and other services. GDP growth corresponds to the 4 year growth of the "Real GDP in chained Dollars" series from the BEA. Payroll and wage growth correspond to the 4 year growth of these variables as given by the 2014 and the 2018 County Business Patterns; the 2014 series were de ated using the national CPI from March 2014.

ables that equal one if the county lies along the Canadian or Mexican border since the local economy in such counties might depend on these neighboring countries (Hanson, 1996); a dummy variable that equals one if the county lies along the coast of an ocean or a Great Lake; average slope and elevation, and total land and water area, as geography could determine the costs of infrastructure provision. We also include the nearest distance from the county to a metropolitan statistical area as a measure of market access as in Michaels (2008), population density as a measure of agglomeration economies, and the 4-year population growth rate. Our main parameter of interest is , which identi es the semi-elasticity of a local economic outcome with respect to the quality of Internet provision across U.S. counties.

The estimation of equation (1) does not yield causal estimates due to the endogeneity between the variables of interest. For instance, counties with higher productivity or better amenities might grow faster and have a larger share of GDP coming from high-skilled services as these counties can attract more productive service establishments and workers. Hence, the willingness to pay for better Internet in these counties is higher, and ISPs will provide better Internet. Thus, ^ would be upward biased. On the other hand, counties with better environmental aesthetic amenities (e.g., rivers, mountains, lakes) might attract more high-skilled workers, which would generate incentives for ISPs to provide high speed Internet. At the same time such amenities could make it more di cult to build physical telecommunications infrastructure, thus lowering the quality of local Internet provision. In this case, ^ would be downward biased.

To recover the causal estimate, we follow an instrumental variable approach using the spatial structure of ARPANET in 1979 (years before the Internet had commercial viability) as the source of exogenous variation.[19] As we document in Section 2, the decisions about the location of the nodes, and thus the spatial structure of ARPANET, were determined by whether (i) the researchers in academic institutions were contractors for the DoD through its ARPA agency in the 1960s; (ii) the research agenda of the institution was more focused on networks, instead of batch-processing, which was the main paradigm at the time; and (iii) the work style of the computer science departments was of a collaborative nature.

Hence, due to historical reasons, it is unlikely that the nodes of ARPANET were selected only in counties with the highest productivity levels. Moreover, the presence of a university does not guarantee high levels of productivity. For example, Rust Belt cities that experienced major losses in amenities and productivity have universities e.g., Detroit, Flint, Camden, Youngstown, Toledo, or Dayton. Moreover, we select the status of the network in 1979, when ARPANET only connected DoD contractors and it was still under military management (Abbate, 2000). This selection guarantees that the structure of ARPANET was more closely related to the research interests of

[19] The rst private Internet Service Provider, The World, appeared in the 1989.

the DoD and the work styles and research agendas of academic departments with which it had contracting relationships. In addition, to avoid remaining concerns, we exclude from our main estimations those counties that had an ARPANET node in 1979. Exploiting the spatial structure of ARPANET, our first stage is defined by

$$\log(\text{InternetQuality}_c) = a + b\ ARPANET_c + X_c^{0} + \ _c \tag{2}$$

where $ARPANET_c$ denotes the log of the minimum distance between county's c centroid and an ARPANET connection line, which were presented in Figure 2. Since government reports and maps do not contain information on the exact location of the physical infrastructure used to connect the nodes, straight lines that mimic the maps of ARPANET (e.g., Figure 2) are our proxy for such physical infrastructure. Given that ARPANET had specific network requirements (reliability and delay) to guarantee interconnection quality between computers, the physical infrastructure connecting the nodes must have been of the highest quality at the time. We estimate different specifications of equation (2) using other instruments, including a dummy variable that equals 1 if a connection line crossed county c, and distance categories to the closest line.

Our identification strategy has a similar logic as Duranton et al. (2014) and Duranton (2015), who use historical routes in the U.S. and Colombia as instruments for the location of modern roads. Due to infrastructure cost reasons, it is easier to build contemporary highways following historical paths. Similarly, it's easier to build the modern Internet backbone on historical ARPANET infrastructure (Abbate, 2000; Leiner et al., 1997; McKenzie and Walden, 1991) because a physical Internet backbone requires an underground optic fiber. In addition, it is cheaper for ISPs to provide high-quality Internet to counties near the modern backbone in the same way that it is cheaper for a construction company to physically connect a county closer to the Interstate Highway System. Since counties closer to ARPANET lines—representing the old Internet backbone—are more likely to be closer to the modern Internet backbone, ISPs are more likely to offer higher Internet speeds in such counties.

A potential concern about our instrumental variable is that the ARPANET connecting lines are located along other types of infrastructure, such as highways or railways. Therefore, we would not be capturing the effect of ARPANET but of these ways. In Figure A-5 we present the maps of the ARPANET network in 1979, together with the primary US roads and the railroad infrastructure in 2019. As the figure shows, there seems to be no correlation between these types of infrastructure.

Our strategy also has similarities to Forman et al. (2012), who use the presence of ARPANET nodes in a county as an instrument for Internet investment by businesses, and to Jiang (2022), who uses the location of NSFNET nodes in a DiD framework. Using only the location of the nodes produces a weak instrument since nodes are scarce (less than 1% of counties had a node). Instead, we use the connection between the nodes as an instrument for the average reported Internet speeds of providers to the FCC within a county, which arguably produces a stronger instrument. We do not consider NSFNET nodes, since historical evidence suggests they could be endogenous in a cross-sectional empirical framework that considers geographical areas as the unit of observation.[20] Therefore, we use the ARPANET historical documentation maps from 1979 for our empirical strategy.

We choose 1979, a year when ARPANET had no connections neither shared resources with NSFNET or its predecessor, (Abbate, 2000; McKenzie and Walden, 1991). We consider that our instrumental variable is relevant since ARPANET is the first precursor of the Internet. The main backbone of the Internet relied on ARPANET nodes and lines from the 1970s and 1980s (Leiner et al., 1997). It is likely that modern backbones took advantage of existing underground infrastructure, in the same logic as the path dependence of roads as in Duranton et al. (2014) and Duranton (2015).

As Figure 1 shows, the data supports the relevance of our instrument. Counties crossed by ARPANET lines seem to have better Internet quality as of 2018, both download and upload speeds. We formally test the relevance of our instrument estimating different specifications of equation (2) using OLS. In particular, Table 2 validates the previous results, even after including the vector of control variables. Notice that those counties that had an ARPANET node in 1979 have Internet speeds that on average double those in counties without a node (Panel A); however, less than 1% of counties had a node, which would render a weak instrument as in Forman et al. (2012). Similarly, counties that had an ARPANET connection line in 1979 have download and upload speeds that are 27% and 38% higher in 2018, respectively, relative to counties without a line (Panel B). Notably, counties farther away from a 1979 ARPANET line have lower Internet speeds (Panel C). In particular, a 10% increase in distance is correlated with speeds between that are between 1.3% and 1.8% lower. This negative relationship with distance holds if we consider distance categories instead of a continuous measure (Panel D). These results support the relevance of our instrumental variable.[21]

The structure of ARPANET as an instrumental variable satisfies the exclusion restriction. In other words, ARPANET impacts the local economic structure in U.S. counties exclusively through modern Internet speeds. Based on the history of the network, it is unlikely that the nodes chosen by

[20] NSF encouraged local NSFNET regional networks to provide service to private companies. Due to increasing returns to scale only some universities created NSFNET regional networks. Most likely, the ones where there was an abundance of private companies that needed advanced telecommunications services in the 1980s, which would threaten causal identification in county level cross-sectional regressions. Nonetheless, NSFNET nodes can be used as IV for firm level outcomes like the ones in Jiang (2022).

[21] These results are robust if we use the 1988 structure of ARPANET instead of the 1979 structure, or if we use Internet speeds in 2014 instead of speeds in 2018. Both sets of results are presented in Table A-2 and Table A-3, respectively.

Table 2: OLS Estimates . The Impact of ARPANET on Current Internet Speeds O ered to Businesses

Panel A : County has a node	Mean download speed		Mean upload speed	
	(1)	(2)	(3)	(4)
Had node in 1979	1.314***	0.900***	1.605***	1.083***
	(0.154)	(0.183)	(0.166)	(0.196)
Constant	3.599***	3.881***	3.236***	3.593**
	(0.039)	(1.315)	(0.046)	(1.478)
Panel B : County has a line	Mean download speed		Mean upload speed	
	(1)	(2)	(3)	(4)
Had connection line in 1979	0.352***	0.271***	0.468***	0.380***
	(0.079)	(0.078)	(0.090)	(0.089)
Constant	3.551***	3.504**	3.174***	3.127**
	(0.041)	(1.362)	(0.048)	(1.532)
Panel C : Distance to a line	Mean download speed		Mean upload speed	
	(1)	(2)	(3)	(4)
Distance to Line in 1979	-0.168***	-0.132***	-0.213***	-0.184***
	(0.026)	(0.026)	(0.029)	(0.030)
Constant	5.540***	3.590***	5.691***	3.247**
	(0.299)	(1.388)	(0.344)	(1.564)
Panel D : Distance to a line	Mean download speed		Mean upload speed	
(Categories)	(1)	(2)	(3)	(4)
Has a line	0.638***	0.543***	0.812***	0.743***
	(0.108)	(0.108)	(0.122)	(0.121)
No line and dist 2 (0km; 73:4km]	0.604***	0.544***	0.733***	0.717***
	(0.106)	(0.110)	(0.121)	(0.123)
No line and dist 2 (73:4km; 151:5km]	0.386***	0.358***	0.484***	0.498***
	(0.103)	(0.104)	(0.117)	(0.117)
No line and dist 2 (151:5km; 290:8km]	0.152	0.157	0.159	0.202*
	(0.103)	(0.102)	(0.118)	(0.116)
Constant	3.266***	1.448	2.830***	0.385
	(0.076)	(1.388)	(0.084)	(1.540)
Observations	3,077	3,072	3,077	3,072
Controls		X		X

Notes: This table shows the impact of ARPANET in 1979 structure on modern Internet speed o ered to businesses by county. We display the impact on the mean download and the mean upload speeds. The dependent variables are: a dummy variable that equals 1 if a county has a node (Panel A); a dummy variable that equals 1 if a county has a connection line (Panel B); the log of the distance between the country's centroid and a connection line (Panel C); and distance quartiles from a connection line (Panel D), where counties between 290.8km and 1,118km belong to the omitted category. Panels B, C and D do not include counties with a node in 1979 (32) and the number of observations correspond to these panels. Controls include geographic and economic characteristics. SHAC adjusted standard errors (Conley, 1999) are in parenthesis, with a radius of 28.55km, corresponding to the radius of the metropolitan area in the median of the distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

the DoD through contractor processes, academic research agendas (interested in network research or not), and research work styles in the 1960s (collaborative or uncooperative) are related to productivity or amenities in counties today. Moreover, the presence of a university does not guarantee high levels of productivity and amenities (e.g., Detroit, MI, Camden, NJ). The agendas and collaboration style of computer science departments depend on their internal dynamics, and their contracting relationships depend on professional networks. Moreover, national defense computing research locations were decided purely based on the interests of DARPA and the DoD. Lastly, to guarantee that the exclusion restriction is satisfied, we drop from our samples any counties with ARPANET nodes.

Our instrumental variable is unlikely to be correlated with omitted variables that influence local economic structure; that is, it satisfies the exogeneity assumption. Using ARPANET lines connecting the nodes, instead of the nodes themselves supports this assumption. Even if there are remaining concerns that the ARPANET nodes are located in counties with high amenities or productivity, the lines that connect them are likely to be exogenous to such unobserved factors. For example, an ARPANET line connects the Argonne National Laboratory in Argonne, IL, to the Wright-Patterson Air Force Base in Green County, OH. Such line passes over several Indiana counties, including Wabash County. Even under the strong assumption that military nodes in Argonne, IL, or Green County, OH, were selected due to high productivity of the counties, the line passing over Wabash County, IN, is exogenous to the productivity of these counties with nodes. The lines themselves predict the location of the ARPANET backbone, which itself predicts the geographical layout of the equipment that forms the modern Internet backbone. In Table A-4, we show that even though the location of the nodes in 1979 is strongly correlated with the share of employment in high-skilled and business services in the 1970s, whether a county is crossed by a connection line is not correlated by such shares.

Our last concern for identification is whether our strategy satisfies the Stable Unit Treatment Value Assumption (SUTVA). In our case, SUTVA implies that potential outcomes for a given county respond only to its own Internet quality and are unrelated to the treatment status of other counties. Due to engineering related reasons, the local provision of Internet does not directly influence another county's provision because ISPs are in charge of the last mile and can differentiate the Internet service quality at granular levels.[22] Such targeting capacity by ISPs dissipates concerns

[22] The provision of Internet has some similarities to electricity or water, cases in which the provider decides the quality of the service in the last mile, hence the quality of the service provision can be targeted (counties, neighborhoods, blocks, buildings, etc.). For example, quality can vary across tracts within the same city or even across neighborhoods within the same county. Some illustrative cases are Detroit (Wayne County) today or Chattanooga (Hamilton County) in the past, cities where national commercial ISPs did not provide high-speed Internet in specific neighborhoods, even though the rest of the city had access to broadband services. The importance of the last mile has been documented in detail by journalists, economists and non-profits (Lobo et al., 2008; Peñarroyo et al., 2022; Thornton and Mars, 2022).

about SUTVA violations.

5 Results

In this section, we present the main results of the paper. In particular, we show the impact that better provision of Internet has on local economic activity. Moreover, we show its impacts on local structural transformation. Finally, we present the results regarding the two main proposed mechanisms behind such impact, together with the impact of internet quality of earnings inequality within counties.

5.1 Internet Quality and Local Economic Growth in the Short-run

To study the e ect of better provision of Internet on local economic growth, we start by estimating equation (1) using as dependent variables the 4-year growth rate of the county's real GDP, total annual payroll, total employment and average wages, de ned as the ratio of the two previous variables. We estimate such equations using two-stage least squares (2SLS). For our rst stage we estimate equation (2) using the (log) distance between a country's centroid and an ARPANET connection line in 1979. In Table 3 we present the results of such estimation, using the (log) mean download speed o ered to businesses in 2018 in Panel A and the (log) mean upload speed o ered to businesses in 2018 in Panel B. We report two types of standard errors: robust, and spatial heteroskedasticity and autocorrelation consistent standard errors (Conley, 1999), together with the rst-stage Kleibergen-Paap F-test for weak instruments (Kleibergen and Paap, 2006).

We highlight three results from Table 3. First, better Internet has a slight positive e ect on economic growth as shown in column 1. In particular, doubling Internet speeds can lead to an increase of 2.1 to 3.7 percentage points (pp) in the 4-year GDP growth rate, which is a relevant magnitude if we consider that the interquartile range (IQR) for this rate is 14.8 percentage points. Although robust to di erent speci cations, it is worth noticing that this estimate is somewhat imprecise. Second, notice in column 3 that there is a positive and signi cant e ect on short-run employment growth: doubling the quality of Internet increases the 4-year employment growth by 2.8 to 4.9 pp, which has an IQR of 12.5. That is, faster Internet leads to job growth and employment reallocation towards those regions with better access to it.

Third, since there is not a change in the growth rate of total payroll (column 2), results from column 4 suggest that faster Internet leads to a reduction in the growth rate of wages{though, not necessarily a decrease in wage levels. In particular, doubling the speed of Internet decreases the growth rate of wages by 1.6 to 2.8 percentage points. In principle, these results seem to di er from in Forman et al. (2012), who nd a null impact of better Internet on average wages. We do not

Table 3: 2SLS Estimates . E ect of Better Internet Speed O ered to Businesses on Local Economic Aggregates

Panel A: Download Speed	Growth rate of			
	GDP	Annual Payroll	Employment	Average Wage
	(1)	(2)	(3)	(4)
Log(Mean Download Speed	0.037*	0.024	0.049***	-0.028**
O ered to Businesses)	(0.021)	(0.018)	(0.016)	(0.011)
	[0.022]	[0.019]	[0.017]	[0.012]
Constant	0.074	0.211	0.281***	-0.052
	(0.174)	(0.131)	(0.096)	(0.081)
	[0.182]	[0.142]	[0.111]	[0.084]
FS F-Test	44.446	46.000	45.192	45.192
Panel B: Upload Speed	Growth rate of			
	GDP	Annual Payroll	Employment	Average Wage
	(1)	(2)	(3)	(4)
Log(Mean Upload Speed	0.021*	0.013	0.028***	-0.016**
O ered to Businesses)	(0.012)	(0.010)	(0.009)	(0.006)
	[0.012]	[0.011]	[0.009]	[0.007]
Constant	0.166	0.271**	0.404***	-0.123
	(0.169)	(0.125)	(0.090)	(0.076)
	[0.176]	[0.139]	[0.106]	[0.079]
FS F-Test	56.421	58.614	57.594	57.594
Observations	3,022	3,047	3,037	3,037
Controls	X	X	X	X

Notes: This table shows the results of regression of di erent local economic outcomes on current Internet o ered to businesses (both mean download and upload speeds). We use as an instrument the county centroid's distance to an ARPANET connection line in 1979. We include geographic and economic controls in our rst stage. Average wages are measure as the ratio between total payroll and total employment. Regressions do not include counties with ARPANET nodes in 1979. The FS F-Test corresponds to the Kleibergen & Paap F-test for weak instruments. Robust standard errors are in parentheses and SHAC adjusted standard errors (Conley, 1999) are in brackets, with p-values * $p < 0{:}10$, ** $p < 0{:}05$, *** $p < 0{:}01$. The radius for SHAC errors is set at 28.55km, corresponding to the radius of the metropolitan area in the median of the distribution.

consider both set of results to contradict each other as they analyze the e ects of rms' investment on Internet shapes county labor markets in levels during the ve year period when the adoption of the technology was in its initial stages. Di erently, we evaluate quality of Internet provision on growth rates 20 years after, in a period when the technology was already widely adopted. The short-run impacts on wage growth might occur if better Internet led to a positive labor supply shock in a county in this period in the absence of a shift in labor demand.

Moreover, the rst stage KP F-statistic for weak instruments shows a value of around 45 in Panel

A and around 58 in Panel B. Thus, the distance to an ARPANET line in 1979 is not a weak instrument for current Internet speeds. Moreover, our results are robust to speci cations where we (i) further restrict the spatial correlation patterns of the error term (Table A-5); (ii) use the 1988 ARPANET structure as our instrumental variable (Table A-6); (iii) include all counties in the continental US (Table A-7); (iv) exclude 6 Metropolitan Areas with a large concentration of nodes (Table A-8);[23] (v) use median download speed as a measure for Internet quality (Table A-7).

Finally, we estimate equation (1) using OLS to grasp the direction of the original bias. As can be seen in Table A-9, coe cients from the OLS estimation are negative and some of them signi cantly di erent from zero; that is, they are lower than those in Table 3. This comparison suggests that ignoring the endogeneity problem of these regressions would lead to downward biased estimators, as hypothesized in Section 4. In the following section, we discuss how the quality of communication infrastructure can also a ect local structural transformation and the distribution of economic activity across US counties.

5.2 Internet Quality and Local Structural Transformation

We want to understand if and how a better provision of communications infrastructure alters the sectoral economic structure of counties; that is, if it leads to local structural transformation. Since research has already shown that better communication infrastructure can lead to a better transmission of ideas (Carlino et al., 2007), it would be natural to conclude that better communication infrastructure favors local activity in sectors where ideas are key for the growth of the sector.

To explore local structural transformation, we estimate equation (1) using as dependent variable the share of employment and the share of GDP in each sector within a county in 2018, the quality of Internet in 2018 as our variable of interest, and the (log) distance to an ARPANET connection line in 1979 as its instrument. We start by analyzing these shares at the 2-digit NAICS sector. We present the results of these regressions in Figure 3, where we show the point estimates and their 95% con dence intervals. The horizontal axis of this gure presents the 2-digit NAICS sector codes sorted by the size of the point estimate; the full name of each sector can be found in Table A-1.

Results from Panel A in Figure 3 show that faster Internet download speeds have a positive impact in the employment shares of ve 2-digit sectors: (i) Educational Services; (ii) Administrative, Support, Waste Management, and Remediation Services; (iii) Professional, Scienti c, and Technical Services; (iv) Arts, Entertainment, and Recreation; and (v) Management of Companies and Enterprises. In addition, it also has a small but signi cant positive e ect on Real Estate and Rental

[23] From this speci cation, we exclude the following metropolitan areas: Boston-Cambridge-Newton, Los Angeles-Long Beach-Anaheim, New York-Newark-Jersey City, San Francisco-Oakland-Hayward, San Jose-Sunnyvale-Santa Clara, and Washington-Arlington-Alexandria.

Figure 3: Internet and Local Structural Transformation

(a) Employment Shares

(b) GDP Shares

Note: these gures show the 2SLS estimates that measure the causal impact of better Internet provision (measured as faster download speeds o ered by Internet providers to businesses) on the share of employment or GDP on 2-digit NAICS sectors. We use as instrumental variable the distance to an ARPANET line in 1979 , together with geographic and economic characteristics. The horizontal axis shows the sectors sorted by the size of the estimate. The full sector names are shown in Table A-1.

and Leasing and Utilities. Notice that these positive effects are concentrated in the sectors denominated Skilled-Scalable Services by Eckert et al. (2020), or Prime Services by Ahlfeldt et al. (2020). As discussed in these papers, such services are intensive in high-skilled workers and information services, and have been responsible for the relatively faster growth of larger cities in the past two decades.

On the other hand, faster download speeds affects negatively employment shares in three aggregate sectors: (i) Health Care and Social Assistance; (ii) Retail Trade; and (iii) Finance and Insurance. These effects are statistically significant at the 95% confidence level. Notice that these three sectors can be found in almost every town and city across the U.S., which could suggest that regions with worse quality of Internet might have only service firms to serve the local market (e.g., physician practice, social assistance offices, physical bank, retail store). It is possible that in rural and low-populated counties establishments such as retail stores, physician practices and government social assistance offices might be major employers. Hence, the arrival of firms that use ICT are more likely to impact the share of employment of such sectors. Moreover, we find heterogeneous impacts of better Internet on Healthcare and Social Assistance and Finance and Insurance when we consider the 3-digit NAICS sub-sectors in these groups of industries (see table 5 for more details). In addition, the effect on Manufacturing (31) also appears to have a negative point estimate, but it is quite imprecisely estimated, which could be explained by the heterogeneous effects of Internet quality on different manufacturing sub-sectors. This sectoral ranking is robust to using (i) payroll shares instead of employment shares as our dependent variable (Figure A-6); (ii) upload speeds as the measure of Internet quality (Figure A-7); or (iii) the 1988 ARPANET structure for our IV strategy (Figure A-8).

For more specific magnitudes, consider the point estimate for Professional, Scientific, and Technical Services ($\hat{\beta} = 0.012$), and the median county, which has an average download speed of 39 Mbps and 2.6% of employment in this sector. Our estimates suggest that, if this county improves its Internet quality to 135 Mbps (a 350%, equivalent to the county in the 75th percentile), this county could see an increase in the relative importance of this sector of 4.2 pp, going from 2.6% to 6.8% of the total county level employment. Contrarily, consider the point estimate for Retail Trade ($\hat{\beta} = -0.015$). If this same median county experiences that 350% in download speeds, it could see a reduction in the relative importance of retail employment of 5.3pp, going from 15.4pp to 10.1pp.

In Panel B from Figure 3, we present the results of the same estimations but using sectoral GDP shares as the main dependent variables. In this case, the positive effects for high-skilled services hold: four of the five services in this category are in the top 6 of the sectors positively affected by better Internet quality, including Finance and Insurance. However, we would like to note two important differences compared to the previous figure. First, manufacturing appears to be positively affected in this case, but its confidence intervals remain quite broad and the effects are not significant. Second, there is a sizeable negative effect of Internet quality on the relative importance of Agriculture. We also observe negative impacts on *Wholesale Trade*; *Health Care and Social Assistance*; and *Transportation and Warehousing*.

To disentangle the impacts of better Internet provision on sectoral outcomes at the county level, we re-estimated the model using employment shares for each one of the the 92 NAICS sectors at the 3-digit level. Given the large dimensionality of these results, we report them separately for Agriculture, Mining and Manufacturing sub-sectors in Table 4, and for all Services sub-sectors in Table 5. Moreover, we categorize all 3-digit sectors depending on whether their point estimate is significant at the 95% confidence level, positive or negative. We display the table with the estimates and the standard errors for each sub-sector in the online appendix.[24]

With respect to manufacturing, we highlight two findings. First, only *Wood Product Manufacturing* is negatively and significantly affected by higher Internet quality. On the other hand, only four sub-sectors are positively impacted. In particular, the following manufacturing 3-digit NAICS sub-sectors are positively impacted by better download speeds: *Textile Mills*; *Fabricated Metal Product Manufacturing*; *Electrical Equipment, Appliance, and Component Manufacturing*; and *Miscellaneous Manufacturing*. These impacts could be explained by the fact that better communication technologies benefits industries with just-in-time manufacturing. Differently, none of the agricultural or mining 3-digit NAICS sub-sectors are positively affected by better Internet. In particular, the effect on all of them is null, except for activities belonging to *Oil and Gas Extraction* which requires good quality communications, particularly for drilling activities.

For the case of services, our findings show interesting patterns. First within information services, the most favoured sub-sectors are the ones related to Internet and communication: *Telecommunications*, and *Internet service providers, Web search data processing services*. Second, the disaggregated impacts on *Finance and Insurance* and *Healthcare and Social Assistance* are illustrative of the local structural change generated by the Internet. At the aggregate level we found evidence that better Internet quality has a negative impact on these two sectors, as shown in Figure 3.

However, when we consider 3-digit NAICS sub-sectors, we find that the reduction in the employment share of *Finance and Insurance* is driven by *Credit Intermediation and Related Activities*, which experiences a significant negative impact. This sub-sector includes commercial banking and all its branches located across all the country, from the largest cities to the most rural towns. Differently, the impact of better Internet provision on *Securities, Commodity Contracts, and Other Financial*

[24] The sub-sectors Crop Production (111), Animal Production (112), Rail Transportation (482), Postal Service (491), Internet Publishing and Broadcasting (516), and Private Households (814) are not included in the CBP.

Table 4: Direction of 2SLS Estimates . Impact of Internet Speed O ered to Businesses on Agriculture and Manufacturing Subsectors

Positive Impacts (95%)	Zero or Statistically Insigni - cant Impacts	Negative Impacts (95%)
313 - Textile Mills	113 Forestry & Logging	211 - Oil & Gas Extraction
332 - Fabricated Metal Product Manufacturing	114 - Fishing, Hunting, Trapping	321 - Wood Product Manufact.
335 - Electrical Equipment, Appliance, & Component Manufact.	115 - Support Activities for Agriculture & Forestry	
339 - Miscellaneous Manufacturing	212 - Mining (except Oil & Gas)	
	213 - Support Act. for Mining	
	221 - Utilities	
	236 - Construction of Buildings	
	237 - Heavy & Civil Engineering Construction	
	238 - Specialty Trade Contractors	
	311 - Food Manufacturing	
	312 - Beverage & Tobacco Product Manufacturing	
	314 - Textile Product Mills	
	315 - Apparel Manufacturing	
	316 - Leather & Allied Product Manufacturing	
	322 - Paper Manufacturing	
	323 - Printing & Related Support Activities	
	324 - Petroleum & Coal Products Manufacturing	
	325 - Chemical Manufacturing	
	326 - Plastics & Rubber Products Manufacturing	
	327 - Nonmetallic Mineral Product Manufacturing	
	331 - Primary Metal Manufact.	
	333 - Machinery Manufacturing	
	334 - Computer & Electronic Product Manufacturing	
	336 - Transportation Equipment Manufacturing	
	337 - Furniture & Related Product Manufacturing	

Notes: We classify the NAICS 3-digit sub-sectors according to 2SLS estimates. The estimates are obtained by a regression of the share of employment in each sector on the quality of Internet provision (measured as higher download speeds o ered by Internet providers to businesses). The IV is the distance of a county to an ARPANET line in 1979. We include geographic economic characteristics in the rst stage. The classi cation depends on whether the 2SLS estimate was signi cant at the 95% con dence level. Regressions do not include counties with ARPANET nodes in 1979. Exact estimates are provided on request.

Table 5: Direction of 2SLS Estimates . Impact of Internet Speed on Services Subsectors

Positive Impacts (95%)	Zero or Statistically Insigni - cant Impacts	Negative Impacts (95%)
425 - Wholesale Electronic Markets & Agents & Brokers	423 - Merchant Wholesalers, Durable Goods	424 - Merchant Wholesalers, Nondurable Goods
442 - Furniture & Home Furnishings Stores	441 - Motor Vehicle & Parts Dealers	447 - Gasoline Stations
448 - Clothing & Clothing Accessories Stores	443 - Electronics & Appliances Stores	484 - Truck Transportation
451 - Sporting Goods, Hobby, Book, & Music Stores	444 - Building Material & Garden Equipment & Supplies Dealers	522 - Credit Intermediation & Related Activities
485 - Transit & Ground Passenger Transportation	445 - Food & Beverage Stores	621 - Ambulatory Health Care Services
493 - Warehousing & Storage	446 - Health & Personal Care St.	624 - Social Assistance
517 - Telecommunications	452 - General Merchandise Stores	722 - Food & Drinking Services
518 - ISPs, Web Search Portals, and Data Processing Services	453 - Miscellaneous Store Retailers	
519 - Other Information Services	454 - Nonstore Retailers	
523 - Securities, Commodity Contracts & Other Financial Invest.	481 - Air Transportation	
532 - Rental & Leasing Services	483 - Water Transportation	
551 - Management of Companies & Enterprises	486 - Pipeline Transportation	
561 - Administrative & Support Services	487 - Scenic & Sightseeing Transportation	
611 - Educational Services	488 - Support Activities for Transportation	
622 - Hospitals	511 - Publishing (except Internet)	
711 - Performing Arts, Spectator Sports, & Related Industries	512 - Motion Picture & Sound Recording Industries	
812 - Personal & Laundry Services	515 - Broadcasting (exc. Internet)	
	521 - Monetary Authorities	
	524 - Insurance Carriers & Related Activities	
	525 - Funds, Trusts, & Other Financial Vehicles	
	531 - Real Estate	
	533 - Lessors of Non nancial Intangible Assets	
	541 - Professional, Scienti c, & Technical Services	
	562 - Waste Management & Remediation Services	
	623 - Nursing & Residential Care Facilities	
	712 - Museums, Historical Sites, & Similar Institutions	
	713 - Amusement, Gambling, & Recreation Industries	
	721 - Accommodation	
	811 - Repair & Maintenance	
	813 - Religious, Grantmaking, Civic, Professional, & Similar	

Notes: We classify the NAICS 3-digit sub-sectors according to 2SLS estimates. The estimates are obtained by a regression of the share of employment in each sector on the quality of Internet provision (measured as higher download speeds o ered by Internet providers to businesses). The IV is the distance of a county to an ARPANET line in 1979. We include geographic economic characteristics in the rst stage. The classi cation depends on whether the 2SLS estimate was signi cant at the 95% con dence level. Regressions do not include counties with ARPANET nodes in 1979. Exact estimates are provided on request.

Investments is positive and statistically significant. This sub-sector includes trading and other high-tech finance activities. Similarly, the negative impact of faster Internet on Healthcare and Social Assistance comes from the impacts on Social Assistance and Ambulatory Healthcare Services. These results contrast with the null impacts on Nursing and Residential Care Facilities, and the positive and significant impacts on Hospitals.

Third, within the 2-digit NAICS sector 56 that corresponds to Administrative and Support and Waste Management and Remediation Services the impacts are also heterogeneous across sub-sectors. While Administrative and Support Services is positively affected by higher Internet speeds, Waste Management and Remediation Services is not impacted. Fourth, we also find heterogeneity within the 2-digit NAICS sub-sectors Retail Trade and Wholesale Trade. Some sub-sectors, such as Wholesale Electronic Markets, Agents & Brokers or Warehousing and Storage, are positively affected by better Internet provision. Contrarily, sub-sectors like Merchant wholesalers of non-durable goods, Gasoline stations, Truck transportation, and Food Services and Drinking Places, face a negative impact.

5.3 Mechanism 1: Input-Output Linkages

We explore input-output linkages as a potential mechanism for the observed changes in local structural transformation. Specifically, we consider the case of industries that purchase a significant amount of inputs from ICT sectors in the economy. If ISPs offer higher quality Internet across counties, this can increase the quality and quantity of the inputs purchased from ICT industries. One example is audiovisual communication via Internet. If firms can have access to higher quality of Internet, they are more likely to purchase Skype or Zoom for their business communications, which is an improvement relative to mobile phone calls. A second example is data storage services. High-speed Internet allows firms to purchase relatively inexpensive cloud services, which are cheaper compared to purchasing servers and hiring computer technicians just to obtain data storage capacity.

We use the input-output accounts from the Bureau of Economic Analysis to quantify the dependence of every sector on Information and Communication Technologies. Specifically, we use the "Total Requirements Table" that shows the inputs that are required directly and indirectly to deliver a dollar of output to final uses. First, we select the 2-digit NAICS sectors Broadcasting and Telecommunication and Information Services and Data Processing Services as the main sectors representing ICT technologies. Second, we rank sectors depending on their absolute dependence and relative dependence on ICT inputs as per the "Total Requirements Table".[25] Third, we rank

[25] For absolute dependence, we use the coefficients from the total requirements table. For example to produce $1 of Computer and electronic products we need $0.0297021 of direct and indirect inputs from Broadcasting and telecommunications. For relative dependence, we estimate the ratio of the value of inputs coming from ICT sectors

sectors in quartiles depending on these measures, with quartile 1 containing those sectors at the top quartile of the distribution.

We estimate the e ect of faster Internet on the share of employment at the county level in all sectors within in the same quartile. Our ndings show that the share of employment in those sectors with higher dependence on inputs related to ICT-sectors are more sensitive to improvements on Internet quality, as shown in Table 6. In fact, when we consider the absolute dependence on ICT-inputs, we observe an increasing e ect across quartiles, ranging from a null e ect in quartile 4 to a positive impact in quartile 1 (^ = 0:047). Following the same interpretation as in Section 5.2, the latter coe cient represents an increase in 16.5pp in the share of employment in sectors belonging to this category, following an increase of 350% in the average Internet speed o ered to businesses within a county (the di erence between the counties in the percentiles 50th and 75th of the Internet speed distribution across U.S. counties). Our results are robust to using absolute or relative relevance, or download or upload speeds. They are also robust to using payroll shares instead of employment, as shown in Table A-10. Moreover, we use a di erent ranking based on tertiles or deciles, instead of quartiles, and the ndings remain similar (Table A-11). Thus, we conclude that our results are driven by industry linkages between industries sensitive to Internet quality and ICT-related inputs.

5.4 Mechanism 2: ICT Related Occupations

Local structural transformation could be driven by an in ow of workers in occupations that might bene t from the use of ICTs, as the Rybczynski theorem would predict. For instance, rms in sectors that bene t the most from higher Internet quality, such as business and administrative services, could decide to locate in places with a relatively higher abundance of workers in ICT-driven occupations, e.g., computer scientists and administrative support workers. This can lead to a virtuous cycle in which these growing sectors end up attracting more ICT-driven workers.

To test this hypothesis, we use data from the 5-year American Community Survey for 2013-2017. The dataset contains the number of workers in di erent occupation categories for each county (Manson et al., 2022). For each location, we compute the total number and the share of workers in occupations that are more prone to using Internet. In particular, we include: management, business, and nancial occupations, computer and mathematical occupations, architecture and engineering occupations, and o ce and administrative support occupations. Using these data, we estimate regressions equivalent to those in Section 5.2, that is, we regress the log (and the share) of workers in ICT-driven occupations on current Internet speeds o ered to businesses, instrumented with the county distance to an ARPANET connection line.

with respect to the total value of inputs.

Table 6: 2SLS Estimates, Heterogeneity Analysis . Impacts of Internet Speeds on Sectors due to Industry Linkages with ICT-related Industries - Dependent Variable: Employment Shares

Panel A	Absolute relevance of ICT-inputs			
	Quartile 4	Quartile 3	Quartile 2	Quartile 1
	(1)	(2)	(3)	(4)
Log(Mean download speed	-0.004	0.013	-0.061**	0.047**
O ered to Businesses)	(0.008)	(0.015)	(0.025)	(0.020)
Constant	-0.036	0.037	1.265***	-0.110
	(0.064)	(0.102)	(0.274)	(0.145)
Panel B	Absolute relevance of ICT-inputs			
	Quartile 4	Quartile 3	Quartile 2	Quartile 1
	(1)	(2)	(3)	(4)
Log(Mean upload speed	-0.003	0.009	-0.044**	0.034**
O ered to Businesses)	(0.006)	(0.010)	(0.017)	(0.014)
Constant	-0.041	0.054	1.187***	-0.049
	(0.060)	(0.093)	(0.257)	(0.143)
Panel C	Relative relevance of ICT-inputs			
	Quartile 4	Quartile 3	Quartile 2	Quartile 1
	(1)	(2)	(3)	(4)
Log(Mean download speed	0.002	-0.031	-0.026	0.050**
O ered to Businesses)	(0.009)	(0.020)	(0.019)	(0.022)
Constant	0.049	0.259	0.890***	-0.041
	(0.071)	(0.175)	(0.177)	(0.168)
Panel D	Relative relevance of ICT-inputs			
	Quartile 4	Quartile 3	Quartile 2	Quartile 1
	(1)	(2)	(3)	(4)
Log(Mean upload speed	0.002	-0.022	-0.019	0.036**
O ered to Businesses)	(0.007)	(0.014)	(0.014)	(0.016)
Constant	0.052	0.220	0.856***	0.023
	(0.066)	(0.170)	(0.167)	(0.166)
Observations	3,054	3,054	3,054	3,054
Controls	X	X	X	X

Notes: Quartile 1 are sectors with high dependence on ICT inputs. Quartile 4 are sectors with low dependence. This table shows the results of estimating a regression of employment shares within di erent sectoral categories on the quality of internet o ered to businesses in 2018 (either the average download or upload speed), instrumented with the distance of a county centroid to an ARPANET line in 1979. We include geographic and economic controls in the rst stage. Sectoral categories are due to industry linkages. We split the sectors in quartiles de ned by rankings built using the absolute or relative dependence of a 2-digit NAICS sector on ICT-related inputs. We compute such quartiles using input-output tables from the BEA. Regressions do not include counties with ARPANET nodes in 1979. SHAC adjusted standard errors (Conley, 1999) are reported in parenthesis, with a radius of 28.55km, corresponding to the radius of the metropolitan area in the median of the distribution. * p < 0:10, ** p < 0:05, *** p < 0:01

The results of these regressions suggest that better Internet provision in a county increases the number of workers in occupations that might bene t from the use of ICTs. In particular, estimates from Panel A in Table 7 imply that a 1% increase in download speeds lead to a 1.7% increase in the total number of workers in ICT-intensive occupations. These results also hold if, instead of occupations, we consider workers with an educational attainment higher than a bachelor's degree: a master's, a professional school or a doctorate degree. Results from Panel B suggest that a 1% increase in download speeds lead to a 2% increase in the total number of workers with more than a bachelor's degree. Better Internet also increases the relative importance of both types of workers within a county, as we show in column (1) of both panels. Table A-12 shows that these results are robust when we use upload speeds to measure the quality of Internet provision o ered to businesses.

Table 7: 2SLS Estimates . The Impact of Internet on the prevalence of ICT-intensive workers

Panel A: Workers in ICT related occupations		
	Share	Logs
	(1)	(2)
Log (Mean Download Speed	0.013*	1.664***
O ered to Businesses)	(0.007)	(0.312)
Constant	0.119**	-1.623
	(0.053)	(2.339)
Panel B: Workers with more than a Bachelor's Degree		
	Share	Logs
	(1)	(2)
Log (Mean Download Speed	0.034***	1.977***
O ered to Businesses)	(0.007)	(0.372)
Constant	0.176***	-0.604
	(0.057)	(2.760)
Observations	3,072	3,071
Controls	X	X

Notes: This table shows the results of regressions studying the impact of mean download speeds on the prevalence of ICT-intensive (Panel A) or high-skilled workers (Panel B), both using shares and logs. We use as instrumental variable the distance of a country centroid to an ARPANET connection line in 1979. We include geographic and economic controls in the rst stage. Data for the dependent variables comes from the NHGIS 2013-2017 5-year ACS (NHGIS code: AH04 and AH3S). Regressions do not include counties with ARPANET nodes in 1979. SHAC adjusted standard errors (Conley, 1999) are reported in parenthesis, with a radius of 28.55km, corresponding to the radius of the metropolitan area in the median of the distribution. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

These results are consistent with the Rybczynski theorem from the HOV model: higher endowments of one factor lead to a more than proportional expansion of the output in the sector that uses such factor intensively, and a decline in the output of the other sector. In our case, since higher quality of Internet increases the number of workers who use ICTs more frequently, the sectors that expand are those who hire these workers more intensively. To support our argument, in Figure 4 we show the correlation between the size of the estimates that measure the impacts of better Internet speeds on GDP sectoral shares (from Figure 3) and the share of ICT-related occupations for 2-digits NAICS sector. This correlation is positive, hence sectors that employ more ICT-related workers see a larger expansion of their output share due to the provision of faster Internet.

Figure 4: Estimates of the Impact of Faster Internet on Sectoral Economic Activity vs. Share of ICT-Related Occupations

Note: this gure shows the correlation between the 2SLS estimates that measure the causal impact of better Internet provision (i.e., faster download speeds o ered by ISPs to businesses) on the share of GDP on 2-digit NAICS sectors (from Figure 3) and the share of workers in ICT-related occupations for each sector: management, business, and nancial occupations, computer and mathematical occupations, architecture and engineering occupations, and o ce and administrative support occupations. The gure excludes the estimates for agriculture.

We also test for the presence of capital-skill complementarities, for the particular case of computer equipment. We estimate correlations at the county level between a proxy of a county's use of computer equipment and its share of ICT-related workers, and average Internet download speeds o ered to business. Since we do not have data on the sales or use of computers, we use the county's total payroll in two 4-digit sectors { Electronics and Appliance Stores and Professional and Commercial Equipment and Supplies Merchant Wholesalers { as proxies for the use or sales of home and o ce computers, respectively. We show the results of these correlations in Figure 5. In the top panel, we observe that counties with a larger share of workers in ICT-driven occupations have more activity in those sectors related to the retail and wholesale of computers and electronic equipment. Similarly, the bottom panel shows that Internet speeds in a county are positively correlated with the size of these sub-sectors. These results point to presence of capital-skill complementarity in counties with better communications infrastructure.

Figure 5: Use of PC in a County (Proxy), Workers in ICT-driven Occupations and Internet Speeds

(a) Use O ce PC (Proxy) vs. ICT Workers

(b) Use Home PC (Proxy) vs. ICT Workers

(c) Use O ce PC (Proxy) vs. Internet Speed

(d) Use Home PC (Proxy) vs. Internet Speed

Note: These gures show the correlation between the use of personal computers, the share of workers in ICT-related occupations in a county, and Internet quality (measured as the mean download speed o ered by ISPs to rms). We use as proxy for \Use of O ce PC" in a county the total payroll of the sub-sector Electronics and Appliance Stores (NAICS code 4232), and for \Use of Home PC" in a county the total payroll of the sub-sector Professional and Commercial Equipment and Supplies Merchant Wholesalers (NAICS code 4431). Due to con dentiality restrictions, the CBP data only reports this information for 893 counties for sector with NAICS code 4234, and for 1322 counties for the sector with NAICS code 4431.

5.5 Internet Quality and Inequality

Our previous ndings show that better Internet rises the share of high-skilled workers in a county. If the demand for these types of workers is increasing, it is possible that Internet speeds could a ect local inequality through a relative increase in the wages of high-skilled workers. To formally test this hypothesis, we retrieve data from the 5-year 2015-2019 ACS on the median yearly earnings by educational attainment at the county level (Manson et al., 2022). Using these data, we compute average wage for workers in 3 distinct categories: individuals with less than a bachelor's degree[26], with a bachelor's degree, or with a graduate or a professional degree. Afterward, we estimate the same regression as in Table 7, but using the log of earnings as the main dependent variable.

Table 8: 2SLS Estimates. The Impact of Internet Speed O ered to Businesses on the Earnings of Workers by Educational Attainment

Panel A	Earnings of Workers with		
	Less than Bachelor's	Bachelor's	Graduate or Professional
	(1)	(2)	(3)
Log (Mean Download Speed O ered to Businesses)	0.031	0.157***	0.165***
	(0.020)	(0.035)	(0.038)
Constant	9.622***	10.285***	10.733***
	(0.199)	(0.271)	(0.268)
Panel B	**Earnings of Workers with**		
	Less than Bachelor's	Bachelor's	Graduate or Professional
	(1)	(2)	(3)
Log (Mean Upload Speed O ered to Businesses)	0.022	0.112***	0.116***
	(0.014)	(0.022)	(0.023)
Constant	9.661***	10.492***	10.919***
	(0.188)	(0.230)	(0.229)
Observations	3,071	3,048	3,001
Controls	X	X	X

Notes: This table shows the results of regressions studying the impact of internet speeds on the earnings of workers by educational attainment. Internet speeds are measured with mean download speeds (Panel A) or mean upload speeds (Panel B) o ered by Internet service providers to businesses in the county. The instrument is the county centroid's distance to an ARPANET connection line in 1979. We include geographic and economic controls in the rst stage. Data for the dependent variables comes from the NHGIS 2015-2019 5-year ACS (NHGIS code: AMFR). Regressions do not include counties with ARPANET nodes in 1979. SHAC adjusted standard errors (Conley, 1999) are reported in parenthesis, with a radius of 28.55km, corresponding to the radius of the metropolitan area in the median of the distribution. * p < 0:10, ** p < 0:05, *** p < 0:01

Results from Table 8 show that better quality of Internet o ered to businesses leads to an increase in the earnings of workers with at least a bachelor's degree: a 1% increase in the quality of Internet increases median earnings by between 0.11% to 0.17%. On the other hand, better Internet seems to have a small{but insigni cant{ e ect on the earnings of individuals with less than a bachelor's degree. Thus, our results show that the reduction in communication costs, induced by better Internet provision, leads to a rise in local inequality. These results are in line to previous ndings

[26] This category includes individuals with less than or with a high school degree (including equivalencies), some college, or an associate's degree.

on how reductions in trade costs can increase local inequality (Hanson, 1996; Goldberg and Pavcnik, 2007; Topalova, 2010; Verhoogen, 2008).

6 Conclusions

Higher quality communication infrastructure can have major impacts on local economic outcomes. Differently from other types of infrastructure, better communications infrastructure can increase the transmission speed of ideas between individuals at relatively low costs. Moreover, the reduction in trade costs derived from better communications infrastructure is different compared to the decrease in trade costs caused by roads or railroads. Research has also shown that improvements in communications access can lead to more innovation, entrepreneurship, larger firms, higher housing prices and higher trade flows. However, their effects on local economies and local structural transformation remained understudied.

In this paper, we document how differences in the quality of communication infrastructure influence the structural transformation of local economies, their short-run economic growth, and their wage inequality. In particular, we use economic and Internet provision data for all counties in the United States for 2018, to explore the relationship between better these variables. For identification, we use the distance from each county to one of the lines connecting ARPANET nodes, a network that was the precursor of the Internet and later became the backbone of Internet in its initial stage. These connection lines represent the actual telecommunications equipment installed to connect the old network nodes. We obtain such information from historical government reports documenting the early history of Internet and combine them with different geographic and economic characteristics.

Our estimates suggest that if a county improves the Internet it provides to businesses, its short-run GDP and employment growth increases. Moreover, better Internet favours local employment and GDP in high-skilled services, such as, management, information, professional services and educational services. Nonetheless, better Internet also leads to a decrease in the relative importance of other sectors, such as, retail, food services, health services, and financial services. Even though the negative effect on the financial or the health services might seem puzzling, they appear natural when we explore a more disaggregated sectoral structure. Specifically, better Internet reduces the county share of employment in *Credit Intermediation and Related Activities*, which mostly includes physical banks, while it increases employment in those subsectors related with high-tech financial products. Similarly, faster Internet reduces the local share of employment in social assistance and ambulatory healthcare services, but increases the employment shares of *Hospitals*.

Two mechanisms explain our results. First, we find that industry linkages account for some of the observed changes in local structural transformation. Specifically, we show that faster Internet favours sectors that purchase a higher amount of inputs from ICT-related industries. Second, we find that better Internet induces local sorting of high-skilled workers and workers in ICT-driven occupations towards these better connected regions. Thus, our results are consistent with the Rybczynski theorem from the Heckscher-Ohlin-Vanek model and with the presence of capital-skill complementarity. Lastly, we show that reductions in communication costs induced by better provision of Internet increase local wage inequality. That is, subsidies to improve the quality of Internet may favor high-skilled workers relatively more.

Our results have implications for public expenditures on infrastructure, as they suggest that higher Internet quality explains regional development and local inequality. Advanced and middle-income economies have spent public resources on regional Internet subsidies like Canada (Government of Canada, 2019), the United States (The White House, 2022), Germany (European Commission, 2022), United Kingdom (Hutton, 2022), Colombia (Gobierno de Colombia, 2022) or Brazil (Governo Federal do Brasil, 2020). The motivation behind these policies has been to reduce the technology education gap and bring more development to the farther away regions. Nevertheless, such subsidies might lead to unintended consequences regarding local industry structure and inequality.

Three issues are beyond the scope of this paper as they would require a different econometric model and data. First, our empirical strategy does not allow us to study whether Internet improvements are zero sum. That is, whether counties with faster Internet are becoming more intensive in high-skilled services at the expense of other counties. Such question would require a dynamic panel data and an exogenous rollout of high quality Internet service. Second, our paper does not explain whether local structural transformation occurs due to the expansion of existing businesses or changes in the location decisions of services firms. These issues are addressed by Jiang (2022) for manufacturing establishments. Third, our paper does not study the impact of Internet and connectivity on inequality and convergence across regions. From our point of view, these three topics require further research.

References

Abbate, J. (2000). Inventing the Internet. MIT press.

Acosta, C. and Lyngemark, D. H. (2020). Spatial Wage Di erentials, Geographic Frictions, and the Organization of Labor within Firms. EAFIT University Working Paper .

Ahlfeldt, G., Koutroumpis, P., and Valletti, T. (2017). Speed 2.0: Evaluating Access to Universal Digital Highways. Journal of the European Economic Association, 15(3):586{625.

Ahlfeldt, G. M., Albers, T. N., and Behrens, K. (2020). Prime Locations. Working paper.

Alder, S. (2016). Chinese Roads in India: The E ect of Transport Infrastructure on Economic Development. Available at SSRN 2856050.

Allen, T. (2014). Information Frictions in Trade. Econometrica, 82(6):2041{2083.

Allen, T. and Arkolakis, C. (2022). The Welfare E ects of Transportation Infrastructure Improvements. Review of Economic Studies.

Allen, T. and Atkin, D. (2016). Volatility and the Gains from Trade. National Bureau of Economic Research.

Ar , W. B. and Hikkerova, L. (2021). Corporate Entrepreneurship, Product Innovation, and Knowledge Conversion: the Role of Digital Platforms. Small Business Economics, 56(3):1191{1204.

Asturias, J., Garca-Santana, M., and Ramos, R. (2019). Competition and the Welfare Gains from Transportation Infrastructure: Evidence from the Golden Quadrilateral of India. Journal of the European Economic Association, 17(6):1881{1940.

Atolia, M. (2007). Trade liberalization and rising wage inequality in latin america: Reconciliation with hos theory. Journal of International Economics, 71(2):467{494.

Baldomero-Quintana, L. (2022). How Infrastructure Shapes Comparative Advantage. University of Michigan Working Paper.

Beem, R. (2022). Broadband Internet and Business Activity. University of Tennesse Working Paper.

Beracha, E. and Wintoki, M. B. (2013). Forecasting Residential Real Estate Price Changes from Online Search Activity. Journal of Real Estate Research, 35(3):283{312.

Blum, B. S. and Goldfarb, A. (2006). Does the Internet Defy the Law of Gravity? Journal of International Economics, 70(2):384{405.

Bonadio, B. (2016). Ports vs. Roads: Infrastructure, Market Access and Regional Outcomes . Available at SSRN 2856050.

Bowen, H. P., Leamer, E. E., and Sveikauskas, L. A. (1986). Multicountry, multifactor tests of the factor abundance theory. National Bureau of Economic Research Cambridge, Mass., USA.

Breinlich, H. and Criscuolo, C. (2011). International Trade in Services: A Portrait of Importers and Exporters. Journal of International Economics, 84(2):188{206.

Burstein, A., Cravino, J., and Vogel, J. (2013). Importing skill-biased technology. American Economic Journal: Macroeconomics, 5(2):32{71.

Carlino, G. A., Chatterjee, S., and Hunt, R. M. (2007). Urban Density and the Rate of Invention. 61(3):389{419.

Cavallo, A. (2018). More Amazon E ects: Online Competition and Pricing Behaviors. National Bureau of Economic Research.

Cerf, V. and Khan, R. (1990). Selected ARPANET Maps 1969-1990. SIGCOMM Computer Communication Review, 20(5).

Charlot, S. and Duranton, G. (2006). Cities and Workplace Communication: Some Quantitative French Evidence. Urban Studies, 43(8):1365{1394. _eprint: https://doi.org/10.1080/00420980600776459.

Conley, T. G. (1999). GMM Estimation with Cross Sectional Dependence. Journal of Econometrics, 92(1):1{45.

Cosar, A. K., Demir, B., Ghose, D., and Young, N. (2022). Road capacity, domestic trade and regional outcomes. Journal of Economic Geography, 22(5):901{929.

Cristea, A. D. (2011). Buyer-Seller Relationships in International Trade: Evidence from U.S. States' Exports and Business-Class Travel. Journal of International Economics, 84(2):207{220.

DARPA (1981). A History of the ARPANET: The First Decade. Technical Report 4799, Defense Advanced Research Projects Agency, 1400 Wilson Bldv, Arlington, VA, 22209.

Davis, D. R. and Weinstein, D. E. (2001). An account of global factor trade. American Economic Review, 91(5):1423{1453.

Deardor , A. V. (1984). Testing trade theories and predicting trade ows. Handbook of International Economics, 1:467{517.

DeStefano, T., Kneller, R., and Timmis, J. (2022). The (Fuzzy) Digital Divide: the E ect of Universal Broadband on Firm Performance*. Journal of Economic Geography. _eprint: https://academic.oup.com/joeg/advance-article-pdf/doi/10.1093/jeg/lbac006/43438904/lbac006.pdf.

Dietzel, M. A. (2016). Sentiment-Based Predictions of Housing Market Turning Points with Google Trends. International Journal of Housing Markets and Analysis.

Dingel, J. and Neiman, B. (2020). How Many Jobs Can be Done at Home? Journal of Public Economics, 189:104235.

Donaldson, D. (2018). Railroads of the Raj: Estimating the Impact of Transportation Infrastructure. American Economic Review, 108(4-5):899{934.

Donaldson, D. and Hornbeck, R. (2016). Railroads and American Economic Growth: A \Market Access" Approach. The Quarterly Journal of Economics, 131(2):799{858.

Duranton, G. (2015). Roads and Trade in Colombia. Economics of Transportation, 4(1):16{36.

Duranton, G., Morrow, P., and Turner, M. (2014). Roads and Trade: Evidence from the U.S. Review of Economic Studies, 81(2):681–724.

Eckert, F., Ganapati, S., and Walsh, C. (2020). Skilled Scalable Services: The New Urban Bias in Economic Growth. Available at SSRN 3736487.

Encyclopedia.com (2022). Node.

Esquivel, G. and Rodríguez-López, J. A. (2003). Technology, trade, and wage inequality in mexico before and after nafta. Journal of Development Economics, 72(2):543–565.

European Commission (2022). Broadband in Germany.

Faber, B. (2014). Trade Integration, Market Size, and Industrialization: Evidence from China's National Trunk Highway System. Review of Economic Studies, 81(3):1046–1070.

Fajgelbaum, P. and Redding, S. J. (2022). Trade, Structural Transformation, and Development: Evidence from Argentina 1869–1914. Journal of Political Economy, 130(5):1249–1318.

Federal Communications Commission (2019). FCC Form 477 Local Telephone Competition and Broadband Reporting. Instructions for Filings as of December 31, 2019 and Beyond.

Fink, C., Mattoo, A., and Neagu, I. C. (2005). Assessing the Impact of Communication Costs on International Trade. Journal of International Economics, 67(2):428–445.

Ford, J. S., Rutherford, R. C., and Yavas, A. (2005a). The Effects of the Internet on Marketing Residential Real Estate. Journal of Housing Economics, 14(2):92–108.

Ford, J. S., Rutherford, R. C., and Yavas, A. (2005b). The Effects of the Internet on Marketing Residential Real Estate. Journal of Housing Economics, 14(2):92–108.

Forman, C., Goldfarb, A., and Greenstein, S. (2012). The Internet and Local Wages: A Puzzle. American Economic Review, 102(1):556–75.

Forman, C. and Van Zeebroeck, N. (2012). From Wires to Partners: How the Internet Has Fostered R&D Collaborations Within Firms. Management Science, 58:1549–1568.

Freund, C. L. and Weinhold, D. (2004). The Effect of the Internet on International Trade. Journal of International Economics, 62(1):171–189.

Gaspar, J. and Glaeser, E. L. (1998). Information Technology and the Future of Cities. Journal of Urban Economics, 43(1):136–156.

Gertler, P., Gonzalez-Navarro, M., Gracner, T., and Rothenberg, A. D. (2022). Road Maintenance and Local Economic Development: Evidence from Indonesia's Highways. Technical report, National Bureau of Economic Research.

Glaeser, E. L. and Ponzetto, G. A. (2007). Did the Death of Distance Hurt Detroit and Help New York? Technical report, National Bureau of Economic Research.

Glaeser, E. L., Saiz, A., Burtless, G., and Strange, W. C. (2004). The Rise of the Skilled City. Brookings-Wharton Papers on Urban Affairs, pages 47–105. Publisher: Brookings Institution Press.

Gobierno de Colombia (2022). Centros Digitales.

Goldberg, P. K. and Pavcnik, N. (2007). Distributional e ects of globalization in developing countries. Journal of economic Literature, 45(1):39{82.

Goldin, C. and Katz, L. F. (1998). The origins of technology-skill complementarity. The Quarterly Journal of Economics, 113(3):693{732.

Government of Canada (2019). High-Speed Access for All: Canada's Connectivity Strategy.

Governo Federal do Brasil (2020). Plano Nacional de Banda Larga.

Hanson, G. H. (1996). Economic integration, intraindustry trade, and frontier regions. European economic review, 40(3-5):941{949.

Hardy, A. P. (1980). The Role of the Telephone in Economic Development. Telecommunications Policy, 4(4):278{286.

Harrigan, J. (1997). Technology, Factor Supplies, and International Specialization: Estimating the Neoclassical Model. American Economic Review, 87(4):475{494.

Harris, S. and Gerich, E. (1996). Retiring the NSFNET Backbone Service: Chronicling the End of an Era. ConneXions, 10(4).

Hauben, M., Hauben, R., and Truscott, T. (1998). Behind the Net: The Untold Story of the ARPANET and Computer Science. Wiley-IEEE Computer Society Press.

Hutton, G. (2022). The Universal Service Obligation (USO) for Broadband.

Jiang, X. (2022). Information and Communication Technology and Firm Geographic Expansion. Duke University Working Paper.

Juhasz, R. and Steinwender, C. (2018). Spinning the web: Codi ability, Information Frictions and Trade. University of British Columbia Working Paper.

Kantor, S. and Whalley, A. (2019). Research Proximity and Productivity: Long-Term Evidence from Agriculture. Journal of Political Economy, 127(2):819{854. _eprint: https://doi.org/10.1086/701035.

Kleibergen, F. and Paap, R. (2006). Generalized Reduced Rank Tests Using the Singular Value Decomposition. Journal of Econometrics, 133(1):97{126.

Leamer, E. E. (1980). The leontief paradox, reconsidered. Journal of Political Economy, 88(3):495{503.

Leiner, B., Cerf, V., Clark, D., Kahn, R., Kleinrock, L., Lynch, D., Postel, J., Roberts, L., and Wol , S. (1997). A Brief History of the Internet.

Leontief, W. (1953). Domestic production and foreign trade; the American capital position re-examined. Proceedings of the American Philosophical Society, 97(4):332{349.

Licklider, J. C. and Taylor, R. W. (1968). The Computer as a Communication Device. Science and Technology, 76(2):1{3.

Lin, J. (2011). Technological Adaptation, Cities, and New Work. The Review of Economics and Statistics, 93(2):554{574.

Lobo, B., Novobilski, A., and Ghosh, S. (2008). The Economic Impact Of Broadband: Estimates From A Regional Input-Output Model. Journal of Applied Business Research, 24(2).

Malecki, E. J. (2002). The Economic Geography of the Internet's Infrastructure. Economic Geography, 78(4):399{424. Publisher: Routledge _eprint: https://www.tandfonline.com/doi/pdf/10.1111/j.1944-8287.2002.tb00193.x.

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., and Ruggles, S. (2022). IPUMS National Historical Geographic Information System: Version 17.0 [dataset].

Marinoni, A. and Roche, M. (2022). You've Got Mail! Communication Infrastructure, Firm Entry and Performance - Evidence from the US Postal Service Expansion 1880-1900. Mimeo, Georgia Institute of Technology and Harvard University.

McKenzie, A. and Walden, D. (1991). The ARPANET, the Defense Data Network, and the Internet. Froehlich, F. & Kent, A., Encyclopedia of Telecommunications. Marcel Dekker, New York, pages 365{367.

Michaels, G. (2008). The E ect of Trade on the Demand for Skill: Evidence from the Interstate Highway System. The Review of Economics and Statistics, 90(4):683{701.

Network Information Center, SRI International (1978). ARPANET Directory. Technical report, Defense Communications Agency, Menlo Park, California, 84025.

Oestmann, M. and Bennöhr, L. (2015). Determinants of House Price Dynamics. What Can We Learn from Search Engine Data? Review of Economics, 66(1):99{127.

Parro, F. (2013). Capital-skill complementarity and the skill premium in a quantitative model of trade. American Economic Journal: Macroeconomics, 5(2):72{117.

Peñarroyo, C., Lindquist, S., and Miller, R. (2022). Mapping Detroit's Digital Divide. University of Michigan Urban Laboratory.

Rosenblat, T. S. and Mobius, M. M. (2004). Getting Closer or Drifting Apart? The Quarterly Journal of Economics, 119(3):971{1009. Publisher: Oxford University Press.

Rybczynski, T. M. (1955). Factor endowment and relative commodity prices. Economica, 22(88):336{341.

Sotelo, S. (2020). Domestic Trade Frictions and Agriculture. Journal of Political Economy, 128(7):2690{2738.

Steinwender, C. (2018). Real E ects of Information Frictions: "When the States and the Kingdom became United". American Economic Review, 108(3):657{696.

The White House (2022). Biden-Harris Administration Announces Over $25 Billion in American Rescue Plan Funding to Help Ensure Every American Has Access to High Speed, A ordable Internet.

Thornton, K. and Mars, R. (2022). The Future of the Final Mile. Podcast 99 Percent Invisible, Episode 481

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from india. American Economic Journal: Applied Economics, 2(4):1{41.

Tre er, D. (1995). The case of the missing trade and other mysteries. The American Economic Review, pages 1029{1046.

U.S. Congress (1992). Management of NSFNET. Hearing Before the Subcommittee on Science, Space, and Technology, U.S. House of Representatives, 102nd Congress, Second Session. page 190.

Verhoogen, E. A. (2008). Trade, quality upgrading, and wage inequality in the mexican manufacturing sector. The Quarterly Journal of Economics, 123(2):489{530.

Wall, C. (2021). Invisible and Vital: Undersea Cables and Transatlantic Security. Center for Strategic and International Studies.

Zook, M. A. (2002). Grounded Capital: Venture Financing and the Geography of the Internet Industry, 1994{2000. Journal of Economic Geography 2(2):151{177.

Quality of Communications Infrastructure Provision and Local Structural Transformation

Camilo Acosta, Universidad EAFIT
Luis Baldomero-Quintana, William & Mary

Appendix
(for online publication)

A1 Extra Figures and Tables

Figure A-1: Original description of ARPANET nodes in December 1969

Source : Cerf and Khan (1990)

Figure A-2: ARPANET - 1988, Digitized Map

Note: this gure shows a digitized map of the ARPA network as of April 1988 extracted from Cerf and Khan (1990).

Figure A-3: Employment Shares by Aggregate Sectors

(a) Agriculture and Mining

(b) Manufacturing

(c) High Skilled Services

(d) Other Services

Note: these maps show the shares of employment in each county by quintiles using data from the 2018 County Business Patterns.

Figure A-4: Payroll Shares by Aggregate Sectors

(a) Agriculture and Mining

(b) Manufacturing

(c) High Skilled Services

(d) Other Services

Note: these maps show the shares of aggregate payroll in each county by quintiles using data from the 2018 County Business Patterns.