The creation and diffusion of knowledge: Evidence from the Jet Age

Stefan Pauly* Fernando Stipanicic†

November 8, 2021

Click here for the latest version

This paper studies the impact of travel time on the diffusion and creation of knowledge. We provide causal evidence by exploiting the beginning of the Jet Age as a natural experiment. We digitize airlines' historical flight schedules and construct a novel data set of the flight network in the United States. Between 1951 and 1966, travel time between locations more than 2,000km apart decreased on average 41%. We use patent data as a measure of knowledge creation and diffusion. For research establishments located more than 2,000km apart, the reduction in travel time increased citations by 6.9%, accounting for 32.7% of the observed increase in citations in this distance interval. The decrease in travel time also led to an increase in knowledge access of each location which spurred the creation of new patents. The effect was stronger in initially less innovative locations, leading to a yearly growth rate of patenting 1.2 percentage points higher relative to more innovative locations. The predicted difference accounts for 21% of the convergence rate observed. We uncover one mechanism through which convergence occurred: expansion of multi-establishment firms. We find that the reduction in travel time to a firm's headquarters led to an increase in the amount of subsidiaries in other locations, with a larger relative increase of subsidiaries in initially less innovative locations.

JEL Classification: O31, O33, R41, N72

*Sciences Po, Ph.D. Candidate. Email: *stefan.pauly@sciencespo.fr*

†Toulouse School of Economics, Ph.D. Candidate. Email: *fernando.stipanicic@tse-fr.eu*

1. Introduction

During the 20th century airplanes marked a revolution in time and space: they made it faster than ever to transport people between distant places. Since their inception airplanes continuously evolved but during 1950s there was a breakthrough in the development of aviation: the introduction of jet engines into civil aircrafts. This event marked the beginning of the *Jet Age* in civil aviation and allowed airplanes to go faster, farther and carry more people.

This paper provides new causal evidence of the impact of improvements in air travel during the beginning of the *Jet Age* on the diffusion and creation of knowledge. Accessibility to knowledge is essential to create new innovations (Furman and Stern (2011), Acemoglu et al. (2016)) and innovation has been characterized as one of the main causes of economic growth (Lucas (1993), Aghion and Howitt (1997) and Jones (2002)). Physical proximity is a key driver of the diffusion of knowledge as it enables face to face interactions (Storper and Venables (2004), Glaeser (2011)). In this paper we argue that the developments in air travel improved distant face to face interactions which increased knowledge diffusion and creation.

The beginning of the *Jet Age* happened in an environment with solid regulation. This implied that differences in travel time were mostly due to differences between the new and old flight technology. In consequence, the beginning of the *Jet Age* appears as a natural experiment at the national level which is suitable to study the effect of improved connectivity on other outcomes.

We find three main results. First, the reduction in air travel time increased knowledge diffusion, especially between far away locations. Second, the reduction in travel time led to an improvement in a location's access to knowledge, which further triggered an increase in the production of new patents. The effect on the increase in the production of new patents was disproportionally stronger in initially less innovative locations. We find that most of the increase in new patents comes from two entry margins: entry of new establishments of new firms and new establishments of firms that expand from other locations. Third, we find that travel time affected firms' locations decisions. The decrease in travel time led to an increase in firms' geographical scope, with a larger impact on the relative entry rate in initially less innovative locations. The improvement in access to knowledge acted as a convergence force between locations. Firms' expansion is one of the mechanisms that explains the convergence.

We start by constructing a new dataset of the flight network in the United States during 1950s and 1960s. We digitize historical flight schedules of the major interstate airlines operating in the period and obtain the fastest route between every two airports in the network. We document that between 1951 and 1966 travel time decreased on average by 29%, and the decrease is on average of 41% for airports located more than 2,000km apart.[1]

We decompose the change in travel time and find that 90% of the change is due to the improvement in aircrafts' speed, while 10% is due to a change in the flight routes. This is consistent with the fact that during this period the Civil Aeronautics Board (CAB) was imposing strong regulation in the interstate airline market. With the objective to promote a *stable* airline industry, the CAB determined ticket prices and restricted entry of airlines into new or existing routes.

[1]New York and Boston are about 300km apart, while New York and San Francisco are located about 4,130 km apart. Between 1951 and 1966 we observe a reduction of travel time of 23% (13 minutes reduction) between New York and Boston, while the reduction is of 50% (5 hours 30 minutes reduction) between New York and San Francisco.

Airlines could choose the type of aircraft, frequency of flights and layover time, endogenous variables that represent a threat to identification. We tackle the endogeneity of travel time to airlines' decisions by constructing an instrument that depends only on the nationwide roll out of jet airplanes. Results show that there is no evidence of endogeneity in long distance routes.

Additionally, air travel saw a big increase in the transport of people, while the transport of goods remained negligible. Annual reports of the Interstate Commerce Commission (I.C.C. (1965), I.C.C. (1967)) show that in 1951 airplanes transported one-third the amount of passenger-miles of those transported by trains, while in 1966 airplanes transported three times as much as trains. In the same time period, airplanes did not carry more than 0.1% of the total intercity ton-miles. In other words, the effects of improved connectivity can be attributed to interactions between people and not to trade.

To study knowledge flows, which by their nature are difficult to observe,[2] we follow Jaffe et al. (1993) and use patent citations as our observable measure of knowledge flow. We merge multiple sources of patent data to construct one dataset of all patents granted by the United States Patent and Trademark Office (USPTO) with filing year between 1945 and 1975, which includes for each patent: filing year, technology classification, location of the inventors when they applied for the patent, owner of the patent and citations to other patents which were granted by the USPTO. We consider only patents with inventors located in the US. When a firm has an inventor applying for a patent in certain Metropolitan Statistical Area (MSA), we say that the firm owning that patent has a research establishment in that MSA.

[2] In the words of Krugman (1992) *"they leave no paper trail by which they may be measured and tracked."*

We merge the dataset on travel time with the one on patents and estimate reduced form models to provide causal evidence of the impact of changes in travel time on the diffusion of knowledge, the creation of knowledge and firms' locations decision. We proceed in three steps.

First, we use tools of the international trade literature and estimate a gravity equation to obtain the elasticity of citations to travel time. We identify the elasticity exploiting only within establishment-pair across-time variation in citations and travel time. At the establishment level the elasticity is -0.083 statistically significant at the 1% level. This value implies that citations increased on average 2.4% due to the decrease in travel time between 1951 and 1966. We find that the absolute value of the elasticity is increasing with the distance between the citing and cited establishments. At a distance of more than 2,000km, the change in travel time implies an increase in citations of 6.9%. This accounts for 32.7% of the observed increase in citations at more than 2,000km. In order to rule out the possibility that the opening of new routes or the timing of adoption of jets at the route level is correlated with future knowledge flows, we perform instrumental variables estimation. We instrument the observed travel time with a fictitious travel time that changes only due to the nationwide roll out of jets, hence independent of decisions at the route level. The results do not change significantly, reflecting the reduced scope for endogeneity of travel time.

Second, using the estimated elasticity of diffusion of knowledge, we compute a measure of access to knowledge that is specific to each location-technology. The measure captures changes in access to knowledge that are only consequence of the change in travel time. We use access to knowledge as an input to produce new patents and estimate the elasticity of patents to knowledge access. We identify the elasticity at the establishment level relying only across time variation in patents and knowledge access. We find an elasticity of 10.14, significant at the one percent level. This elasticity implies that new patents increased at an yearly growth rate of 3.5% as consequence of the increase in knowledge access, which accounts for 79.5% of the observed yearly growth rate.

We find that the value of the elasticity of patents to knowledge access is bigger in magnitude for establishments located in initially less innovative locations. Within each technology class, we rank locations according to the amount of patents in the initial time period and split them into four quartiles. We find that the increase in knowledge access predicts an yearly growth rate of patenting of 4.5% in locations in the lowest quartile of initial innovativeness, while it predicts a 3.4% yearly growth rate in the highest quartile. The difference in growth rates indicates that the increase in knowledge access acted as a convergence force between locations, and it has the power to explain 21% of the convergence observed in the data.

Third, we uncover the sources of the increase in patenting. We open the effect of knowledge access on patenting by whether the research establishment and/or the firm existed in the initial time period. We find that most of the effect of knowledge access on new patents happens through two entry margins: entry of establishments of new firms and entry of subsidiaries of firms expanding from other locations. The margin of entry of firms creating a subsidiary is particularly pronounced in destination locations that are in the lowest quartile of initial innovativeness.

We further study if firm's subsidiary's location decision depends on travel time. We estimate a probability model of a location having an establishment of a firm depending on travel time to the firm's headquarters' location. We identify the change in the probability only from changes in travel time and the opening and closure of establishments. We find that locations that receive a higher relative reduction in travel time to a firm's headquarters increase relatively more the probability of receiving a subsidiary of that firm. The impact of the reduction in travel time on the change in the probability relative to the baseline probability is higher in initially less innovative locations. The change in travel time implies that the yearly growth rate of the probability of receiving a subsidiary of a firm headquartered in another location is 6 percentage points higher in receiver locations in the lowest quartile of innovativeness, relative to the receiver locations in the highest quartile. In the data we observe that yearly growth rate of the probability is 4.8 percentage points higher in locations in the lowest quartile. The change in travel time predicts a stronger differential growth rate, as observed in the data.

The results show that jet airplanes changed the geography of innovation. First, due to a stronger reduction in travel time for longer distances, jet airplanes increased the diffusion of knowledge especially towards locations far away from innovation centers. New York and Chicago were the main innovation centers of the United States in 1950s, hence locations in the South and the West were the most benefited. Second, the effect of diffusion of knowledge led to an increase in innovation, and the more so in locations that were initially less innovative. Therefore, jet airplanes led to convergence across locations in terms of innovation. We highlight that one of the mechanisms through which the catching up operated was the geographical expansion of multi-establishment firms.

Our convergence result contrasts with previous studies on improvements in other means of transport. Pascali (2017) finds that the introduction of steam engine vessels in the second half of the 19th led an increase in international trade which contributed to economic divergence between countries. Faber (2014) finds that the expansion of the highway system in China led to a reduction of GDP growth of peripheral counties, with evidence suggesting a trade channel. While both papers emphasize a trade channel, in our set up the trade channel would not be of first order. During 1950s and 1960s, the introduction of jet airplanes represented a big shock to the mobility of people while not affecting significantly the transport of goods. Hence, we uncover a new effect of improved connectivity.

The seminal paper Jaffe et al. (1993) finds that patent citations decay rapidly with distance. Our work is related to other literature which found that business travel affects innovation (Hovhannisyan and Keller (2015)), trade (Söderlund (2020)) and industrial activity (Coscia et al. (2020)). Also, air travel shapes collaboration between researchers (Catalini et al. (2018)). Our result about firms deciding their establishments' locations based on travel time to headquarters is comparable to the one found by Giroud (2013), who finds that a reduction in air travel time to headquarters increases plant level investment and total factor productivity. Similarly, Campante and Yanagizawa-Drott (2017) finds that firms' cross country investment decision depends on connectivity to headquarters.

The paper is structured as follows. First, we present a simple theoretical framework which lays the foundations of how to think about the creation and diffusion of knowledge. The framework shows the two key parameters to estimate. Second, we describe the historical context in which jet airplanes were introduced. Third, we present the two datasets that we use: travel times and patents. Fourth, we perform the analysis to estimate the impact of travel time on the diffusion of knowledge, the creation of knowledge, and firm's location decision. Fifth, we conclude.

2. Conceptual framework

This section lays out a simple theoretical framework to think about the creation of knowledge. The framework clearly shows the two key parameters to estimate empirically: the elasticity of knowledge diffusion to travel time and the elasticity of knowledge creation to knowledge access.

Following Carlino and Kerr (2015) we consider a production function of knowledge which includes external returns in the form of knowledge spillovers. Knowledge output of a firm depends not only on firm's specific characteristics as its idiosyncratic productivity and input decisions, but also on an externality due to knowledge spillovers. We consider a production function of knowledge of the following form:

$$\text{New Knowledge}_{Fi} = f(z_{Fi}, inputs_{Fi}) \times \text{Knowledge Access}_i^{\rho} \tag{1}$$

where New Knowledge$_{Fi}$ is the knowledge created by firm F located in i. The production output of Fi depends on an *internal* component and on an *external* component. The *internal* component is the firm's idiosyncratic productivity z_{Fi} and choice of inputs $inputs_{Fi}$. The *external* component represents the externality to which all firms F in location i are exposed to: Knowledge Access$_i$. This externality, *Knowledge Access*, represents the total amount of knowledge spillovers that the firm is exposed to. The degree to which the externality affects the production of knowledge is governed by the parameter ρ. If ρ is zero then knowledge spillovers have no effect on the creation of new knowledge. On the other hand, a positive ρ implies that, keeping productivity and inputs constant, an increase in the level of knowledge spillovers leads to an increase in firm F's creation of new knowledge.

A long standing literature studies the importance of knowledge spillovers for the creation of new knowledge.[3] The concept of knowledge spillovers goes back at least to Marshall (1920) who explains it as one of the agglomeration forces. Krugman (1991) refers to knowledge spillovers as one of the justifications for external increasing returns, and that the degree of spillovers are dependent on physical distance. The geographic decay of spillovers is grounded in the fact that not all knowledge is easy to codify, usually referred to as *tacit knowledge*, and geographic proximity increases the degree of knowledge spillovers by facilitating face-to-face interactions (Storper and Venables (2004), Glaeser (2011)). Hence, we consider the total amount of knowledge spillovers to which the firm F in location i is exposed to has the following functional form:

$$\text{Knowledge Access}_i = \sum_j \text{Knowledge stock}_j \times \text{distance}_{ij}^{\beta} \tag{2}$$

where Knowledge stock_j is the total amount of knowledge in location j (which is non-negative) that could potentially spill over to location i and distance_{ij} is a measure of distance from j to i. The amount of knowledge that spills over from j to i depends on distance and the degree with which distance impedes spillovers, governed by the parameter β. If β is zero, then distance does not affect knowledge spillovers from j to i and all locations perfectly share the same level of *Knowledge Access*. On the contrary, a negative β implies a decay in knowledge spillovers when distance increases. In other words, a negative β implies that if we reduce the distance from j to i while keeping every other distance constant, the amount of spillovers from j to i will weakly increase.

[3] The chapters of Audretsch and Feldman (2004) and Carlino and Kerr (2015) in the Handbook of Regional and Urban Economics provide an excellent review on the literature on knowledge spillovers, their geographic decay and how they affect the creation of knowledge.

This theoretical framework bears resemblance to the concept of *Market Access* presented in Donaldson and Hornbeck (2016) and Redding and Venables (2004). If we interpret *Knowledge Access* as one of the inputs in the production function of knowledge, then Knowledge Access$_i$ could be interpreted as a measure of *Input Market Access*. This measure captures how cheaply firms in location i can access pre-existing knowledge, where the cost of accessing knowledge depends on distance between i and j. Also, *Knowledge Access* is similar to a measure of network centrality. The centrality of each location i (node) is the weighted sum of distance (edges) to every location, where the weight of each location is given by its knowledge stock.

The theoretical framework highlights the two parameters to estimate: ρ and β. Empirically, we use travel time as a measure of distance to first estimate β and then conditional on β we estimate ρ. Changes in travel time due to improvements in commercial aviation allow us to estimate both parameters. First, we use citations between patents as a proxy for the diffusion of knowledge. We estimate β by relating changes in travel time between research establishments to changes in citations between research establishments. Second, we use the stock of patents filed by inventors in each location as proxy for each location's stock of knowledge. We construct a measure of knowledge access using the patent stock, travel times and the value of β. New patents in each location proxy for new knowledge. Changes in travel time lead to changes in knowledge access which allow us to estimate ρ.

3. Historical context

3.1. Air transport: jet arrival

The jet aircraft was first invented in 1939 for military use, with the German Heinkel He 178 being the first jet aircraft to fly. The first commercial flight by a jet aircraft was in 1952 by the British Overseas Airways Corporation (BOAC) from London, UK to Johannesburg, South Africa with a Havilland Comet 1. Nonetheless, given the amount of accidents of the Havilland Comet 1 due to metal fatigue, jet commercial aviation did not truly take off until the Boeing 707 entered commercial service in late 1958. The 24th of January of 1959 represented a major milestone in the jet era: American Airlines Flight 2 flew with a Boeing 707 jet aircraft from Los Angeles to New York, the first non-stop transcontinental commercial jet flight.[4]

In 1951 New York City and Los Angeles were connected with a one-stop flight in 10 hours and 20 minutes. The flight had a forced stop in Chicago and was operated with the propeller aircraft Douglas DC-6, which had a cruise of 500 kmh. By 1956, New York City and Los Angeles were connected with a non-stop flight in 8 hours and 30 minutes. This was accomplished due to the introduction of the propeller aircraft Douglas DC-7 which had a cruise speed of 550kmh, and a change in regulation which increased maximum flight time of a crew from 8 to 10 hours within a 24-hour window.[5] In 1961, the route was covered with the jet

[4]The reader passionate of aviation history would enjoy reading the following New York Times article which tells the experience of the first transcontinental jet flight: `https://www.nytimes.com/2009/01/26/nyregion/26american.html`

[5]AA and TWA had transcontinental non-stop propeller flights scheduled since at least 1954. These flights were scheduled to take 7 hours 55 minutes, just under the maximum flight time allowed by regulation in domestic flights: regulation impeded pilots from being on duty more than 8 hours within a 24 hours window. Nonetheless, the propeller aircrafts used in these flights, the Douglas DC-7 and the Lockheed Super Constellation, overheated their engines due to excessive demand to cover the route in less than 8 hours. AA fought intensely until

aircraft Boeing 707 in a non-stop flight in 5 hours 15 minutes, reaching 5 hours 10 minutes in 1966. The Boeing 707 had a cruise speed of 1000kmh, cutting travel time from New York City to Los Angeles in half between 1951 and 1966.

3.2. Air transport: moving people, not goods

During the 1950s and 1960s, air transportation served to transport people but not goods. Figures 1 and 2 are images (edited for better readability) from annual reports of the Interstate Commerce Commission of 1967 and 1965 respectively. Figure 1 displays the amount of passenger-miles[6] for Air, Motor and Rail transportation from 1949 to 1966. We observe that, while transport of people by rail decreased and by motor remained relatively constant, transport of people by air multiplied by 6 in a 16-year period, which translates to around 12% compound annual growth. In 1966, air transport accounted for more passenger-miles than both rail and motor transportation together, reflecting the growing importance of this mean of transport.

Figure 2 shows shipments in ton-miles for the period 1939 to 1964 by mean of transport: Airways, Pipelines, Inland Waterways, Motor, Railroads. Interestingly, we observe that air transport of goods, even if it increased, it accounted for less than 0.1% of transport of goods in 1964.[7]

the CAB approved a waiver that allowed non-stop transcontinental flights to take up to 10 hours to accomplish the non-stop transcontinental flight. See page 16 of the edition of 21st of June 1954 of the Aviation Week magazine `https://archive.org/details/Aviation_Week_1954-06-21/page/n7/mode/2up`

[6] Passenger-miles is a standard unit of measurement in transport, where one passenger-mile accounts for one person traveling one mile. The reasoning is the same for ton-miles, with one ton of goods traveling one mile.

[7] We have not found data about shipments by mean of transport measured in monetary values.

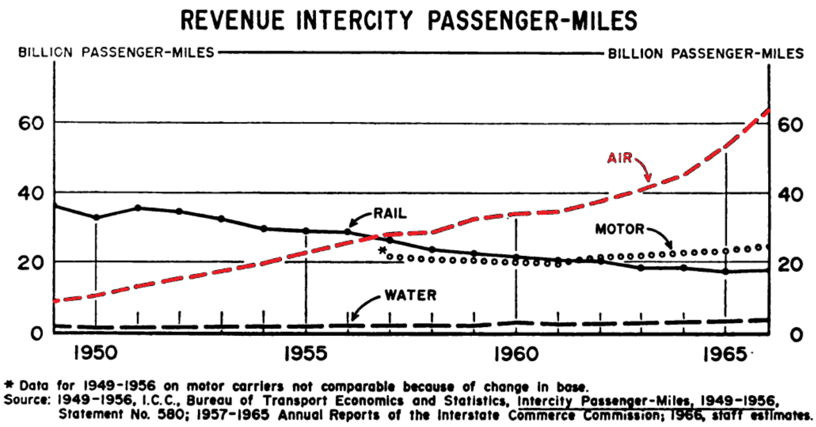

Figure 1: Passenger Miles
Source: Interstate Commerce Commission, Annual Report 1967
Edited by the authors

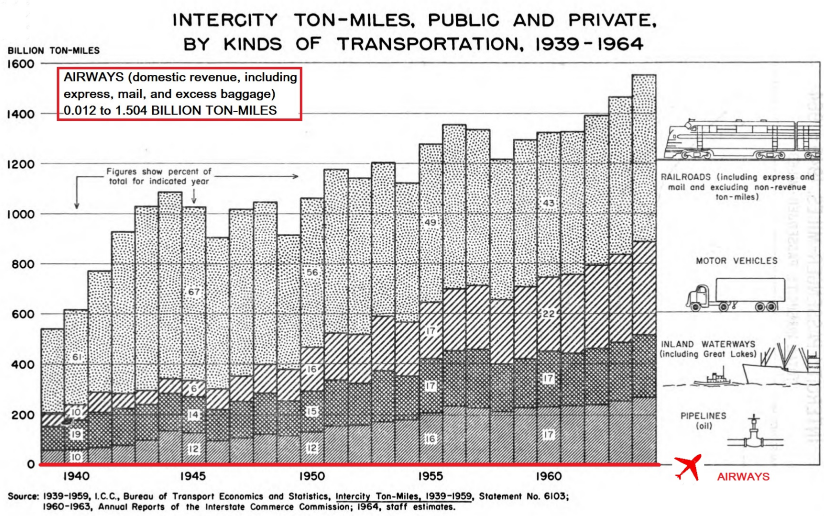

Figure 2: Ton Miles
Source: Interstate Commerce Commission, Annual Report 1965
Edited by the authors

3.3. Regulation

As explained in Borenstein and Rose (2014), in the 1930s the airline industry was seen as suffering from coordination issues, destructive competition and entry. Additionally, the industry was developing in a context of financial instability and increasing military concerns post Great Depression. A strong domestic airline industry was perceived as an interest of national defense. As consequence, the Civil Aeronautics Board (CAB) was created in 1938 with the objective to promote, encourage and develop civil aeronautics.[8] It was empowered to control entry, fares, subsidies and mergers.[9] In other words, the CAB regulated the market by deciding which airlines could fly, in which routes they could operate, the price that they charged in each route, the structure of subsidies and merger decisions. The CAB regulated the airline industry in a barely unchanged manner until it ceased to exist in 1985.

When the CAB was created, it conceived special rights to the existing airlines over the connections they were operating. The CAB did not permit entry of new airlines on interstate routes and gradually allowed current airlines to expand their routes. The CAB controlled both the system and each airline's network. The network was designed to maintain industry stability and minimize subsidies, leading to a system where each route was mainly operated by one or two airlines.[10] Importantly, Borenstein and Rose (2014) in pages 68-69 explain that *"the regulatory route award process largely prevented airlines from reoptimizing their*

[8]The CAB was a federal agency hence, in principle, would not have control over intrastate routes. Nonetheless, according to Borenstein and Rose (2014) the CAB managed to have some intrastate markets under its control using legal arguments.

[9]Safety regulation was under the control of the Federal Aviation Administration.

[10]Borenstein and Rose (2014) in page 68, based on Caves (1962), expose *"In 1958, for example, twenty-three of the hundred largest city-pair markets were effectively monopolies; another fifty-seven were effectively duopolies; and in only two did the three largest carriers have less than a 90 percent share."*

networks to reduce operation costs or improve service as technology and travel patterns changed." As a consequence, any technological improvement such as increases in aircraft speed, capacity or range would not affect each airline's flight network in the short term.

By regulating fares, the CAB explicitly encouraged airlines to adopt new aircraft. Airlines, when operating an older aircraft, would apply for a fare reduction arguing that it is needed in order to preserve demand for low quality service. The CAB would refuse this application, hence airlines would have to adopt new aircraft or risk losing consumers who would choose another airline which flies newer aircrafts.

4. Air travel data

We construct a new data set of the flight network in the United States during the 1950s and 1960s. We collected and digitized information of all the flights operated by the main airlines and obtained the fastest route and travel time between every two airports in the network.

To construct the flight network we use historical flight schedules of the main airlines operating in 1950s and 1960s. Figure 3 is a fragment from an example page of the 1961 flight schedule of American Airlines. In the flight schedule we observe in the center column the name of departure and arrival cities (which we match to airports using airlines' historical ticket office geographical location), while the small columns on the sides depict flights. In the top of the small columns we observe the type of service provided (first class, coach or both), aircraft operated, days operated (daily if information is missing) and flight number. The content of the small columns displays the departure and arrival time (local time, bold numbers represent PM) at each city, including all intermediate stops.

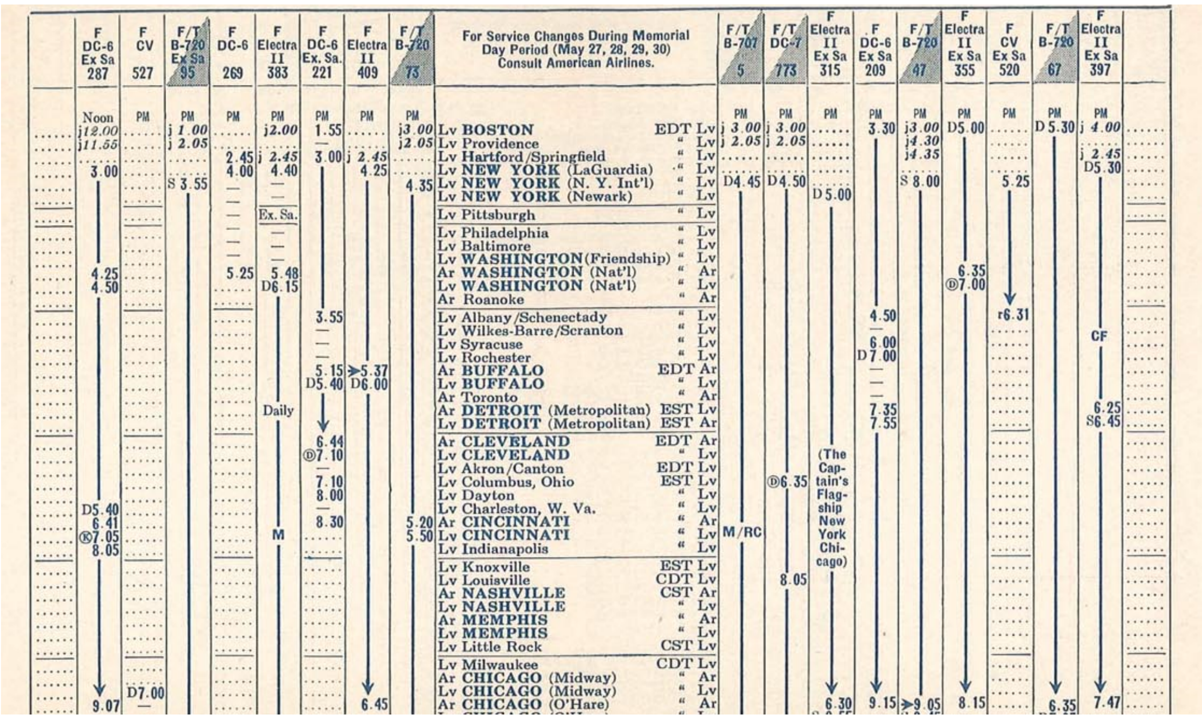

Figure 3: Flight schedule American Airlines 1961
The center column displays the name of departure and arrival cities. The small columns on the sides display flights with departure and arrival time (local time, bold numbers represent PM). The top of the small columns shows the type of service provided (first class, coach or both), aircraft operated, days operated (daily if information is missing) and flight number.

We digitize flight schedules for the years 1951, 1956, 1961 and 1966 of six domestic airlines: American Airlines (AA), Eastern Airlines (EA), United Airlines (UA), Trans World Airlines (TWA), Braniff International Airways (BN), Northwest Airlines (NW),[11] and one international airline: Pan American Airways (PA). This group of airlines includes the *Big 4*: AA, EA, UA and TWA, which

[11] These are six of the fifteen trunk (interstate) airlines operating in 1951. Many of the remaining trunk airlines would merge with another trunk airline over the years, and there would be zero entry of new airlines. We are currently digitizing the remaining trunk airlines and we plan to add them to the travel time dataset in the future. We have already digitized: Allegheny

accounted for between 69% and 74% of interstate air revenue passenger miles in the US in the years collected. BN and NW were digitized in order to have a wide geographical coverage, while PA provides international flights. Based on C.A.B. (1966), in the years collected, the six domestic airlines together account for between 77% and 81% of interstate air revenue passenger miles.

In total we have digitized 6,143 US flights (unique combinations of flight number-year, 7,007 worldwide). However, flights often have multiple stops. If we count each non-stop part (*leg*) of these flights separately, our sample contains 17,737 legs in the US and 21,210 worldwide. Our data connects 275 US airports (434 worldwide) creating 2,563 unique origin-destination (directional) airport links (3,466 worldwide). Figure 4 displays the flight network in continental United States pooling all years together. In Appendix A.2 we show the US flight network by year, around 80% of the non-stop flights remain year-on-year.

Using departure and arrival time of each flight at each airport, we obtain the fastest route and corresponding travel time between every two airports in our data. To obtain the fastest route and travel time we modify the Dijkstra algorithm to account for layover time in case the fastest route includes connecting flights.[12]

Once the fastest route between every two airports is computed, we match

Airlines, Capital Airlines, Colonial Airlines, Continental Airlines and Delta Air Lines. We have also digitized the year 1970 for the six airlines used in this paper and Pan American. Due to a time constraint we have not included them in the current analysis. We are planning to digitize BOAC to obtain more international flights, and to cover for all airlines possible a 70 time period: 1930 to 2000.

[12]We are currently working on setting a minimum waiting time for switching airplanes, such that the change is not permitted unless waiting time is more than the minimum. For the time being we have set the minimum waiting time to zero, meaning that in our calculation one passenger would be able to switch from one airplane to another if departure of the following flight is one minute later than arrival of the previous flight. This a rather implausible assumption and we are estimating the minimum waiting time in each airport depending on the airport's congestion.

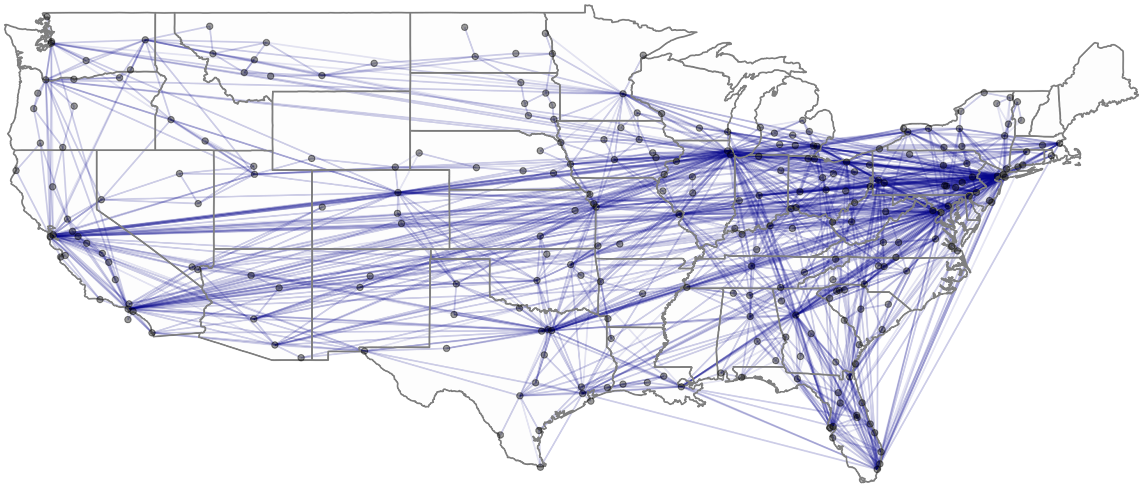

Figure 4: United States flight network 1951-1966

every airport to 1950 Metropolitan Statistical Areas (MSA) using the shape file from Manson et al. (2020). We consider only MSAs in contiguous United States. We use MSAs as the geographical unit of analysis because they are constructed taking into account commuting flows. We assume that people in an MSA would use, for each desired route, the most appropriate airport lying inside or nearby the MSA. We match each airport to all MSAs for which it lies inside the MSA boundary or is at most 15km away from the MSA boundary.[13] 176 out of 275 US airports are matched to at least one MSA. Meanwhile, 142 out of 168 MSAs are matched to one or more airports in at least one year, and 108 MSAs are matched to one or more airports in the four years. We use the sample of 108 MSAs that had an airport in the four years as our baseline travel time data.[14]

[13]The 15km distance was chosen after inspecting airports near the border that should arguably be matched, as for example, Atlanta ATL airport.

[14]In Appendix A.2 we include a table with the 168 MSAs, those connected at least once and those connected in the four years. Among the MSAs not connected is San Jose, California, which in our patent sample accounted for around 2% of patents. San Jose had an airport (SJC) during our time period but it was not served by any of our airlines, so it is not included in our analysis.

4.1. Descriptive statistics: Air travel

To understand the changes in travel time we will first study travel time of non-stop flights and then of all routes including connecting flights. Figure 5 displays the non-stop fastest flight within each MSA pair that was operating in each year. In 1951 the longest non-stop flight across MSAs was between Chicago and San Francisco using the Douglas DC-6, covering a distance of 2,960 km in 7 hours 40 minutes. This travel time was just under 8 hours, the maximum flight time allowed for a crew in a 24-hour period.[15] In 1956, new regulation allowed up to 10 hour flights for transcontinental flights, the longest non-stop flight between MSAs was New York to San Francisco with the Douglas DC-7, covering a distance of 4,151 km in 9 hours. Between 1951 and 1956, while we observe an increase in average flight speed that went up to 17%, the main change observed is that longer non-stop routes were possible.

In 1961, the first year in which we have jet aircrafts in the travel time data, there is a reduction in travel time between MSA-pairs, especially for those far apart from each other. In 1966, there is a further decrease in travel time due to a widespread adoption of jet aircrafts in shorter distances. In Appendix Figure 22 we show the jet adoption rate by distance for MSAs connected with a non-stop flight. All MSA-pairs more than 3,000km apart connected with a non-stop flight operate at least one jet flight in 1961, and this expands to all those more than 2,000km apart in 1966. The speed gain of jets relative to propeller aircrafts is increasing with the amount of time that the jet can fly at its cruise speed, arguing in favor of an adoption that is increasing with the distance between origin and destination.[16]

[15]Honolulu was not concerned by the regulation. Honolulu was connected with non-stop flights to San Francisco (9 hours 40 minutes), Los Angeles (11 hours) and Portland (12 hours 55 minutes).

[16]We are currently exploring the differential timing of jet adoption across airlines. Differences

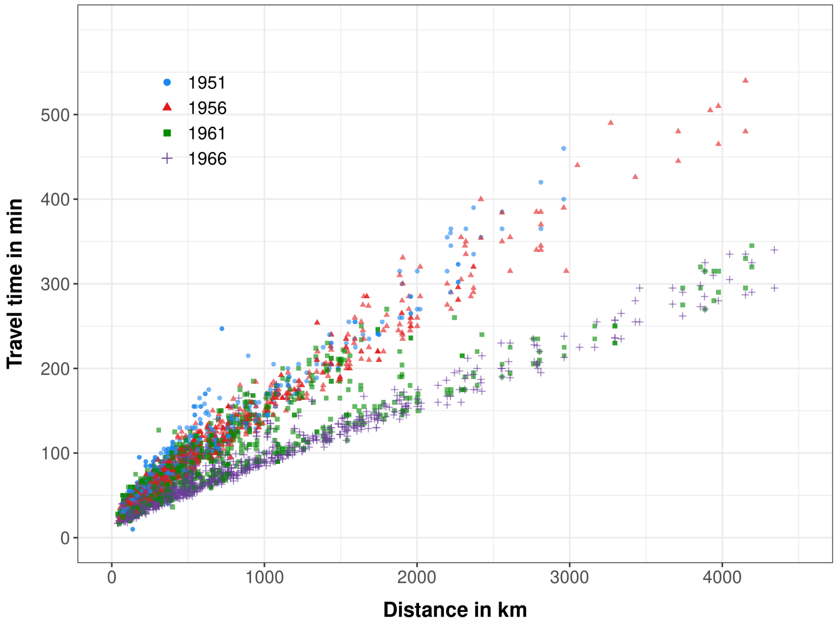

Figure 5: Non-stop fastest flights United States MSAs

The change in travel time in non-stop flights is also reflected in the travel time for connecting flights. Figure 6 shows, relative to 1951, the average and standard deviation change in travel time for all MSA-pairs, including non-stop and connecting flights.[17] Between 1951 and 1956, there is an average reduction in travel time of 9.2% which is roughly constant for all distances over 500km. Between 1951 and 1961, there is a reduction in travel time that is increasing with distance. The average decrease in travel time is of 16.8%, while the reduction is of 29.4% for a distance of more than 2,000km and 39.2% for a distance of

in (pre-existing) route distance and past contractual relationships with aircraft suppliers potentially led to different adoption rates at each time period. For example, Eastern Airlines' routes were particularly shorter than for other airlines. Also, those committed to buy Douglas airplanes (the leader US commercial aircraft supplier pre-jet era) would have adopted jets later, as Douglas launched jet airplanes later than Boeing.

[17]The plot includes only MSA-pairs with travel time in all time periods. The standard deviation for MSA-pairs less than 250km apart is big relative to the ones at other distances. Hence we decided not to include it because it distorts the visualization of the rest of the plot.

4,250-4,500km. Between 1951 and 1966, there is an even stronger decrease in travel time at all distances. The average reduction in travel time is 28.7% across all distances, 40.8% for a distance of more than 2,000km and 48.4% for a distance of 4,250-4,500km. The increased adoption of jets for short distance flights implied that both non-stop flights at short distance and connecting flights at farther distance had a decrease in travel time.

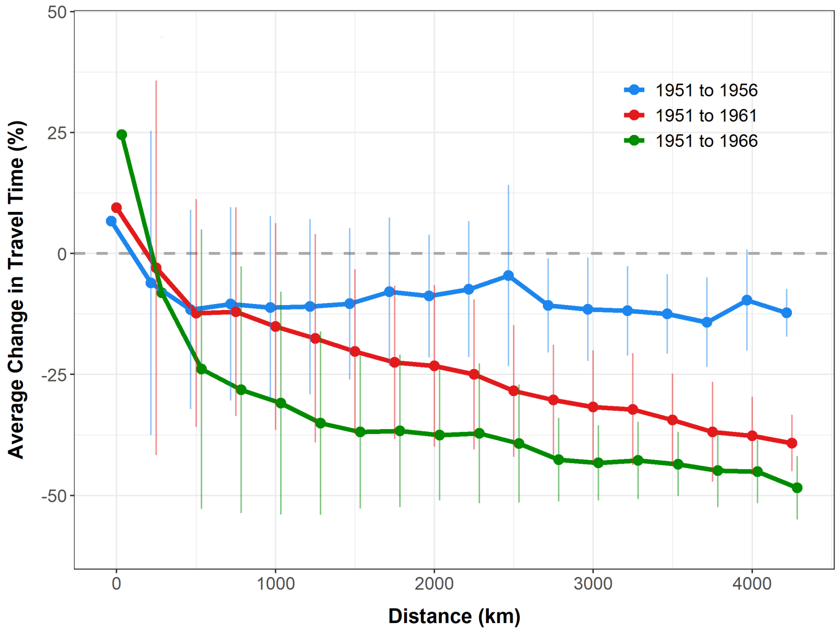

Figure 6: Change in MSAs travel time

Figure 25 in Appendix A.2 shows that the change in travel time is accompanied by a reduction of the amount of legs needed to connect two MSAs at every distance. This reduction is especially marked between 1951 and 1956, and 1961 and 1966. Between 1956 and 1961, we do not observe a big reduction in the amount of legs, implying that the decrease in travel time observed in Figure 6 between 1956 and 1961 comes from a source other than the amount of legs. In

Appendix Figure 26 we open up the change in travel time by the way an MSA pair was connected in 1951 and 1966: either directly (non-stop flight) or indirectly (connecting flight). We observe that much of the increase in travel time for MSA pairs less than 250km apart comes from routes that in 1951 were operated non-stop while in 1966 were operated with connecting flights.[18] Interestingly, for MSA-pairs more than 2,000km apart travel time reduced on average 42% for those pairs that were connected indirectly in both periods, and 51% for those that switched from indirect to direct. This fact shows the relevance of improvements in flight technology even for MSAs that were not directly connected.

It could be the case that a reduction in the amount of legs or an increase in frequency of flights reduces layover time, which then translates into a reduction of travel time. In Appendix Figure 28 we compare the change in travel time from 1951 to 1966 with a counterfactual change in travel time in which we eliminate layover time in both time periods. We observe that the average change in travel time is stronger at every distance in the fictitious scenario without layover time. This implies that the relative importance of layover time to total travel time within a route increased between 1951 and 1966, so total travel time did not decrease proportionally to the change of in-flight travel time. In short, layover time attenuated the reduction in travel time.

[18]Appendix Figure 27 repeats the exercise discarding layover time in all time periods. By comparing Figure 26 and Figure 27 we can disentangle the effect of layover time and the change in in-flight time. For MSA pairs less than 250km that changed from direct to indirect connection, 80% of the increase in travel time is due to the increase in layover time (which was previously zero as it was a non-stop flight), and 20% is due to the increase of in-flight time.

4.2. Constructing an instrument

In this section we construct an instrumental travel time that is based on the pre-existing flight routes and the time-varying nation-wide roll out of jets. In this way, the instrument abstracts from the endogenous decisions of two agents: First, regulator's decision on the opening/closure of routes. Second, airlines' decision about to which routes allocate jet vs propeller airplanes and scheduling (frequency of flights and layover time). We first explain the idea and identifying assumptions of the instrument, and then we detail how it is constructed.

In Borenstein and Rose (2014) it is argued that, due to strict regulation, it was difficult for airlines to adapt their flight network when technology to fly changed. However, we may be concerned that the decision of the regulator to grant new routes could be targeted to specific pairs or correlated with unobservable variables that also affect the creation and diffusion of knowledge.[19] Hence, as the first step in the construction of our instrument, we *fix routes* to the ones we observe in 1951. In this way the instrumental travel time is computed only using non-stop flights present in 1951, and does not consider appearance or disappearance of non-stop flights in the data. The identifying assumption is that the network of flight routes in 1951 did not yet include the changes that would be optimal to operate with jet airplanes. In other words, we require that the regulator did not change routes already by 1951 in anticipation of the arrival of jet airplanes.[20]

Airlines could decide on two factors that affect travel time: the type of airplane (jet vs. propeller) operated in each route[21] and scheduling, which consists on

[19]For example, the regulator could have targeted the opening of new routes between places in order to boost their economic activity.

[20]For example, in the instrument there are no non-stop transcontinental routes.

[21]In 1961, all non-stop flights of more than 3,000km had at least one jet operating within them,

the frequency of flights and layover time in case of connecting flights. We may be concerned that, as with the regulator, airlines' decisions could be correlated with unobservables that also affect the creation and diffusion of knowledge.[22] The second step in the construction of our instrument is to discard layover time (hence discarding all scheduling decisions) in all time periods, and that in each year all routes are operated with a *fictitious average airplane* of the year. Hence, the change in instrumental travel time in a route is independent of the type of airplane used in the route and it only depends on the nation-wide roll out of jets. The identifying assumption is that no single route had the power to shift the average speed of the year.

To construct the instrumental travel time we first estimate, separately for each year, a linear regression of travel time on flight distance using only the fastest non-stop flight in each origin-destination airport pairs.[23] These yearly regressions provide us with the fictitious average airplane of each year: the intercept gives the take-off and landing time of the airplane while the slope provides the (inverse) speed. Second, we fit these regressions to obtain predicted travel time in each non-stop flight and year. Third, for each year, we compute the fastest travel time using the Dijkstra algorithm. The Dijkstra algorithm looks for the fastest path using only 1951 non-stop flights, while the travel time in each non-stop flight in each year is given by the predicted travel time from the previous step. Layover time is set to zero in all years.

Figure 7 shows the percentage change in observed and instrumental travel

while in 1966 it was the case in all non-stop flights of more than 2,000km. Therefore the endogeneity of jet adoption is a smaller concern for long distance flights.

[22]For example, airlines may have decided to prioritize the allocation of jets to routes which had a higher share of business travel, which may be correlated with the diffusion of knowledge.

[23]The use of a linear regression is motivated by the linearity between travel time and distance displayed in Figure 5. To estimate these regressions we use all routes appearing in each year.

time relative to 1951. We compute the percentage change within each MSA-pair for each year and then take averages within 250km bins. We observe that the instrumental travel time follows pretty closely the observed change in travel time in each year. Especially, it replicates the pattern of a stronger decrease in travel time for MSAs located farther apart. It is only for MSAs less than 250-500km apart that the change in the instrumental travel time departs from the observed change.[24] This finding shows that for most of the change in travel time that we observe is due to the change in speed of airplanes, and that the endogeneity concern is limited for MSAs located far away from each other.

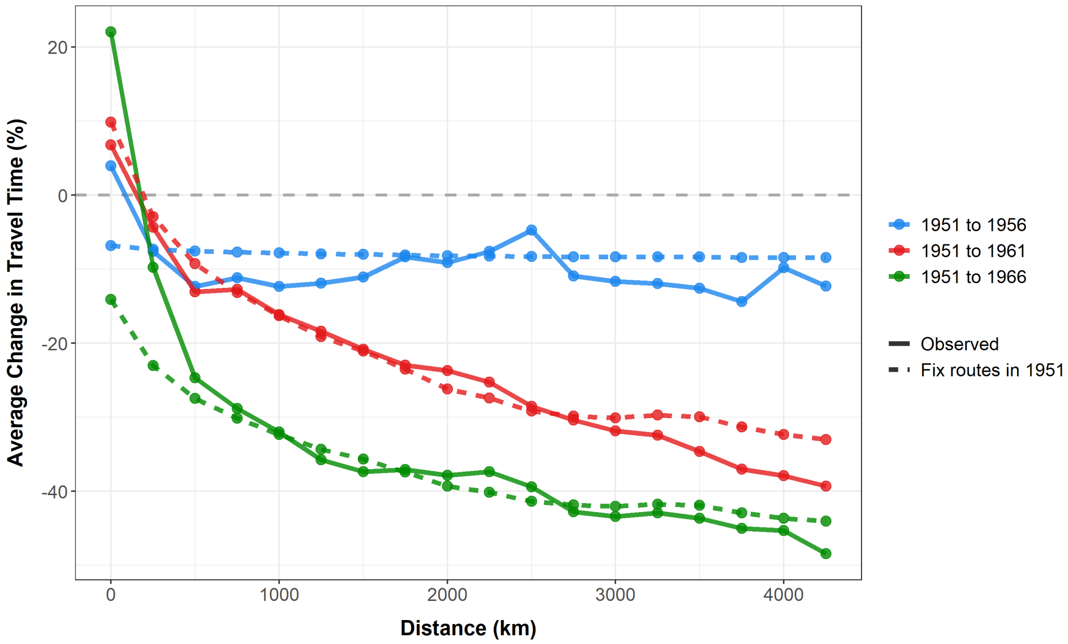

Figure 7: Instrumental Travel Time between US MSAs.

In Appendix A.2 we present other two counterfactual travel times: one in which we fix airplanes to be the average airplane of 1951 and allow routes to

[24]We observe an increase in travel time for short distances in 1961 relative to 1951. Given that non-stop routes are fixed and that for longer distances there is a decrease in travel time, the increase in travel time in short distances comes from an increase in the value of the intercept relative to the slope in 1961, relative to 1951.

evolve, and another in which both the average airplane and routes are varying. These two counterfactuals together with the one presented in this section allow us to decompose the change in travel time by the change in routes and the change in speed of airplanes. We obtain that around 90% of the change in travel time is due to the change in speed of aircrafts, while around 10% of the change is due to the change in the flight routes. Appendix Figure 30 shows that the share is roughly constant for all distances. This finding confirms that most of the observed changes in travel time are due to improvements in flight technology.

5. Patent data

We use patent data as our source of innovation information. We construct a dataset of all patents granted by the United States Patent and Trademark Office (USPTO) with filing year[25] between 1949 and 1968, which includes for each patent: filing year, technology classification, location of the inventors when they applied for the patent, owner of the patent and citations to other patents also granted in the United States. This dataset provides the distribution of patents and citations over the geographic space, allowing to take into account ownership structure.

To construct the patent dataset we downloaded from Google Patents all patents granted by the USPTO with filing year between 1949 and 1968. This dataset

[25]Filing year, also called application year, is the closest date to the date of invention that is present in the data and it represent the date of the first administrative event in order to obtain a patent. In the other hand, the publishing (also called granting year) is a later year in which the patent is granted. The difference between filing and publishing year depends on diverse non-innovation related factors (as capacity of the patent office to revise applications) and changes over time. Hence filing year is the date in our data that approximates the best to the date of invention.

contains patent number, filing year and citations.[26,27] Based on the patent number we merge it with multiple datasets. First, we obtained technology class from the USPTO Master Classification File[28] and we aggregated them to the six technology categories of Hall et al. (2001). Second, we obtained geographic location of inventors from three datasets: HistPat (Petralia et al. (2016)) and HistPat International (Petralia (2019)) for patents published until 1975, Fung Institute (Balsmeier et al. (2018)) for patents published after 1975.[29] We match all inventors' locations to 1950 Metropolitan Statistical Areas (MSAs) in contiguous United States. To do the match we obtain geographical coordinates from the GeoNames US Gazetteer file and Open Street Maps, and use the MSAs shape file from Manson et al. (2020). Third, we obtain ownership of patents from two sources: Kogan et al. (2017) for patents owned by firms listed in the US stock market and Patstat (Magerman et al. (2006)) for the remaining unmatched patents.[30]

For the descriptives presented below and the posterior analysis we truncate and aggregate the data in the following way. We drop patents that are owned by universities or government organizations. To count patents that are classified

[26] Very few patents had missing information on filing year. We complemented both missing filing year and citations with the OCR USPTO dataset.

[27] We note that the patent citation record starts in 1947, year in which the USPTO made it compulsory to have front page citations of prior art. Gross (2019)

[28] `https://www.google.com/googlebooks/uspto-patents-class.html`

[29] Due to the gap between the filing year and publishing year we also do the matching to patents published after 1968. Our underlying patent data actually covers a longer time period of filing years, which we need for example to construct forward and backward citation lags. However, there are limitations to use the geographic data in filing years 1971-1972. In Appendix B.4 we show that during filing years 1971-1972 the rate of unmatched patents to inventors' location increases. This is probably due to Histpat and Fung data not being a perfect continuation one of the other.

[30] Patent ownership in both datasets comes from the patent text, which is self declared by the patent applicant. Particularly, Kogan et al. (2017) does not explicitly state if it takes into account firm-ownership structure to determine the ultimate owner of a patent, neither does Patstat.

into multiple technology categories, we do a fractional count by assigning proportionally a part of the patent to each category. Citations are counted as the multiplication of the technology weight of the citing and cited patents. We drop patents (and their citations) that have inventors in multiple MSAs[31] and citations in which the citing owner is the same as the cited owner.[32]

We aggregate the patent data to 4 time periods of 5 years each, with the center of the period being the year of travel time data collected. The periods are: 1951 (which contains the years 1949-1953), 1956 (1954-1958), 1961 (1959-1963) and 1966 (1964-1968). We consider only patents in Metropolitan Statistical Areas (MSAs) that are matched to an airport in the four periods.[33] The final dataset has 108 Metropolitan Statistical Areas (MSAs) with patents and travel time.

5.1. Descriptive statistics: Patents

This section presents three facts about US patents over our sample period: First, initially less innovative locations had a higher patenting growth rate. The average yearly growth rate of locations in the lowest quartile of initial innovativeness was 7.2% while it was 1.9% for those in the highest quartile. High growth locations were also primarily in the South and the West of the US. The South and the West grew three times as fast as the Midwest and the Northeast. Second, over time firms grew larger as measured by the amount of MSAs in which they had research establishments, and their share of patents increased. The amount of

[31] Working with multi-MSA patents requires an assumption on how to compute distance and travel time between the citing and cited patents, as it is not a single origin-destination location pair. We hence prefer to abstract from multi-MSA patents. In the other hand, collaboration of inventors located in different MSAs is a interesting subject and it is part of our research agenda.

[32] Incentives to self-cite may be different than to cite patents of other owners.

[33] We drop around 9% of patents that are in an MSA that is not matched to an airport in the 4 time periods. Descriptive statistics including those patents are similar to the ones presented.

firms with research establishments in more than 10 MSAs almost tripled over the time period and their share of patents doubled. Third, the mass of citations shifted towards longer distances. While the first quartile of citation distance remained relative stable over the time period, the third quartile increased its distance by 39%. At the same time, the share of citations at more than 2,000km increased by 30%.

We compute descriptives by technology. In here we present descriptives of averages across technologies. Technology specific descriptives are included in Appendix B.4.

Fact 1.a.: Initially less innovative locations had a higher patenting growth rate

In the period 1951 to 1966 we observe that the highest growth of patenting takes place in locations that were initially less innovative. The differential growth rate implies a convergence rate of 5.3% per year.

Figure 8 shows the geographic distribution of patenting in 1951. Darker colors refer to a higher level of *initial innovativeness*, which is defined as the amount of patents filed by inventors in the MSA in 1951.[34] We observe that MSAs in the top quartile of patenting are concentrated in the Northeast (which includes New York) and the Midwest (which includes Chicago), with few additional MSAs in the West.[35,36]

[34]We aggregate patents to the MSA-technology level and then compute the quantile-position of each MSA in the technology. Each MSA has a different value of quantile-position in each of the 6 technology categories. To obtain the MSA level quantile we take the average quantile across technologies within the MSA. Finally we classify MSAs into quartiles depending on whether the average quantile is higher or lower than the thresholds 0.25, 0.50, 0.75.

[35]In Appendix B.4 we show that the 1951 geographic distribution of patents looks similar across technology categories.

[36]The top 5 patenting MSAs in 1951 were: New York City (25% of all patents), Chicago (11%), Los Angeles (8%), Philadelphia (6%) and Boston (4%).

Figure 9 shows the geographic distribution of patenting growth in 1951-1966.[37] We observe a striking pattern relative to Figure 8: high growth MSAs were those that were initially less innovative. High growth happens in initially less innovative locations the South and the West but also in the Northeast. We confirm this pattern in Figure 10, which shows the MSA's ranking of innovativeness in 1951 and its subsequent patenting growth rate in 1951-1966. Figure 10 shows that MSAs that were initially more innovative (lower values in the ranking) are those that saw lower values of subsequent patenting growth.[38,39] We estimate a linear regression with an intercept and a slope, and find that the slope is positive and statistically different from zero. At the mean, lowering initial innovativeness by 10 positions in the ranking was associated with a subsequent 0.42 percentage points higher yearly growth rate of patenting.

Figure 10 presents average growth rates across technologies within a MSA. If we compute the average growth rates across MSAs within a technology and quartile of initial innovativeness, and then take the average across technologies we obtain a result that goes in the same direction. The average yearly growth rate of MSA-technologies in the lowest quartile of initial innovativeness is 7.2% while it is 1.9% in the highest quartile. The percentage point difference between

[37] We compute the growth rate of patenting in each technology within a MSA and then take the average across technologies within the MSA.

[38] Each dot in Figure 10 is an MSA. To compute the MSA ranking we need to double-rank MSAs. First we rank all MSAs in each technology. Second we take the across-technology average ranking of each MSA. Third we rank all MSA's averages. To compute the MSA's yearly growth rate we first take the 1951-1966 growth rate for each technology in the MSA. We then take the average across technology. Finally we obtain the MSA's yearly growth rate by computing: $yearly\_growth\_rate = (1 + 19\_year\_growth\_rate)^{(1/19)} - 1$ (the 1951 to 1966 period is a 20 year window, we take growth rates as being from the first year 1949 to the last one 1968, which is 19 year growth).

[39] In Appendix Figure **??** we show replicate the plot differentiating geographic regions. MSAs that were initially less innovative and had high subsequent growth were located in all four regions, although they were primarily located in the South and the West.

the two growth rates implies that locations in the lowest quartile converged towards locations in the highest quartile at a speed of 5.3% per year. Appendix Figure **??** shows that the Herfindahl index of patent concentration across MSAs decreases during our sample period, a finding in line with *The Postwar Decline in Concentration, 1945-1990* described in Andrews and Whalley (2021).

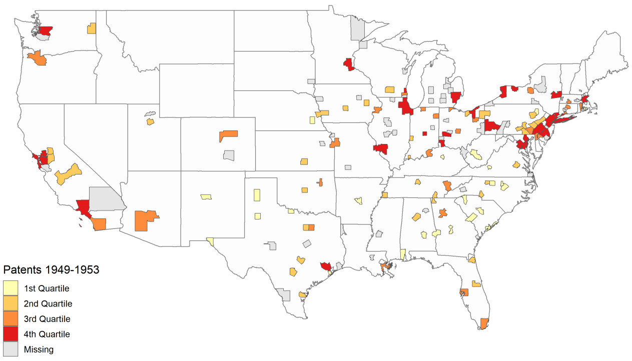

Figure 8: Geography of Patenting 1951

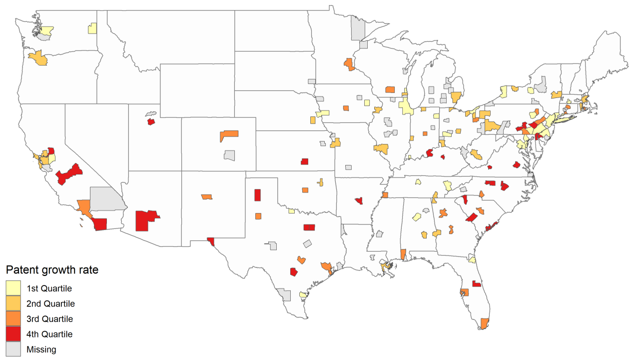

Figure 9: Patent growth 1951-1966

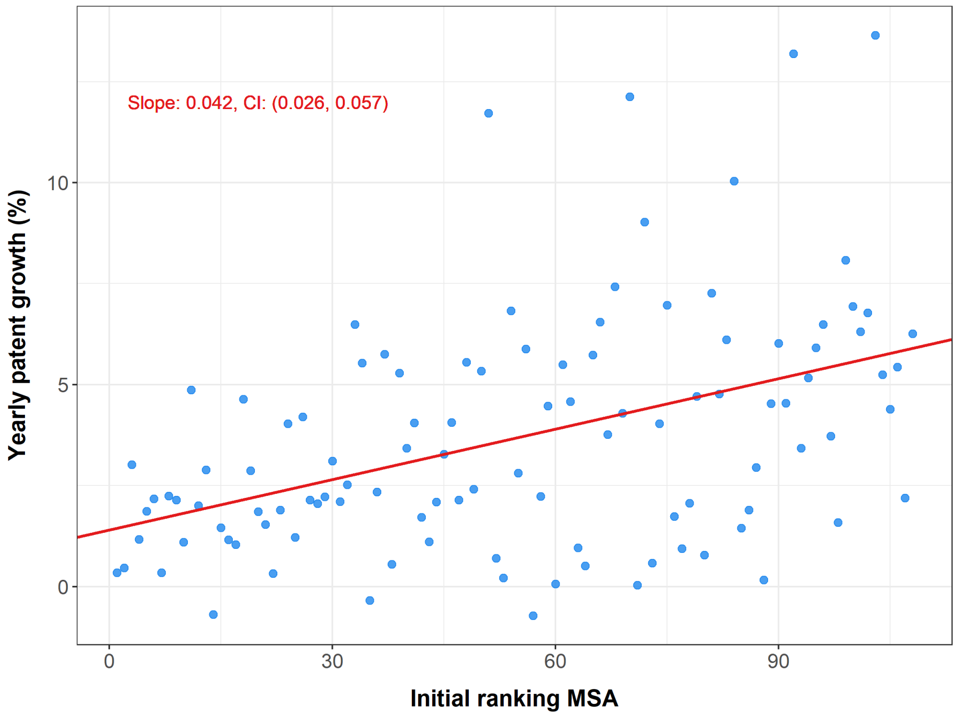

Figure 10: Patent growth by initial innovativeness ranking of MSA

Fact 1.b.: The South and the West of the US had a higher patenting growth rate

Figure 9 shows that MSAs located in the South and the West of the US had a higher patenting growth rate in 1951-1966. We classify MSAs using Census Regions of the US (Midwest, Northeast, South and West)[40] and aggregate patents within each region-technology-year. Figures 11 and 12 present averages across technologies within a region-year. Figure 11 shows that the share of patents filed by inventors located in the Midwest and the Northeast decreased from 75% in 1951 to 68% in 1966, while the share of patents filed in the South and the West increased from 25% to 32%. The opposite change in the shares implies that the South and the West had a higher growth rate of patenting relative to the Midwest and the Northeast.

[40]In Appendix Figure **??** we present a map with the four Census Regions.

Figure 12 shows that in the period 1951-1966 the South and the West increased their amount of patenting by 80%, while the Midwest and the Northeast had a 22% growth.[41] Translated into yearly growth rates, the South and the West grew three times as fast as the Midwest and the Northeast (3.14% vs. 1.05% per year).[42]

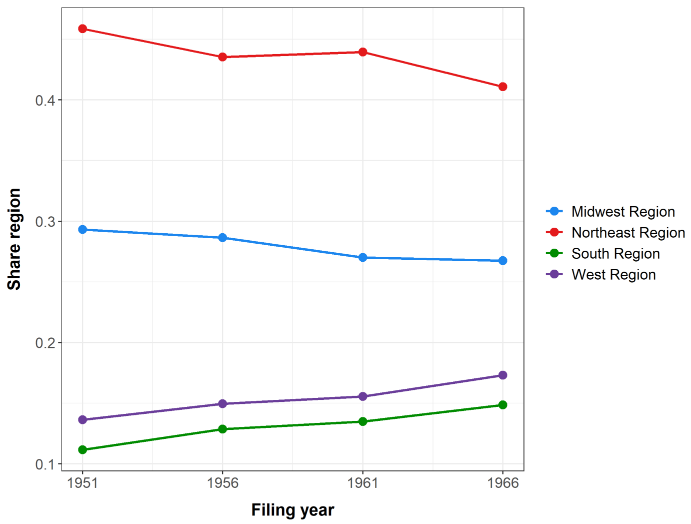

Figure 11: Share of patents by region

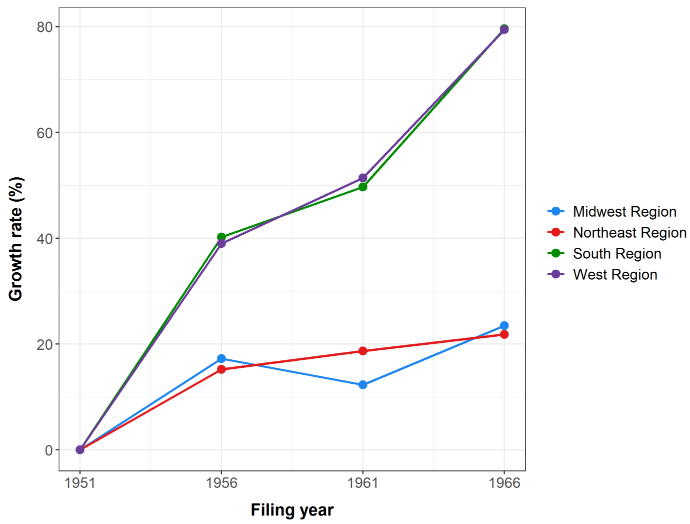

Figure 12: Patent growth by region

Fact 2: Multi-establishment firms expanded geographically and accounted for a higher share of patents

Using the patent owner identifier of patents we identify all locations in which a patent owner had inventors applying for patents. We label a patent owner a *firm* and assume that a firm has a *research establishment* in the MSAs in which it has inventors applying for patents. Combining all patents belonging to the same firm we know if a firm has research establishments in multiple MSAs, if a firm

[41]Growth rates are computed by region-technology and then averaged across technologies within region.

[42]$3.14\% = 1.80^{(1/19)} \times 100$, $1.05\% = 1.22^{(1/19)} \times 100$

expands over time and where it locates its establishments.

In Table 1 we count the number of firms and compute their share of patents according to whether the firm had 1, 2 to 5, 6 to 10, 11 to 20, or more than 20 establishments in each respective year. As we can see, the vast majority of firms had one establishment (95.8% in 1951), while very few had 11 or more establishments (0.1% in 1951). In 1951, single-establishment firms accounted for 57%. At the same time, firms with 11 or more establishments (42 firms, 0.1% of all firms) accounted for 15% of all patents.

From 1951 to 1966, the amount of single establishment firms declined by 1% while the amount of firms with 11 or more establishments increased by 283%. In other words, the amount of firms with presence in 11 MSAs or more grew from 42 to 119 firms. At the same time, the share of patents accounted by firms with 11 or more establishments increased from 15% to 31%. Simultaneously, the share of patents of single-establishment firms decreased from 57% to 46%. Hence, Table 1 illustrates that both the amount of multi-establishment firms and their share of patents grew over time.[43] In Appendix B.4 Figure **??** we show that multi-establishment firms increased their share of patents in all quartiles of MSAs' initial innovativeness, with a stronger increase in initially less innovative MSAs.

While we observe an increase in the number of multi-establishment firms, we also observe an increase in the distance between establishments of the same firm. Figure 13 shows that, for firms that have multiple establishments in the respective

[43]Within each year and bin of firm size, we compute the share of patents by technology and then take the average across technologies. We have computed the across-firms Herfindahl index within technology (so it shows the level of across-firm concentration within a technology) and we do not observe a clear pattern of either concentration or deconcentration.

Year \ N. estab.	Number of firms					Share of patents				
	1	2 to 5	6 to 10	11 to 20	+20	1	2 to 5	6 to 10	11 to 20	+20
1951	41,133	1,684	75	34	8	0.57	0.19	0.08	0.07	0.08
1956	42,590	1,927	111	60	12	0.52	0.19	0.09	0.11	0.08
1961	37,366	2,112	131	80	18	0.48	0.19	0.09	0.13	0.12
1966	40,711	2,086	132	89	30	0.46	0.15	0.09	0.14	0.17

Table 1: Number of firms and share of patents by firm's geographic coverage
Geographic coverage is computed as the amount of Metropolitan Statistical Areas (MSAs) in which the firm has inventors applying for patents (*research establishments*) in a certain year. Bins of geographic coverage are 1 MSA, 2 to 5 MSAs, 6 to 10 MSAs, 11 to 20 MSAs, more than 20 MSAs. The maximum possible is 108 MSAs.

year, the average distance across establishments within the firm increased over time.[44]

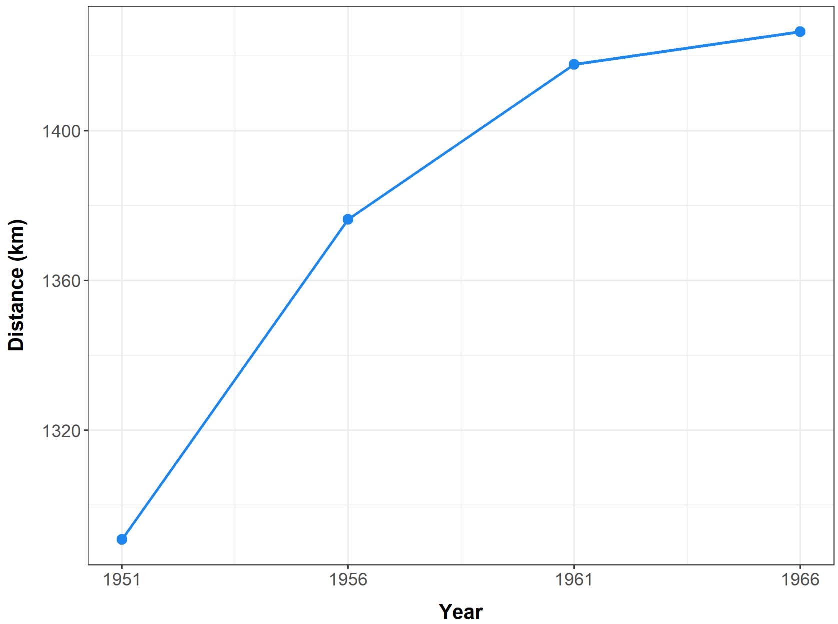

Figure 13: Average distance across establishments within the firm

[44]The increase in distance across establishments within firms could well be the result of firms that are growing and randomly producing new patents in different locations. However, in Section 8 we show that the process firms' geographic expansion was not random: firm's expansion was directed towards locations that got larger reductions in travel time to the firm's headquarters.

Fact 3: Distance of citations increased

In our analysis we use citations as a proxy for knowledge diffusion. According to Jaffe et al. (1993) *"a citation of Patent X by Patent Y means that X represents a piece of previously existing knowledge upon which Y builds."* (page 580).[45] We compute the distance between the citing inventor and the cited inventor. Figure 14 shows the evolution over time of the first, second and third quartile of citation distance.[46] We observe that 25% of citations happened between inventors located less than 300km apart throughout our sample period. For the middle 50% of citations we observe that over time inventors cited other inventors located farther away. The third quartile of citation distance increased from 1,642km in 1951 to 2,284km in 1961, a 39% increase in the distance.[47] In other words, the mass of citations shifted towards longer distances.

In Figure 15 we present the share of citations by distance range between the citing and cited inventors.[48] The distance cutoffs where chosen in order to have a balanced shared of citations in the initial time period, and considering the changes in travel time presented in Section 4.1. The share of citations that happen between inventors located more than 2,000km apart grew from 21.5% in 1951 to 27.9% in 1966. The 6.4 percentage points increase represents an increase of 30% of the share of citations at more than 2,000km.

[45] Jaffe et al. (1993) discusses the reasons why to cite and why not to cite. Using a survey of inventors, Jaffe et al. (2000) find that there is communication among inventors and citations are a *"noisy signal of the presence of spillovers."*

[46] We compute distance between MSA centroids.

[47] As a reference, the straight line distance from New York City NY to other places is: Boston MA 300km, Chicago IL 1,140km, Dallas TX 2,200km, San Francisco CA 4,130km. The quantile 0.10 of was at 0km in every period, implying that 10% of citations took place within MSA. The quantile 0.90 was around 3,611km to 3,789km over the time sample period.

[48] While Figure 14 shows how the distance of each quantile changes over time, Figure 15 shows the mass of citations (and hence the quantile to which belongs) in a certain distance cutoff. For example, in 1951 the share of citations in the 0-300km range was 31.6%, which is equal to saying that the quantile 0.316 in 1951 was 300km.

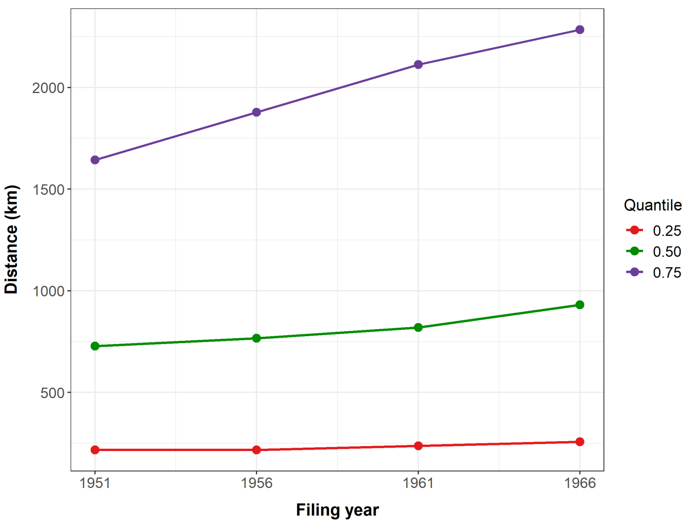

Figure 14: Quantiles of citation distance

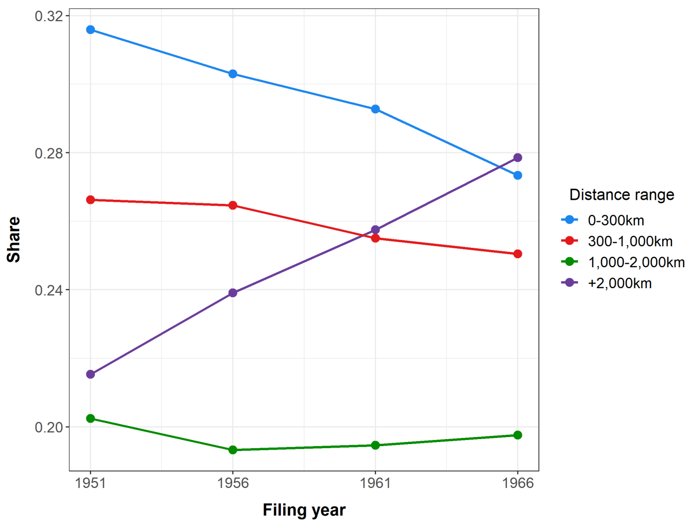

Figure 15: Share of citations by distance

6. Diffusion of knowledge

In this section we show that the reduction in travel time led to an increase in knowledge diffusion, especially over long distances. In doing so we estimate the parameter β highlighted in equation (2): the elasticity of knowledge diffusion to travel time.

To perform the analysis we merge the Air Travel and Patent datasets to obtain a final dataset that contains for each patent owner-location, the amount of patents filed in a certain 5-year period and technology class, the amount of citations to other patents with their respective owner identifier, location and technology class, and the travel time to every location. We aggregate citations to the citing-cited establishment-technology within each period. We assume that passengers take a return flight, hence we make travel times symmetric.[49]

[49] $travel\ time_{ijt} = (travel\ time^{original}_{ijt} + travel\ time^{original}_{jit})/2$ where $travel\ time^{original}_{ijt}$ stands for the travel time between MSA i and j at time period t.

We estimate a gravity equation which relates citations between two establishments-technologies with their pairwise travel time.[50] We estimate the following regression:

$$citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt} \tag{3}$$

where $citations_{FiGjhkt}$ is the amount of citations from patents filed by the establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. We call Fi the research establishment of firm F in location i. travel time$_{ijt}$ is the air travel time (in minutes) between location i and j at time period t. The parameter of interest in the regression is $\boldsymbol{\beta}$, which represents the elasticity of citations to travel time.[51] If citations are affected negatively by travel time we would expect a negative value of β.

Given the panel structure of our data, we can include the fixed effect FE_{FiGjhk} that absorbs any time invariant citation behavior within the *citing establishment-technology and cited establishment-technology*. This fixed effect flexibly controls for persistent relationships within an establishment pair that would lead to relatively more (or less) citations. That includes characteristics like physical distance, but also pre-existing commercial relationships between establishments. The fixed effects FE_{Fiht} and FE_{Gjkt} control for the time changing general level of citations specific to each establishment and technology. For example FE_{Fiht} controls for the fact that if Fih files more patents in a given period, it would mechanically make more citations to every establishment. On the other hand, FE_{Gjkt} controls

[50] For explanation and micro foundations of the gravity equation see Head and Mayer (2014) and references thereof.

[51] A 1 percent increase in travel time has an effect of β percent increase (or decrease in the case of a negative β) in citations.

for Gjk filing more patents or higher quality patents that would receive more citations from every establishment.[52]

The inclusion of FE_{FiGjhk} implies that only variation across time within an establishment-pair is used for identification. By additionally including the fixed effect FE_{Fiht}, the across-time variation is compared only between citing-cited establishment-technology pairs $FiGjhk$ within a citing establishment-technology Fih in period t. As we also include FE_{Gjkt}, the comparison is done while controlling for the size of the cited establishment-technology Gjk in period t. Put differently and simplifying slightly, the identification of $\boldsymbol{\beta}$ relies on changes in citations and travel time within an establishment-pair, relative to another establishment-pair with the same citing establishment, conditional on the two cited establishments' sizes.

Following Silva and Tenreyro (2006), we estimate the gravity equation by Poisson Pseudo Maximum Likelihood (PPML).[53] This estimation methodology has two advantages over a multiplicative model that is then log-linearized to obtain a log-log specification. First, it only requires the conditional mean of the dependent variable to be correctly specified, while the OLS estimation of the log-linearized model would lead to biased estimates in the presence of heteroskedascity. Second, it allows to include zeros in the dependent variable, which is especially relevant when using disaggregated data. One downside of estimating PPML with the fixed effects that we include is that both coefficients and standard errors have to be corrected due to the incidental parameter problem (Weidner and Zylkin (2021)). We follow Weidner and Zylkin (2021) to use split-panel jackknife

[52]In the International Trade literature, the parallel of the fixed effects (simplified for exposition) would be: FE_{ij} country-pair fixed effect, FE_{jt} origin-time fixed effect and FE_{it} destination-time fixed effect.

[53]We use the package *fixest* (Bergé (2018)) in R to estimate high dimensional fixed effects generalized linear models *feglm* with Poisson link function.

bias-correction on the coefficients and Dhaene and Jochmans (2015) to bootstrap standard errors which we also bias-correct with split-panel jackknife.[54]

Whenever *FiGjhk* has positive citations in at least one period and missing value in another, we impute zero citations in the missing period.[55] Travel time is set to one minute whenever $i = j$.[56]

Column (1) in Table 2 presents the results of estimating equation (3). The value of the elasticity of citations to travel time is estimated to be -0.083, statistically significant at the 1% level. Given the average reduction in travel time of 31.4% in the full estimating sample, the elasticity implies that citations increased on average 2.6% as consequence of the reduction in travel time. If we consider the average decrease in travel time across all MSAs in the baseline travel time data, the implied increase is 2.4%.[57]

The importance of air transport relative to other means of transport potentially depends on the distance to travel. Also, we observed in section 4.1 that the improvements in air travel time depended on the distance to travel, with a difference in jet adoption for travel distances under and over 2,000km. Taking

[54] Details on the bias correction and bootstrap procedures are provided in Appendix **??**.

[55] We do not impute zeros in *FiGjhk* that are always zero, as those observations would be dropped due to not being able to identify FE_{FiGjhk}.

[56] We measure air travel time in minutes. In our sample 13% of citations happen within the same MSA. The inclusion of those citations in the estimation increases the amount of observations available to identify of FE_{Fiht} and FE_{Gjkt}, and hence keeping them increases the amount of *FiGjhkt* that remain in the effective sample to identify β. In order to include them we then need to impute a within-location travel time. We assume that within-location (air) travel time is not changing across time periods. Nonetheless, the identification of β is not affected by the value chosen for the within-location (time invariant) travel time, as β is identified by across time variation. In the appendix we show results using other values of (time invariant) within MSA travel time and the coefficients remain equal.

[57] These values come from the multiplication of the elasticity of citations to travel time 0.083 and the average decrease in travel time between 1951 and 1966: 31.4% in the full estimating sample and 28.7% in the raw data of travel time across MSAs.

	PPML		IV PPML	
Dep. variable: *citations*	$cit_{FiGjhkt}$		$cit_{FiGjhkt}$	
	(1)	(2)	(3)	(4)
log(travel time)	-0.083^{***} (0.019)		-0.152^{***} (0.029)	
log(travel time) × 0-300km		0.019 (0.036)		-0.076 (0.221)
log(travel time) × 300-1,000km		-0.089^{***} (0.023)		-0.134^{***} (0.044)
log(travel time) × 1,000-2,000km		-0.094^{***} (0.033)		-0.112^{**} (0.047)
log(travel time) × +2,000km		-0.169^{***} (0.039)		-0.203^{***} (0.043)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010
R2	0.88	0.88	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 2: Elasticity of citations to travel time

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\boldsymbol{\beta}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (2) includes the interaction of travel time$_{ijt}$ with a dummy for distance bin between the citing establishment Fi and the cited establishment Gj. Column (3) and (4) show the result of two step instrumental variables estimation, where $log(\text{travel time}_{ijt})$ is instrumented with $log(\text{travel time}_{ijt}^{\text{fix routes}})$, the travel time that would have taken place if routes were fixed to the ones observed in 1951 and in each year routes were operated with the average airplane of the year. Bootstrap standard errors clustered are presented in parentheses. The coefficients and standard errors in columns (1) and (2) are jackknife bias-corrected. R2 is computed as the squared correlation between observed and fitted values.

these two characteristics into account, we estimate a variation of equation (3) in which we allow the elasticity of citations to travel time to vary by distance interval between the locations of citing and cited establishments.[58] Column (2) in Table 2 shows the result of this estimation.[59] The estimated value of the elasticity in absolute terms increases with distance, reaching -0.169 for distances of more than 2,000km. Between 1951 and 1966 the average change in travel time in the full estimating sample is 47.7% for a distance of more than 2,000km. The estimated elasticity implies that citations between establishments at more than 2,000km apart increased by 8.1% due to the decrease in travel time. In total citations at more than 2,000km increased by 21%, implying that the change in travel time can account accounts for 38.2% of the observed increase. If instead we consider the 40.8% average reduction in travel time across MSAs in the raw data, the elasticity implies an increase in citations of 6.9%, accounting for 32.7% of the total citation increase.

In Appendix B.4 we investigate different heterogeneous effects. We study how travel time affects the extensive margin of citations (whether an establishment cites another establishment or not) and the intensive margin (conditional on citing, how much it cites). We find the effect comes from both margins. We estimate an heterogeneous elasticity depending on the level of spatial concentration of the citing technology and the cited technology, we do not find a statistical difference. We also look at whether it is older patents or younger patents that get diffused, finding some slight evidence that it is technologies that take longer time to diffuse that increase more their diffusion with the reduction in travel time. We study citations to and from government patents, and self citations, on the whole we do not find a different pattern from the baseline. We also do not

[58] We compute distance between the geographical center of each MSA.

[59] The share of observations (citations) in each distance interval is: 0-300km 26.1% (28.5%), 300-1,000km 30.7% (28.5%), 1,000-2,000km 19.7% (23.4%), +2,000km 23.4% (19.6%).

find a particular pattern of the elasticity depending on the citing *firm's size* as measured by the amount of patents filed in 1949-1953. Finally, we estimate the elasticity by citing and cited technology and most of the effect seems to come when the citing and cited technologies are the same.

As mentioned in Section 4.2, we may be concerned that the timing and allocation of jets to routes and that the opening/closure of routes were not random. In case there is an omitted variable that affects the change in travel time at the MSA pair level and is correlated with citations across establishments within the same MSA, we would estimate biased coefficients. In order to tackle the endogeneity concern due to omitted variable we do an instrumental variables estimation using the instrument proposed in Section 4.2. To implement the instrumental variable estimation we follow a control function approach described in Wooldridge (2014). We proceed in two steps estimating the following two equations:

$$\begin{aligned} log(\text{travel time})_{FiGjhkt} &= \lambda_2\, log(\text{travel time}^{\text{fix routes}}_{FiGjhkt}) \\ &\quad + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + u_{FiGjhkt} \end{aligned} \tag{4}$$

$$\begin{aligned} citations_{FiGjhkt} &= \exp\,[\beta\, log(\text{travel time}_{ijt}) + \lambda\, \hat{u}_{FiGjhkt} \\ &\quad + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}] \times v_{FiGjhkt} \end{aligned} \tag{5}$$

In a first step we estimate equation (4) and obtain estimated residuals $\hat{u}_{FiGjhkt}$. In a second step we use the estimated residuals as a regressor in equation (5) which *controls* for the endogenous component of travel time. To perform inference we bootstrap standard errors.[60] According to Wooldridge (2014), there would be evidence of endogeneity if the parameter λ in equation (5) is estimated to be statistically different from zero.

[60] Appendix **??** includes details on the bootstrap procedure.

Columns (3) and (4) of Table 2 show the results of the instrumental variables estimation. If airlines were allocating jet airplanes to routes that would have witnessed a higher degree of exchange of knowledge even in the absence of jets, then we would expect the instrumental variables estimate to be smaller in absolute terms relative to the baseline coefficient. On the other hand, if the regulator targeted the opening of new routes between places that were in a lower trend of exchange of knowledge, we would expect the instrumented coefficient to be larger in absolute terms. Column (3) estimates the elasticity to be -0.152, bigger in absolute value compared to the non-instrumented estimate. The instrumental variables corrects for a downward bias in absolute terms, which represents evidence in favor of the regulator targeting the opening of new routes between places that had a lower degree of exchange of knowledge.

In column (4) of Table 2 we see the coefficients of the instrumental variable estimation by distance between the citing and cited establishments. We observe the presence of a bias in the same direction as in column (3), however the magnitude of the bias is smaller except for the distance bin 0-300km, which is not precisely estimated. In particular, at more than 2,000km, the coefficient is relatively similar to the baseline estimation. In Appendix Table **??** we show the regression including coefficients on the residual *controls*. If the coefficients on controls are statistically significant, that is evidence of endogeneity. While the control is statistically significant when using only one coefficient for all distance, none of them is statistically significant when opening the coefficient by distance range. In other words, we do not find evidence of endogeneity at long distances, especially at +2,000km.

6.1. Diffusion of knowledge: Robustness

We may be concerned that there are other variables that could drive the diffusion of knowledge and at the same time be correlated with the change in travel time. The omission to control for this variables, if they exist, would bias the estimated coefficients.[61] We consider three potential variables that could bias our estimates: improvements in highways, improvements in telephone, flight ticket prices. In Table 3 we show the results controlling for this variables separately, while in Appendix Table **??** we include them simultaneously. Estimates are robust to including these controls.

Column (1) of Table 3 repeats the result of column (2) in Table 2. Column (2) repeats column (1) without bias correction.[62] We observe that not doing the bias correction does not qualitatively affect results. Columns (3) to (6) include the additional controls and should be compared to column (2).

First, in 1947 the Congress published the official plan for the Interstate Highway System, a nation-wide infrastructure plan to improve existing highways and build new ones (see Baum-Snow (2007), Michaels (2008), Jaworski and Kitchens (2019) and Herzog (2021)). In case the change in travel time by air is correlated with the change in travel time by highway, we would have an omitted variable bias if we include only one of them in the estimation. Taylor Jaworksi has graciously shared with us data on county-to-county highway travel time and travel costs for 1950, 1960 and 1970, which we converted to MSA-to-MSA and linearly interpolated to convert to the same years of our air travel data. Hence we have

[61] Variables that are not time changing or that are time changing at the MSA or establishment level do not represent a threat to identification, as they are flexibly controlled for with the fixed effects.

[62] The jackknife bias-correction due to the incidental parameter problem is computationally intensive. Due to the computational burden, we have still not bias-corrected all coefficients. Columns (2) to (6) of Table 3 do not include bias-correction.

a MSA-to-MSA time-varying measure of travel time. In Appendix we show the correlation of MSA-to-MSA change in air travel time and highway travel time. Column (3) of Table 3 includes log highway time as control, the estimated coefficients remain relatively unchanged.

Second, we may believe that other means of communication like telephone lines expanded or changed their price during the period of analysis. Haines et al. (2010) data file contains information on the share of households with telephone lines in 1960.[63] We aggregate the variable to the MSA level. For each MSA-pair, we take the log of the mean share of households with telephone lines[64]. To include the variable as control, we make the measure time variant by interacting it with a time dummy. The assumption behind the interaction is that, if telephone lines expanded or changed their price over the time period, this time-change specific to each year was proportional to the 1960 log mean share of the MSA-pair. Column (4) shows the results.

Third, we control for a time varying effect of distance on citations. We may believe that other variables may have an effect on the diffusion of knowledge, and those variables are related to the distance between the citing and cited establishments. In column (5) we include as control $log(distance)$ interacted with a time dummy. We observe that the coefficients reduce in magnitude, potentially due to the fact that the change in travel time is also correlated with distance, hence controlling for a time-varying effect of distance absorbs part of the effect. In spite of that, the coefficient for distance of more than 2,000km remains statistically significant at the 5%.

[63] Data from the 1962 City Data Book which comes from the US Bureau of the Census.

[64] log(mean telephone share$_{ij}$ = log((telephone share$_i$+telephone share$_j$)/2). We take the log of the mean share because the share is a linear combination of origin MSA and destination MSA characteristics, hence perfectly explained by origin and destination fixed effects. Taking the log prevents this.

Fourth, during the period of analysis ticket prices were set by the Civil Aeronautics Board, so airlines could not set prices of their own tickets. Some airlines included a sample of prices in the last page of their booklet of flight schedules a sample of prices, which we digitized. In appendix A.2 we document multiple facts about prices. The relevant fact for this section is that prices were relatively constant until 1962-1963, years in which we observe a drop in prices of around 20% for routes of more than 1,000km distance. We may be concerned that the change in flow of knowledge is actually consequence of a change in prices, which happens to be correlated with the change in travel time. Given that we do not have ticket prices for each route and year, we use an estimated route price which is time varying. We obtain estimated prices by using the sample of prices that we digitized and fitting, for each year, price on a third degree polynomial of distance between origin and destination. Column (6) shows the result of controlling for log ticket prices: coefficients barely change.[65]

Finally, as we will see in section 8.2, during the time period there was entry and exit of research establishments that was not uniform across locations. We may then think that the change in diffusion of knowledge is only consequence of the change in the geographical location of innovation. In Appendix Table **??** we re-estimate equation 3 with different samples: first, using only citing establishments that were present in 1949-1953, and second using only citing and cited establishments that were present in 1949-1953. We find that the coefficient at more than 2,000km remains comparable to the one in the baseline regression, statistically significant at the 1%.

[65] In order to perform inference we should adjust standard errors by the fact that we have a predicted regressor as control variable.

	PPML bias-corrected	PPML not bias-corrected	Highway time	Telephone × time	Distance × time	Price
Dep. variable: *citations*			$cit_{FiGjhkt}$			
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time) × 0-300km	0.019	0.021	0.023	0.0198	0.032	0.025
	(0.036)	(0.039)	(0.039)	(0.039)	(0.040)	(0.038)
log(travel time) × 300-1,000km	-0.089^{***}	-0.099^{***}	-0.096^{***}	-0.094^{***}	-0.075^{**}	-0.102^{***}
	(0.023)	(0.027)	(0.028)	(0.027)	(0.030)	(0.027)
log(travel time) × 1,000-2,000km	-0.094^{***}	-0.093^{**}	-0.089^{**}	-0.071^{*}	-0.040	-0.104^{**}
	(0.033)	(0.042)	(0.044)	(0.042)	(0.052)	(0.042)
log(travel time) × +2,000km	-0.169^{***}	-0.185^{***}	-0.180^{***}	-0.172^{***}	-0.124^{**}	-0.196^{***}
	(0.039)	(0.049)	(0.050)	(0.050)	(0.059)	(0.049)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010
R2	0.88	0.88	0.88	0.88	0.88	0.88
Controls:						
log(highway time)	-	-	Yes	-	-	-
log(telephone share) × time	-	-	-	Yes	-	-
log(distance) × time	-	-	-	-	Yes	-
log(price)	-	-	-	-	-	Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 3: Robustness: Elasticity of citations to travel time
Part 1

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp[\sum_d \boldsymbol{\beta}_d \mathbb{1}\{distance_{ij} \in d\} \, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \mathbb{1}\{distance_{ij} \in d\} \mathbb{1}\{X_{FiGjhkt}\} \, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0 - 300km]$, $(300km - 1000km]$, $(1000km - 2000km]$, $(2000km - max]$. Column (1) presents jackknife bias-corrected coefficients and bias-corrected bootstrap standard errors. Column (2) repeats column (1) without bias-correction. Relative to (2), columns (3) through (6) contain additional controls. Column (3) controls for log highway time between i and j at period t. Column (4) controls for the log of the mean share of households with telephone line in 1960 in ij pair interacted with a time dummy. Column (5) controls for log distance ij interacted with a time dummy. Column (5) controls for log flight ticket price between i and j at period t. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Columns (2) through (6) present standard errors clustered at the non-directional location in parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

7. Creation of knowledge

In this section we show that the reduction in travel time led to an increase in knowledge creation by increasing knowledge access. We show that the effect on the increase in new knowledge is stronger in initially less innovative locations, leading to convergence across locations in terms of innovation. This convergence contributes to a change in the geographic distribution of knowledge creation, increasing the relative importance of locations in the South and the West of the United States.

We construct a measure of *Knowledge Access* by adapting equation (2) to an empirical set up with multiple technology categories and time periods. The measure of *Knowledge Access* (KA_{iht}) shows how *easy* it is in time period t for location i and technology h to access knowledge created in other locations.[66] We compute *Knowledge Access* as follows:

$$KA_{iht} = \sum_{k} \omega_{hk} \sum_{j, j \neq i} \text{Patent stock}_{jk,t=1953} \times \text{travel time}_{ijt}^{\beta} \tag{6}$$

Where, from right to left, travel time$_{ijt}^{\beta}$ is the travel time between locations i and j at time period t, to the power of the elasticity of diffusion of knowledge to travel time. Patent stock$_{jk,t=1953}$ is the discounted sum of patents produced in location j and technology k between 1941 and 1953.[67] ω_{hk} is the share of citations of technology h that go to technology k at the aggregate level in 1949-1953,

[66] This measure is also related to measures of network centrality: nodes are locations weighted by their patent stocks and edges are travel time between locations. A higher value of the centrality measure is associated with a more central node.

[67] Patent stock$_{jk,t=1953} = \sum_{y \in [1941,1953]} \text{Patents}_{jky} \times (1 - \text{depreciation rate})^{1953-y}$. We use a depreciation rate of 5%, which is the the range of average depreciation rates of R&D found by De Rassenfosse and Jaffe (2017). We decided to fix the patent stock and not to allow it to change over time, as changes in travel time will potentially lead to changes in patent stock creating a dynamic reinforcing effect between knowledge access and new knowledge. In this sense, we abstract from *dynamic* externalities that could be at play.

similar to an input-output weight.[68] Then, KA_{iht} is a weighted sum of the patent stock in each other location and technology, where the weights are how easy it is to access that patent stock (travel time$^{\beta}_{ijt}$) multiplied by how relevant that knowledge is (ω_{kh}). In order to reduce concerns of potential endogeneity of accessing knowledge and creating knowledge, we exclude the patent stock in the location itself from the sum (we only use $j \neq i$).[69]

The measure of *Knowledge Access* is only time varying due to the change in travel time between locations, every other component of the measure is fixed. The degree with which changes in travel time are reflected in access to knowledge depend on how *important* travel time is to get knowledge to diffuse, which is exactly the elasticity of knowledge diffusion to travel time that we estimated in Section 6. As the baseline we use $\beta = 0.185$, which is the elasticity of citations to travel time at more than 2,000 km not bias corrected. In robustness we use distance-specific β and in Appendix C we do sensitivity analysis of the results to changing the value of β.

Figure 16 depicts the time change in log *Knowledge Access* from 1951 to 1966, averaged across technologies within each MSA. Darker colors represent the higher growth in *Knowledge Access*. We observe that MSAs that had the strongest growth are generally located in the South and the West of the United States,

[68] $\omega_{hk} = citations_{hk,t=[1949,1953]} / citations_{h,t=[1949,1953]}$ is included to weight each *source* technology category k by how important it is for the *destination* technology category h. The inclusion of ω_{hk} provides across-technology variation within a location-time.

[69] The theory makes no distinction on whether the knowledge stock is in i or j, so in principle we would like to include the patent stock of i in the knowledge access of i. However, this could lead to econometric problems. First, we do not have exogenous variation of travel time within i. Second, if knowledge creation in i is a persistent process, by including the patent stock of i we would introduce a mechanical relationship between knowledge access and knowledge creation. Hence, our baseline measure of knowledge access of i does not consider the patent stock of i. In Appendix **??** we show that the inclusion of i's patent stock does not affect the results.

far from the knowledge centers of New York and Chicago. The reduction in travel time was larger between locations far apart, implying that locations which happened to be far from knowledge centers increased relatively more their *Knowledge Access*.

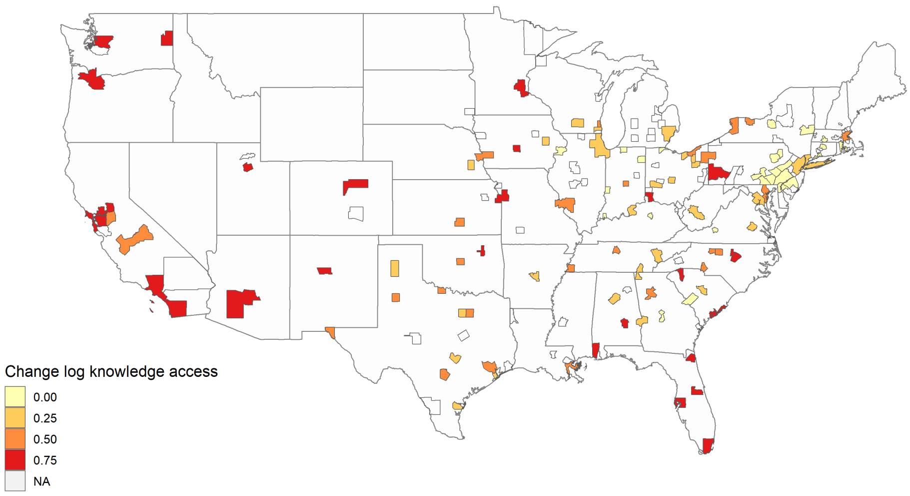

Figure 16: Change in log Knowledge Access 1951 - 1966

With the measure of *Knowledge Access* we then adapt equation (1) to estimate:

$$\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht} \qquad (7)$$

where Patents_{Fiht} are patents applied by establishment of firm F in location i and technology h at time period t. The measure of knowledge access KA_{iht} is at the iht location-technology-time level, implying that all establishments within an iht share the same level of access to knowledge. The parameter of interest ρ is the elasticity of (the creation of new) patents to knowledge access. If knowledge access is relevant for the creation of new knowledge we would expect a positive and significant coefficient.

The fixed effect FE_{Fih} absorbs time invariant characteristics at the firm-location-technology level, as for example the productivity of the establishment-technology. This fixed effect is more fine grained than just a location-technology, which would absorb the comparative advantage of a location in a certain technology. The fixed effect FE_{it} absorbs characteristics that are time variant at the location level. For example, changes in R&D subsidies that are location specific and common across all technologies would be absorbed by this fixed effect. Also, better flight connectivity could spur economic activity as shown in Campante and Yanagizawa-Drott (2017), leading to an increase in patenting activity in the location. If that increase is general across technologies within the location, then FE_{it} would absorb it. Finally, the fixed effect FE_{ht} absorbs characteristics that are time variant at the technology level. If technologies had different time-trends at the national level, then the fixed effect would control for trends in a flexible way.

The inclusion of FE_{Fih} implies that only across-time variation within an establishment-technology is used to identify ρ. The across-time variation in the measure of knowledge access is coming only from the change in travel time. The inclusion of FE_{it} implies that only variation across-technologies within a location-time is used, so across-time variation is compared across establishments within a location, and not across locations. The inclusion of FE_{ht} implies that the identifying across-time variation is conditional on aggregate trends of the sector. In other words, identification of ρ relies on across-time changes in the amount of patents and knowledge access of an establishment, relative to other establishments in the same location, conditional on aggregate technological trends.

	PPML	PPML q innovation	IV PPML	IV PPML q innovation
Dependent Variable: *Patents*		$Patents_{Fiht}$		
	(1)	(2)	(3)	(4)
log(knowledge access)	10.14*** (3.66)	9.36** (3.69)	11.24* (6.35)	10.26 (6.38)
log(knowledge access) × 3rd quartile		2.05*** (0.58)		2.32*** (0.66)
log(knowledge access) × 2nd quartile		3.80*** (0.90)		4.21*** (0.84)
log(knowledge access) × 1st quartile		5.00*** (1.30)		5.77*** (1.11)
R2	0.85	0.85	0.85	0.85
N obs. effective	991,480	991,480	991,480	991,480

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 4: Effect of knowledge access on patents, by MSA innovativeness quartile

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed using patents filed in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Column (3) and (4) show the result of two step instrumental variables estimation, where KA_{iht} is instrumented with $\widetilde{KA}_{iht}$, knowledge access computed using the counterfactual travel time that would have taken place if routes were fixed to the ones in 1951 and each year routes were operated at the average aggregate flying speed of the year. Standard errors are presented in parentheses. Column (1) and (2) present clustered at the location-technology ih. Column (3) and (4) present bootstrap standard errors. R2 is computed as the squared correlation between observed and fitted values.

Column (1) in Table 4 shows the result of estimating equation (7). The elasticity of patents to knowledge access is estimated to be 10.14, significant at the one percent level. The average change in knowledge access at the location-technology level[70] is 9%, implying that on average the change in travel time predicts a 3.5% average yearly growth rate of patents.[71] The observed average yearly growth

[70]Due to entry, we cannot compute the growth rate at the establishment-technology level for 70% of establishment-technology, given that they had 0 patents in the initial time period. In the case of location-technology, 5% did not have patents in the initial period. We the prefer to interpret coefficients using location-technology growth rates, which we compute using the remaining 95% of location-technologies that had positive patents in the initial time period.

[71]The elasticity of 10.14 predicts an increase of 91.3% over the time period of 19 years (10.14 ×

rate of new patents at the location-technology is 4.4%.[72] Comparing the predicted and observed growth rates, the improvement in air travel time has the power to account for 79.5% of the observed average yearly patent growth rate.[73] We note that this is a partial equilibrium analysis and should not be taken as the aggregate effect of travel time on patents. For example, the analysis does not take into account potential reallocation of innovative activity across locations which may happen as consequence of changes in travel time.

In the data we observe that the aggregate amount of new patents grew by 33.8% in 19 years,[74] which is an average yearly growth rate of 1.5%, considerably smaller than the across technology-location unweighted average growth rate. That implies that initially less innovative locations had a higher growth rate of new patents, relative to initially more innovative ones. In column (2) in Table 4 we investigate if the increase in knowledge access had an heterogeneous effect on the amount of new patents created depending on the initial innovativeness of the location.[75] We compute the quartile of innovativeness of location i in technology h in the time period 1949-1953 and interact it with $\log(KA_{iht})$. We use as reference category the highest quartile of initial innovativeness, hence the coefficient on $\log(KA_{iht})$ without interaction is the elasticity for the highest quartile. Coefficients on other quartiles should be interpreted relative to the highest quartile.

0.09 = 0.913), which translates into a 3.5% average yearly growth rate ($(1+0.913)^{1/19}-1 \approx 0.035$).

[72]From 1949 to 1968 we observe an overall growth rate of new patents of 127%. We obtain $0.044 \approx ((1+1.27)^{1/19} - 1$

[73]$79.5 = 3.5/4.4 \times 100$

[74]We compute the aggregate growth rate in each of the technologies and then average across technologies.

[75]The initial innovativeness of a location i refers to the amount of patents that inventors in i applied for in 1949-1953. We compute initial innovativeness by technology. Within a technology h, we rank all locations i according to their level of patents applied in 1949-1953 in technology h. Then we assign them to quartiles according to their position in the ranking. Quartile 0.00 is the lowest quartile of innovativeness while quartile 0.75 is the highest.

We find that the coefficients on lower quartiles of initial innovativeness are positive and statistically different from the coefficient in the highest quartile. Thus, knowledge access had a greater effect on establishments that are located in initially less innovative locations.[76] Given the difference in the coefficients, the increase in knowledge access predicts an average yearly growth of new patents of 4.5% for the initially lowest quartile of innovativeness, while it predicts 3.4% for the highest quartile.[77] The change in knowledge access predicts differential growth rate of 1.1 percentage points. In the data we observe an average yearly growth rate of patents of 7.2% in the lowest quartile and 1.9% in the highest quartile, implying a difference of 5.3 percentage points. Comparing the predicted and observed differential growth rates, the improvement in knowledge access as consequence of the reduction in travel time explains 21% of the difference in growth rates of new patents between locations in the lowest and highest quartile of innovativeness.[78]

In a similar way to in Section 6, we may be concerned that decisions of the regulator or airlines which affect travel time are endogenous to the diffusion of knowledge and consequently to knowledge access. Therefore, we construct an instrument for knowledge access in which instead of using observed travel time, we use the fictitious travel time in which routes were fixed to 1951 and travel

[76] A given change in knowledge access leads to a stronger increase in patenting in initially less innovative locations.

[77] The change in knowledge access for the lowest quartile is on average 9.1%, which multiplied by the coefficient 14.36 (obtained by doing 9.36+5.00=14.36) gives a predicted growth of 131% over 19 years. Translated into average yearly growth it is 4.5% = $[(1 + 1.31)^{(1/19)} - 1] \times 100$. For the highest quartile, knowledge access changed on average 9.5%, which multiplied by the coefficient 9.36 predicts 89% growth rate, which is 3.4% yearly growth rate.

[78] $21\% \approx 1.2/5.1 \times 100$

speed in each route is equal to the national average:

$$\widetilde{KA}_{iht} = \sum_{k} \omega_{hk} \sum_{j,j\neq i} \text{Patent stock}_{jk,t=1953} \times (\text{travel time}^{\text{fix network}}_{ijt})^{\beta} \tag{8}$$

We then implement the instrumental variables estimation by control function as in Section 6. The results are presented in columns (3) and (4) in Table 4. The coefficients do not show an important change and the convergence prediction obtained using non-instrumented PPML remains valid.[79]

Figure 17 shows on the left panel the patent growth observed in the data, while in the right panel it is the predicted patent growth. We compute the prediction using the observed change in travel time and quartile specific elasticities of column (2) in Table 4. Similarly to what is observed in the data, the change in travel time predicts a larger patenting growth rate in the South, the West, and some locations in the East like the Research Triangle (RaleighDurhamCary). At the same time, the change in travel time predicts smaller growth rates in New York, Chicago and their surroundings.

The result in column (2) implies that a given change in *Knowledge Access* had a stronger effect on patenting growth in less innovative locations. In other words, knowledge spillovers as an externality had a more predominant role in the production of knowledge in locations that initially produced relatively fewer patents. Theoretically, this result implies that the parameter ρ in equation (1) varies depending on the level of previous production of knowledge of location i. Empirically the implication is that a given increase in knowledge spillovers

[79]Using IV estimates, the predicted yearly patent growth rate in the lowest quartile is 4.9% while it is 3.7% in the highest quartile. The predicted differential growth rate is then 1.2 percentage points, meaning that the change in knowledge access can explain $(1.2/5.3) \times 100 \approx 23\%$ of the observed differential growth rate.

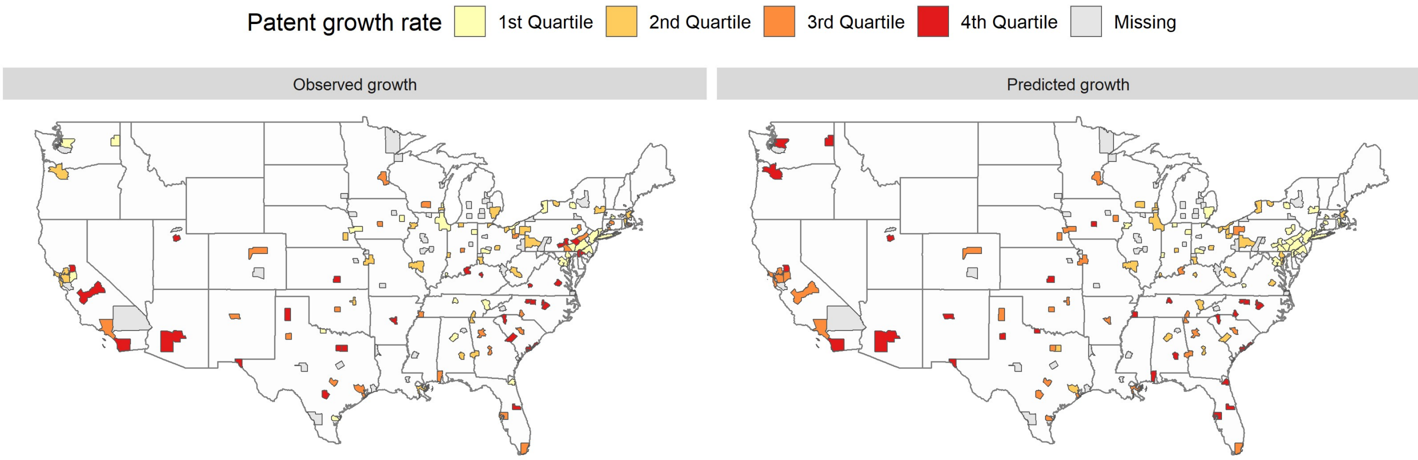

Figure 17: Observed vs. predicted patent growth 1951 - 1966

leads to innovation convergence across locations. In the data we observe that during 1949-1968 there is convergence across locations and that is exactly what the estimated coefficients predict following a reduction in travel time.

In order to understand the convergence result and compare it with other findings in the literature it is important to remember that commercial airplanes during 1950s and 1960s were a mean of transport mainly for people. On the other hand, other transportation improvements as those in water transport, railroads or highways also contain another ingredient: they were used to carry goods. The effects of these other means of transport have been widely studied and the results are different to the ones found in this paper. Pascali (2017) finds that the introduction of the steam engine vessels in the second half of the 19th century had an impact on international trade that led to economic divergence between countries. Faber (2014) finds that the expansion of the highway system in China led to a reduction of GDP growth in peripheral counties, with evidence suggesting a trade channel due to reduction in trade costs. In our setup, the introduction of jet airplanes represented a big shock to the mobility of people while not affecting significantly the transport of merchandise. As consequence, if any, the trade channel would be a second order effect and the improved face to face interactions would be the dominant effect.

7.1. Creation of knowledge: Robustness

In this section we show that the effect of *Knowledge Access* on the creation of new patents and the convergence effect remains with the inclusion of different controls. Table 28 shows the results.

Jaworski and Kitchens (2019) show that improvements in the Interstate Highway System led to local increases in income through an increased market access. In our set up, if the effect of market access affects innovation in the same way across technologies, then it would be absorbed by the MSA-time fixed effect FE_{it} in equation 7. However, if the effect of market access on innovation varies across technologies, then it would be a confounder. To control for this confounder, we compute market access by highway and interact it with a technology dummy. We compute market access as:

$$\text{Market Access}_{it} = \sum_j \text{Population}_{j,t=1950}\, \tau_{ijt}^{\theta} \tag{9}$$

where $\text{Population}_{j,t=1950}$ is population in MSA j in 1950, τ_{ijt} are the shipping costs provided in the data of Taylor Jaworski computed using each year's highway driving distance, highway travel time, petrol cost and truck driver's wage. θ is the elasticity of trade to trade costs which we set to -8.28, the preferred value of Eaton and Kortum (2002) and in the range of many other estimates in the literature (Head and Mayer (2014), Caliendo and Parro (2015), Donaldson and Hornbeck (2016)). Columns (3) and (4) of Table 28 show the results, we do not observe an important difference with the baseline estimates.

Campante and Yanagizawa-Drott (2017) shows that better connectivity by airplane leads to an increase in economic activity as measured by satellite-measured night light. Söderlund (2020) shows that an increase in business travel in the late 1980s and early 1990s led to an increase in trade between countries. In a similar way to market access, we could think that better connectivity by airplane could have led to an increase in market access due to a reduction in information frictions, with goods being shipped by land. Similarly to highway market access, if the effect of market access by airplane is common to all technology categories the effect would be absorbed by the MSA-time fixed effect FE_{it}. In order to account for a technology-specific effect, we construct a measure of airplane market access and interact it with a technology dummy. The measure of airplane market access is similar to equation 9 where τ is the travel time by airplane and θ is set to -1,22, the elasticity of trade to travel time from Söderlund (2020). The results are shown in columns (5) and (6) of Table 28. While the coefficients in all quartiles are reduced, the estimated value of ρ is positive and significant and the result on convergence remains.

We may think that better communication by other means of communication, like telephones, could have spurred the creation of new patents. In columns (7) and (8) we include the log of the MSA's share of households with telephones in 1960 and double-interact it with a technology dummy and a time dummy. The results remain invariant with respect to the baseline.

Another potential explanation for the increase of patenting could be that better connectivity decreased technology-specific financial frictions. However, during 1950s and 1960s interstate lending or bank branching was limited. Prior to 1970s, banks and holdings were restricted in their geographic expansion within and across state borders. Additionally, the Douglas Amendment to the Bank Holding Company Act prevented holding companies from acquiring banks in other states (Jayaratne and Strahan (1996)). Therefore, it is unlikely that interstate bank financing would be a driving force. Nonetheless, other sector-specific modes of financing like venture capital were active, it could be confounding the results. In Appendix **??** we construct multiple measures of access to capital by using market capitalization of patenting firms listed in the stock market. The results present suggestive evidence that access to capital is not driving the results.

Finally, in Appendix **??** we include a additional robustness checks. We compute different versions of *Knowledge Access*: using distance-specific β from section 6, considering patent stock only of locations j far from i, do sensitivity analysis using different values of β. Also, we compute baseline equation computing quartile of initial innovativeness using patents per capita. Last, we re-do the baseline regression using OLS estimation. Results go in the same direction: an increase in knowledge access leads to an increase in patenting and the effect is stronger in initially less innovative locations.

8. Firms' geographic expansion

In this section we track firms and establishments over time and space. We proceed in two steps. First, we show that the increase in patenting is driven by two types of entry: entry of establishments of new firms, and entry of establishments of pre-existing firms. The second type of entry is due to the geographic expansion of firms. Second, we show that the decrease in travel time led firms to expand geographically and this expansion was stronger towards initially less innovative locations. The two results are evidence that firms' geographic expansion was

	PPML		Highway Market Access		Airplane Market Access		Telephone	
Dependent Variable: *Patents*	$Patents_{Fiht}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(knowledge access)	10.14*** (3.66)	9.36** (3.69)	9.28** (3.68)	8.23** (3.69)	6.22* (3.58)	5.84 (3.60)	10.34*** (3.44)	9.25*** (3.43)
log(knowledge access) × 3rd quartile		2.05*** (0.58)		2.16*** (0.57)		2.06*** (0.59)		2.23*** (0.57)
log(knowledge access) × 2nd quartile		3.80*** (0.90)		3.89*** (0.89)		3.75*** (0.88)		3.93*** (0.91)
log(knowledge access) × 1st quartile		5.00*** (1.30)		5.13*** (1.30)		5.08*** (1.29)		5.18*** (1.32)
N obs. effective	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480
R2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85
Controls:								
log(Highway market access) × technology	-	-	Yes	Yes	-	-	-	-
log(Airplane market access) × technology	-	-	-	-	Yes	Yes	-	-
log(Telephone share) × technology × time	-	-	-	-	-	-	Yes	Yes

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 5: Elasticity of new patents to knowledge access, by MSA innovativeness quartile

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\rho\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed using patents in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Relative to columns (1) and (2), columns (3) and (4) control for technology specific effect of log(highway market access), columns (5) and (6) control for technology specific effect of log(airplane market access), columns (7) and (8) control for technology and time specific effect of log(telephone share). Standard errors clustered at the location-technology *ih* are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

one of the driving forces for convergence across locations.

8.1. Entry of new establishments

We exploit the ownership of patents to identify in which locations a firm had research establishments in each time period. We assume that firm F has a research establishment in the inventors' locations that appear in the patents filed by F. In this way, if the research establishment filed patents in a previous time period, we know that the research establishment existed previously. Also, using all patents applied for a firm we know the locations in which the firm has research establishments.[80] Using patents applied in our first time period (1949-1953), we classify all the research establishments that apply for patents in the all time periods (1949-1968). We classify research establishments into three mutually exclusive categories: neither the establishment nor the firm applied for patents in 1949-1953 (*new firm new est*), the establishment did not apply for patents but the firm had establishments in other locations applying for patents in 1949-1953 (*existing firm new est*), the establishment (and hence the firm) applied for patents in 1949-1953 (*existing firm and est*).[81] The dummies *new firm new est* and *existing firm new est* capture two types of entry margin. *new firm new est* captures a new establishment of a new firm, while *existing firm new est* captures entry due to the geographic expansion of firms. The dummy *existing firm and est* captures jointly an intensive and exit margin.

[80] A research establishment is a combination of a patent owner F and a location i where the patent owner has inventors applying for patents. We label a patent owner to be a *firm*. All our *firm* and *research establishment* information comes from the patent data. Hence, we only observe a firm in a certain time period if it applies for patents in that time period.

[81] We define if an establishment exists or not if it applied for patents in any technology h in 1949-1953. By defining an establishment at the Fi level (as opposite to Fih) we are sure that it if we observe entry it is the creation of a new establishment and not an existing establishment patenting in another technology.

We estimate a variation of equation (7) that includes interactions with dummies which indicate the status of the establishment in 1949-1953:

$$\text{Patents}_{Fiht} = \exp\left[\sum_{e} \boldsymbol{\rho_e}\, log(KA_{iht}) \times \mathbb{1}\{Fi \in e\} + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \nu_{Fiht} \tag{10}$$

where Patents_{Fiht} are patents applied by establishment of firm F in location i and technology h at time period t. KA_{iht} is the knowledge access at the location-technology-time level. $\mathbb{1}\{Fi \in e\}$ is an indicator variable that takes value 1 of Fi is of the type $e = \{$*new firm new est*, *existing firm new est*, *existing firm and est*$\}$. The results are displayed in column (2) of Table 6. The results show that the effect of innovation access on the increase of patenting came through the two entry margins: entry of new establishments of new firms and entry of new establishments of firms that previously existed in other locations.

In Table 7 we open up the effect, including a double interaction of Fi establishment type and the quartile of initial innovativeness of location i in technology h, using the highest quartile as the reference category. As in section 7, we classify the level of initial innovativeness of location i in technology h using the amount of patents applied in 1949-1953. The two margins of entry are active in all quartiles of innovativeness, with a stronger effect in lower quartiles. In the case of the entry of establishments that belong to firms that already existed in other locations, the pattern is more prominent. The intensive and exit margin does not appear active in any quartile of innovativeness except for the last one. The combined effect of entry and intensive/exit suggests that, in locations in the lowest quartile of initial innovativeness, the churn rate of patenting firms is increased as consequence of the increase in knowledge access.

Dependent Variable: *Patents*	$Patents_{Fiht}$ (1)	(2)
log(knowledge access)	10.14^{***} (3.66)	
log(knowledge access) × new firm new est		23.71^{***} (4.46)
log(knowledge access) × existing firm new est		23.79^{***} (4.47)
log(knowledge access) × existing firm and est		-0.28 (4.70)
R2	0.85	0.81
N obs. effective	991,480	991,480

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 6: Patents and knowledge access: Entry, exit and continuing firms

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \zeta_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) adds an interaction of $log(KA_{iht})$ with e the type of establishment Fi in a classification on whether the establishment and/or the firm existed in 1949-1953. Standard errors clustered at the location-technology ih are presented in parenthesis. R2 is computed as the squared correlation between observed and fitted values.

The results of Table 6 and Table 7 indicate that one part of the increase in patenting is consequence of multi-establishment firms that expand across locations, and more so in initially less innovative locations. Hence, multi-establishment firms contributed to innovation-convergence across locations by expanding geographically.

Quartile innovativeness \ Establishment type	New firm & New est	Existing firm & New est	Existing firm & Existing est
log(knowledge access)	22.84^{***} (4.40)	22.00^{***} (4.41)	-0.36 (4.67)
log(knowledge access) $\times$ 3rd quartile	3.40^{***} (1.14)	6.35^{***} (1.44)	-1.33 (1.19)
log(knowledge access) $\times$ 2nd quartile	5.95^{***} (1.48)	6.74^{***} (1.67)	-2.20 (2.33)
log(knowledge access) $\times$ 1st quartile	4.88^{**} (1.97)	10.98^{***} (2.15)	-15.62^{***} (3.25)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 7: Patents and knowledge access: entry, exit and continuing firms
The table shows the results of one Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho_{q4,e}}\, log(KA_{iht}) \times \mathbb{1}\{ih \in \text{4th quartile}\} \times \mathbb{1}\{Fi \in e\} + \sum_{q,e} \boldsymbol{\rho_{q,e}}\, log(KA_{iht}) \times \mathbb{1}\{ih \in q\} \times \mathbb{1}\{Fi \in e\} + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \nu_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{ikt} is knowledge access of establishments in location i technology h and time period t. q is the quartile of initial innovativeness of location i within technology h, computed using patents filed in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. e is the type of establishment Fi in a classification on whether the establishment and/or the firm existed in 1949-1953. Standard errors clustered at the location-technology ih are presented in parenthesis. R2 is computed as the squared correlation between observed and fitted values. All columns and rows belong to the same regression. The number of observations is 991,480.

8.2. Geographic expansion of multi-establishment firms

In this subsection we show that the decrease of travel time gave rise to the geographic expansion of multi-establishment firms. We focus on all firms that patented in the initial time period and follow their subsequent opening and closure of establishments. We find that firms directed the opening (closure) of new establishments towards locations that got stronger reductions in travel time to the firm's headquarters.

We define the headquarters location q of firm F as the location in which the firm filed the largest amount of patents in the period 1945-1953. If firm F did not file any patent in 1945-1953, or there is no unique location with the maximum amount of patents (e.g. two locations have the maximum amount of patents), then no headquarters is assigned.[82] Firms with no headquarters assigned are dropped from the estimations that required headquarters location. In Appendix B.4 we include a graphic representation of the geographical distribution of General Electric research establishments and its citations.

We compute the travel time of every firm F's headquarters's location q to each other location j. We then estimate a linear probability model to study if the location decision of establishments of a firm depend on travel time to a firm's headquarters. We estimate the following regression:

$$\mathbb{1}\{establishment_{Fqjt}\} = \gamma \, log(\text{travel time}_{qjt}) + FE_{Fqj} + FE_{Fqt} + FE_{jt} + \zeta_{Fqjt} \quad (11)$$

where $\mathbb{1}\{establishment_{Fqjt}\}$ is a dummy variable that takes value 1 if firm F with headquarters in location q has a research establishment in location j at time period t. The coefficient γ is a semi-elasticity: $\gamma/100$ is the change in percentage points of the probability that firm F has an establishment in location j when travel time increases by one percent. If travel time has a negative impact on the probability then we would expect γ to be negative. The inclusion of the fixed effect FE_{Fqj} implies that γ is identified only from changes in travel time and *opening* and *closure* of research establishments across time.[83] Fixed effects FE_{Fqt} and FE_{jt} control flexibly for changes in firm F expanding and opening establishments everywhere else, and j becoming more attractive for every firm.

Table 8 presents the results jointly with predicted and observed growth rate of

---

[82] Using patents applied in the period 1949-1953 does not significantly affect the results. We use 1945-1953 instead as it allows us to identify headquarters location for 7% more firms.

[83] $\mathbb{1}\{establishment_{Fqjt}\}$ takes value 0 if firm F does not file patents in location j at time period t. The headquarters location q remains fixed for all time periods.

the probability. Column (1) presents the results of estimating equation 11. We find that location j increases the probability of having a subsidiary establishment of firm F when the travel time between the firm's headquarters location q and j decreases. The coefficient is -0.0364, which if we multiply it by the average change in travel time between headquarters' location and every other potential location (-34.7%), the decrease in travel time predicts an increase in the share of existing subsidiaries of 0.0126 percentage points. In the data we observe that the increase in the share of existing subsidiaries is 0.00026 percentage points,[84] meaning that the change in travel time predicts an increase 46 times higher that what is observed.[85] While the magnitude of prediction is substantially different to what we observe in the data, we consider the finding of a negative and significant coefficient is reasonable in economic terms. The result goes in the same direction as Giroud (2013) who finds that a reduction in travel time between a firm's subsidiary and its headquarters leads to an increase in investment in the subsidiary.

In column (2) of Table 8 we estimate the semi-elasticity of the probability of having an establishment to travel time by the quartile of innovativeness of location j in 1949-1953. We compute the quartile of innovativeness at the location level by taking the average quantile across technologies within a location, only for those technologies in which the location has positive patents in 1949-1953. The semi-elasticity is 10 times bigger for the highest quartile of initial innovativeness relative to the lowest quartile, implying that 1% change in travel time has

[84]The share in 1949-1953 is 0.00081 and in 1964-1968 is 0.00107. The shares are computed as the amount of observed subsidiaries (3,967 in 1949-1953 and 5,262 in 1964-1968) divided by the amount of potential subsidiaries (4,893,217 which is time invariant). The amount of potential subsidiaries is the amount of firms (45,731 in the sample for which we identify HQ location) multiplied by the amount of locations other than HQ location (we have 108 locations in the data, meaning that each firm has 107 potential locations for subsidiaries).

[85]If we take the value in the extreme of the 95% confidence interval, $-0.01952 = -0.0364 + 0.88 \times 1.96$, we obtain a predicted change of 0.0066, 25 times higher than what is observed.

Dependent Variable:	Baseline	Quartile receiving location	Initial probability	Change travel time	Predicted yearly growth rate	Observed yearly growth rate
	$\mathbb{1}\{establishment_{Fqjt}\}$					
	(1)	(2)				
log(travel time)	-0.0364*** (0.0088)		0.000810	-34.7%	15.94%	1.50%
log(travel time) × 4th quartile		-0.0749*** (0.0187)	0.001895	-36.0%	15.41%	0.98 %
log(travel time) × 3rd quartile		-0.0150*** (0.0031)	0.000364	-33.4%	15.22%	3.03%
log(travel time) × 2nd quartile		-0.0102*** (0.0028)	0.000145	-35.2%	18.67%	3.86%
log(travel time) × 1st quartile		-0.0079*** (0.0025)	0.000068	-33.8%	21.40%	5.75%
R2	0.49	0.50				
N obs. effective	19,755,792	19,755,792				

*** $p < 0.01$; ** $p < 0.05$; * $p < 0.10$

Table 8: Subsidiaries' location and travel time to headquarters

The table shows the estimation of a linear probability model. The left panel of the table shows estimation results while the right panel shows observed and predicted growth rates of the probability. Column (1) presents the results of OLS estimation of $\mathbb{1}\{establishment_{Fqjt}\} = \gamma\, log(\text{travel time}_{qjt}) + FE_{Fqj} + FE_{Fqt} + FE_{jt} + \zeta_{Fqjt}$ or firm F which has headquarters in location $q$where $\mathbb{1}\{establishment_{Fqjt}\}$ is a dummy that takes value one if firm F which has headquarters in location q has an establishment *open* in location j at time period t. We define an establishment of firm F in location j at time period t as *open* if F has inventors located in j that apply for patents at time period t. travel time$_{qjt}$ is the travel time in minutes between F's headquarters location q and location j at time period t. Column (2) includes an interaction of $log(\text{travel time}_{qjt})$ with the across-technology average quartile of initial level of innovativeness of j. j's quantile of initial innovativeness in technology h is computed using the level of patents of j in 1949-1953 in technology h. Standard errors at the non-directional location pair are presented in parentheses (qj is the same non-directional location pair as jq). Predicted growth rates are obtained using the estimated coefficient and the change in travel time, relative to the initial probability. Yearly growth rates g are obtained by computing $g = [(1 + \text{nineteen year growth rate})^{(1/19)} - 1] \times 100$, where 19 is the amount of years between 1949 and 1968.

10 times a bigger impact in terms of percentage points change in the probability. However, as these are semi-elasticities, in order to compare across quartiles the impact of travel time on the relative change in the probability we have to take into account the initial probability of each quartile. While the semi-elasticity in the highest quartile is around 10 times the one of the lowest quartile, the initial probability is around 27 times the one of the lowest quartile. That implies that, for a given percentage change in travel time, the impact of the change in travel time on the growth rate of the probability in the highest quartile is around 1/3 the one in the lowest quartile.[86] In other words, given the initial very low probability of locations in the lowest quartile of innovativeness to receive a subsidiary from a firm headquartered in another location, the small increase in percentage points represents a big relative increase in the probability.

The yearly growth rate implied by the change in travel time is 21.4% for the lowest quartile while it is 15.4% for the highest quartile, implying a predicted difference of 6 percentage points in the yearly growth rate.[87] In the data we observe an average yearly growth rate which is 4.8 percentage points higher for the lowest quartile relative to the highest quartile.[88] Therefore, using only

[86] These are approximate numbers. The precise computations: the ratio of coefficients is $9.48 = (-0.0749)/(-0.0079)$, the ratio of initial probability is $27.86 = 0.001895/0.000068$, the ratio of the growth rate is $0.34 = (-0.0749/0.001895)/(-0.0079/0.000068)$.

[87] For the lowest quartile, the model predicts a 3,869% increase in the probability over 19 years ($19 = 1968 - 1949$), which translates into an average yearly growth rate of 21.4%. For the highest quartile the predicted increase is 1,422%, an average yearly growth rate of 15.4%. Consistent with the computation of the relative growth rate: $1,422/3,869 = 0.36 \approx 0.34 \times (33.8/36.0)$, where 0.34 has to be adjusted by the fact that the average change in travel time is not the same across quartiles. The 19-year growth rates are obtained by multiplying the change in travel time (-33.8% vs -36.0%) by the coefficient (-0.0079 vs -0.0749) divided by 100, and finally dividing by the initial probability (0.000069 vs 0.001895) and multiplying by 100. For the lowest quartile: $3,869 = [(-33.8) \times (-0.0079/100)/0.000069] \times 100$, and for the highest quartile:$[1,422 = (-36.0) \times (-0.0749/100)/0.001895] \times 100$. The average yearly growth rates are computed as $21.4 \approx [(1+38.69)^{1/19} - 1] \times 100$ and $15.4 \approx [(1+14.22)^{1/19} - 1] \times 100$.

[88] The average yearly growth rate of the probability for the lowest quartile is 5.7% while it is 0.9% for the highest quartile.

variation in travel time, the model correctly predicts a stronger growth rate of entry of multi-establishment firms in initially less innovative locations.

9. Conclusion

This paper constructed a new dataset of the flight network in the United States during the *Jet Age* and studied the impact of improvements of air travel on the diffusion and creation of knowledge. We found that the reduction in travel time lead to a increase in knowledge diffusion, especially between research establishment located far apart. The reduction in travel time also led to an increase in the general access to knowledge, which impacted positively the creation of knowledge. The increase effect in the creation of knowledge was stronger in locations initially less innovative, generating a convergence force which goes in the same direction as what is observed in the data. One of the drivers of the increase the creation of knowledge and convergence is the geographical expansion of firms.

We estimate one new key parameter: the elasticity of diffusion of knowledge to travel time. We provide causal evidence of *standing on the shoulders of giants*: how new knowledge builds upon pre-existing knowledge. We do so by using an economy-wide shock and focusing on the private sector. The results show that knowledge spillovers are more important for locations which are initially less innovative.

Our novel dataset document a historical country wide event that dramatically changed the way we see time and space. Our results provide new evidence on how the introduction of jet airplanes changed the geography of innovation.

Better connectivity to innovation centers in the Northeast led to an increase in innovation in the South and the West of the United States. In this way, jet airplanes were one facilitator in the shift of innovative activity towards the South and the West of the United States.

In future research we plan to a structural model that incorporates the findings of this paper. We have consider two types of model that could potentially account for the increase in the diffusion of knowledge and the increase of innovation in the South and the West. The first option is to modify Donaldson and Hornbeck (2016) including an intermediate sector which produces knowledge. The output of this intermediate sector would depend on *Knowledge Access*: either as an externality or as a composite of knowledge that enters as an input in the production function of knowledge. Depending of the modeling choice, travel time would affect the level of the externality or affect the price index of the composite of knowledge. The second option is to modify Davis and Dingel (2019), who find that a system of cities is an equilibrium outcome in the presence of localized knowledge spillovers. However, Davis and Dingel (2019) only allows for knowledge spillovers within cities. A natural extension would be to allow for knowledge spillovers across cities, where the degree of across-city spillovers depends on the across-city travel time.

The previous two models, however, do not include multi-establishment firms. A model with multi-establishment firms should include the location interdependency of establishments within a firm: the ideal location of an establishment of a firm depends on the location of every other establishment of the firm. To do so we would build upon Oberfield et al. (2020) who present a model of spatial equilibrium with multi-establishment firms.

Finally, we would like to point to the limitations of our analysis. The results found in this paper are identified by exploiting differential time changes across establishments. As consequence, we are able to identify differential impacts and not aggregate ones. The results obtained could be consequence of general increase in the amount of diffusion and creation of knowledge, a relocation of previous diffusion and creation, or a mix of both. At the same time, the potential relocation of resources as consequence of the reduction in travel time may have increased the allocative efficiency and therefore increasing the amount of knowledge creation. Further research should be performed in order to be able to disentangle the two forces and their interactions.

References

Acemoglu, D., U. Akcigit, and W. R. Kerr (2016). Innovation network. *Proceedings of the National Academy of Sciences 113*(41), 11483–11488.

Aghion, P. and P. Howitt (1997). *Endogenous Growth Theory*. The MIT Press.

Andrews, M. J. and A. Whalley (2021). 150 years of the geography of innovation. *Regional Science and Urban Economics*, 103627.

Audretsch, D. B. and M. P. Feldman (2004). Knowledge spillovers and the geography of innovation. In *Handbook of regional and urban economics*, Volume 4, pp. 2713–2739. Elsevier.

Balsmeier, B., M. Assaf, T. Chesebro, G. Fierro, K. Johnson, S. Johnson, G.-C. Li, S. Lck, D. O'Reagan, B. Yeh, G. Zang, and L. Fleming (2018). Machine learning and natural language processing on the patent corpus: Data, tools, and new measures. *Journal of Economics & Management Strategy 27*(3), 535–553.

Baum-Snow, N. (2007). Did highways cause suburbanization? *The quarterly journal of economics 122*(2), 775–805.

Bergé, L. (2018). Efficient estimation of maximum likelihood models with multiple fixed-effects: the R package FENmlm. *CREA Discussion Papers* (13).

Borenstein, S. and N. L. Rose (2014, June). *How Airline Markets Work... or Do They? Regulatory Reform in the Airline Industry*, pp. 63–135. University of Chicago Press.

C.A.B. (1951, 1956, 1961, 1966). *Air carrier traffic statistics.* Civil Aeronautics Board.

Caliendo, L. and F. Parro (2015). Estimates of the trade and welfare effects of nafta. *The Review of Economic Studies 82*(1), 1–44.

Campante, F. and D. Yanagizawa-Drott (2017, 12). Long-Range Growth: Economic Development in the Global Network of Air Links*. *The Quarterly Journal of Economics 133*(3), 1395–1458.

Carlino, G. and W. R. Kerr (2015). Agglomeration and innovation. *Handbook of regional and urban economics 5*, 349–404.

Catalini, C., C. Fons-Rosen, and P. Gaul (2018, June). How do travel costs shape collaboration? Working Paper 24780, National Bureau of Economic Research.

Caves, R. E. (1962). *Air transport and its regulators: an industry study*. Harvard economic studies v. 120. Cambridge: Harvard University Press.

Coscia, M., F. Neffke, and R. Hausmann (2020). Knowledge diffusion in the network of international business travel. *Nature Human Behaviour 4*(10).

Davis, D. R. and J. I. Dingel (2019). A spatial knowledge economy. *American Economic Review 109*(1), 153–70.

De Rassenfosse, G. and A. B. Jaffe (2017). Econometric evidence on the r&d depreciation rate. Technical report, National Bureau of Economic Research.

Dhaene, G. and K. Jochmans (2015). Split-panel jackknife estimation of fixed-effect models. *The Review of Economic Studies 82*(3), 991–1030.

Dijkstra, E. W. et al. (1959). A note on two problems in connexion with graphs. *Numerische mathematik 1*(1), 269–271.

Donaldson, D. and R. Hornbeck (2016). Railroads and american economic growth: A market access approach. *The Quarterly Journal of Economics 131*(2), 799–858.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

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.

Furman, J. L. and S. Stern (2011, August). Climbing atop the shoulders of giants: The impact of institutions on cumulative research. *American Economic Review 101*(5), 1933–63.

Giroud, X. (2013, 03). Proximity and Investment: Evidence from Plant-Level Data *. *The Quarterly Journal of Economics 128*(2), 861–915.

Glaeser, E. (2011). *Triumph of the City*. Macmillan.

Gross, D. P. (2019). The consequences of invention secrecy: Evidence from the uspto patent secrecy program in world war ii. Technical report, National Bureau of Economic Research.

Haines, M. R. et al. (2010). Historical, demographic, economic, and social data: the united states, 1790–2002. *Ann Arbor, MI: Inter-university Consortium for Political and Social Research*.

Hall, B. H., A. B. Jaffe, and M. Trajtenberg (2001, October). The nber patent citation data file: Lessons, insights and methodological tools. Working Paper 8498, National Bureau of Economic Research.

Head, K. and T. Mayer (2014). Chapter 3 - gravity equations: Workhorse,toolkit, and cookbook. In G. Gopinath, E. Helpman, and K. Rogoff (Eds.), *Handbook of International Economics*, Volume 4 of *Handbook of International Economics*, pp. 131–195. Elsevier.

Herzog, I. (2021). National transportation networks, market access, and regional economic growth. *Journal of Urban Economics 122*, 103316.

Hovhannisyan, N. and W. Keller (2015, March). International business travel: an engine of innovation? *Journal of Economic Growth 20*(1), 75–104.

I.C.C. (1965). *79th Annual Report*. Interstate Commerce Commission.

I.C.C. (1967). *81st Annual Report*. Interstate Commerce Commission.

Jaffe, A. B., M. Trajtenberg, and M. S. Fogarty (2000, May). Knowledge spillovers and patent citations: Evidence from a survey of inventors. *American Economic Review 90*(2), 215–218.

Jaffe, A. B., M. Trajtenberg, and R. Henderson (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics 108*(3), 577–598.

Jaworski, T. and C. T. Kitchens (2019). National policy for regional development: Historical evidence from appalachian highways. *Review of Economics and Statistics 101*(5), 777–790.

Jayaratne, J. and P. E. Strahan (1996). The finance-growth nexus: Evidence from bank branch deregulation. *The Quarterly Journal of Economics 111*(3), 639–670.

Jones, C. I. (2002, March). Sources of u.s. economic growth in a world of ideas. *American Economic Review 92*(1), 220–239.

Kogan, L., D. Papanikolaou, A. Seru, and N. Stoffman (2017, 03). Technological Innovation, Resource Allocation, and Growth*. *The Quarterly Journal of Economics 132*(2), 665–712.

Krugman, P. (1992, September). *Geography and Trade*, Volume 1 of *MIT Press Books*. The MIT Press.

Krugman, P. R. (1991). *Geography and trade*. MIT press.

Lucas, R. E. (1993). Making a miracle. *Econometrica 61*(2), 251–272.

Magerman, T., B. Van Looy, and X. Song (2006). Data production methods for harmonized patent statistics : patentee name harmonization.

Manson, S., J. Schroeder, D. Van Riper, T. Kugler, and S. Ruggles (2020). Ipums national historical geographic information system: Version 15.0.

Marshall, A. (1920). Principles of economics 8th ed. *London: McMillan*.

Michaels, G. (2008). The effect of trade on the demand for skill: Evidence from the interstate highway system. *The Review of Economics and Statistics 90*(4), 683–701.

Oberfield, E., E. Rossi-Hansberg, P.-D. Sarte, and N. Trachter (2020). Plants in space. Technical report, National Bureau of Economic Research.

Pascali, L. (2017). The wind of change: Maritime technology, trade, and economic development. *American Economic Review 107*(9), 2821–54.

Petralia, S. (2019). HistPat International Dataset.

Petralia, S., P.-A. Balland, and D. Rigby (2016). HistPat Dataset.

Redding, S. and A. J. Venables (2004). Economic geography and international inequality. *Journal of international Economics 62*(1), 53–82.

Silva, J. M. C. S. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Söderlund, B. (2020). The importance of business travel for trade: Evidence from the liberalization of the soviet airspace. Working Paper Series 1355, Research Institute of Industrial Economics.

Storper, M. and A. J. Venables (2004, 08). Buzz: face-to-face contact and the urban economy. *Journal of Economic Geography 4*(4), 351–370.

Weidner, M. and T. Zylkin (2021). Bias and consistency in three-way gravity models. *Journal of International Economics 132*, 103513.

Wooldridge, J. M. (2014). Quasi-maximum likelihood estimation and testing for nonlinear models with endogenous explanatory variables. *Journal of Econometrics 182*(1), 226–234.

A. Appendix: Travel Time Data

A.1. Data Construction

We construct a dataset of travel times by plane between US MSAs for the years 1951, 1956, 1961, 1966. We get information of direct flights from airline flight schedules and feed this information into an algorithm to allow for indirect flights. For each MSA pair with airports served by at least one of the airlines in our dataset we compute the fastest travel time in each of the four years.

Using images of flight schedules, we digitized the flight network for six major airlines: American Airlines (AA), Eastern Air Lines (EA), Trans World Airlines (TWA), United Airlines (UA), Braniff International Airways (BN) and Northwest Airlines (NW). Note that the first four in this list were often referred to as the *Big Four*, highlighting their dominant position in the market. They alone accounted for 74% of domestic trunk revenue passenger-miles from February 1955 to January 1956. Together the six airlines accounted for 82% of revenue passenger-miles in that same period, 77% from February 1960 to January 1961 and 78% from February 1965 to January 1966 (C.A.B., 1966). Our sample of airlines thus covers a vast share of the domestic market for air transport. In addition, the airlines were chosen to maximize geographic coverage.

In total we obtain a sample of 5,910 flights. These flights often have multiple stops. If we count each origin-destination pair of these flights separately, our sample contains 17,469 legs.

Table 9 lists the exact dates of when flight schedules we digitized became effective. Due to limited data availability not all flight schedules are drawn from the same part of the year. As seasonality of the network seems limited and given the large market share of the airlines we consider, our data is a good approximation of the network in a given year.

Table 9: Date of Digitized Flight Schedules

Airline	1951	1956	1961	1966
AA	September 30	April 29	April 30	April 24
EA	August 1	October 28	April 1	April 24
TWA	August 1	September 1	April 30	May 23
UA	April 29	July 1	June 1	April 24
BN	August	August 15	April 30	April 24
NW	April 29	April 29	May 28	March 1
PA	June 1	July 1	August 1	August 1

Figure 18 shows two pages of the flight schedule published by American Airlines in 1961. Each column corresponds to one flight. As can be seen, one flight often has multiple stops. Departure and arrival times in most flight schedules are indicated using the 12-hour system. PM times can be distinguished from AM times by their bold print. In the process of digitization we converted the flight schedules to the 24-hour system. Times in most tables are in local time. We thus recorded the time zones that are indicated next to the city name and converted them to Eastern Standard Time.

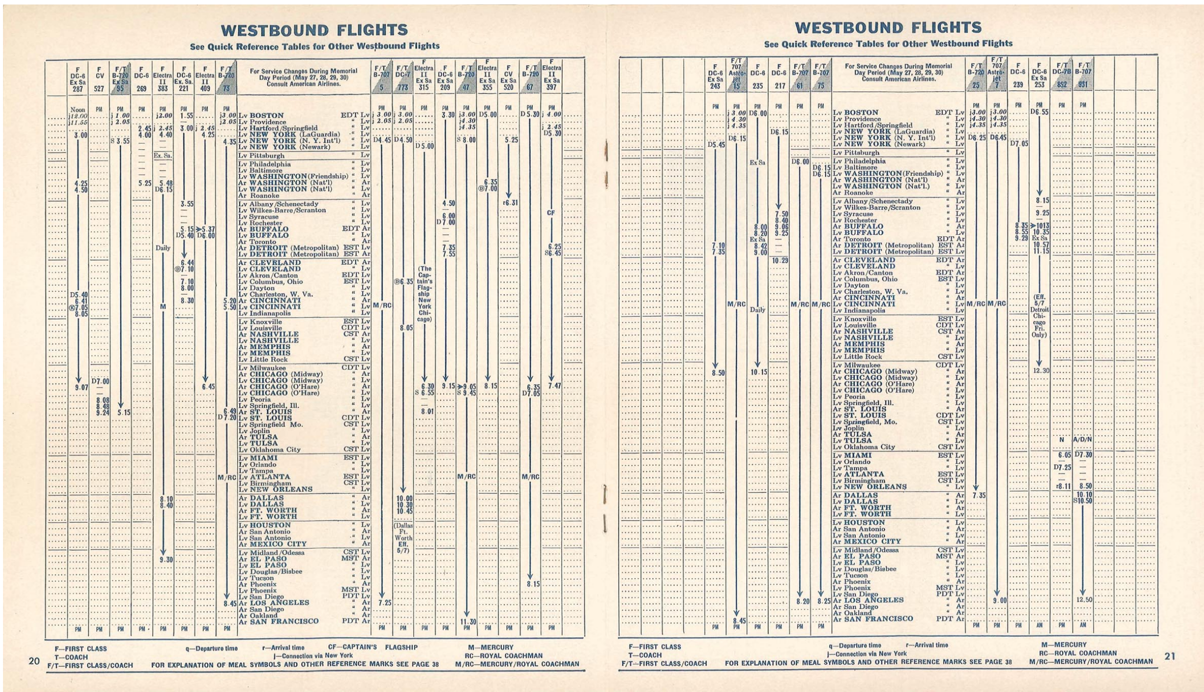

Figure 18: Flight Schedule American Airlines 1961.

To obtain exact geographical information on where airports are located, we match city names to their IATA airport codes. We use the addresses of ticket offices that are indicated on the last pages of the flight schedules. Most of the ticket offices were located directly at the airport, allowing to infer the airport the airline was serving in a given year. For some flight schedules we are missing these last pages and used information from adjacent years in order to identify airports. We also manually verified the airport match using various online sources. We then obtain geographical coordinates from a dataset provided by `https://ourairports.com/` (downloaded July 2020).

From the flight schedule we also collect information on the aircraft model, indicated next to the flight number. Using various online sources, we manually identified aircraft models that are powered by a jet engine. We thus know on which connections airlines were using jet aircraft.

Flight Schedules also contain information on connecting flights. For example, the second column in figure 18 indicates a departure from Boston leaving at 12.00 local time. A footnote is added to the departure time indicating that this departure is a connection via New York. It is thus not operated by flight 287 otherwise described in column 2, but it is just supplementary information for the passenger. As we are interested in the speed of aircraft and the actual travel time on a given link, this information on connecting flights would pollute our data and we thus delete this supplementary information.

As outlined above, the digitization requires human input. It is thus prone error-prone. The travel time calculation relies on each link in the network, and if one important connection has a miscoded flight, it might potentially distort the travel time between many MSA pairs. We thus implement an elaborate method to detect mistakes in the digitization process. In particular, after the initial transcription, we regress the observed duration of the flight on a set of explanatory variables: the full interaction of distance, a set of airline indicators, a set of year indicators and a dummy variable indicating whether the aircraft is powered by a jet engine or not. This linear model yields an R^2 above 95%. We then compute the predicted duration of each flight and obtain the relative deviation from the observed duration. If the deviation is above 50%, we manually check whether the transcribed information is correct. If we find a mistake, we correct the raw data, rerun the regression and recompute relative deviations, until all the observations with more than 50% deviation have been manually verified.

For 15 connections, the information was correctly transcribed from the flight schedule, but the flight time differed a lot from other flights with similar distances that used the same aircraft. The implied aircraft speed for these cases is either unrealistically high or low, in one case the implied flight time is even negative. These cases seem to be typos introduced when the flight schedule was created (e.g. a ”2” becomes a ”3”). Instead of inferring what the true flight schedule was which is not always obvious, we drop these cases. Table 10 lists all 15 cases.

Table 10: Dropped Connections

	Airline	Year	Origin	Destination	Departure Time	Arrival Time
0	UA	66	TYS	DCA	1940	2036
1	UA	66	LAX	BWI	2150	1715
2	UA	66	CHA	TYS	1635	1909
3	PA	66	SFO	LAX	2105	1850
4	PA	66	SEA	PDX	705	935
5	PA	56	PAP	SDQ	830	835
6	PA	51	HAV	MIA	800	903
7	PA	51	SJU	SDQ	825	830
8	NW	66	HND	OKA	655	1135
9	EA	66	ORD	MSP	2340	2340
10	EA	56	SDF	MDW	1352	1418
11	EA	56	GSO	RIC	2207	2204
12	AA	56	PHX	TUS	1630	1655
13	PA	51	STR	FRA	1320	1540
14	EA	66	TPA	JFK	1330	1548

As our analysis is at the MSA level, we match airports to 1950 MSA boundaries. Each airport is matched to all MSAs for which it lies inside the MSA boundary or at most 15km away from the MSA boundary. If we focus only on airports contained within MSA boundaries, we would, for example, drop Atlanta’s airport.

Of 275 US airports, 156 airports are matched to at least one MSA. 18 of these are matched to two MSAs and Harrisburg International Airport is matched to three MSAs: Harrisburg, Lancaster and York. Out of 168 MSAs, 142 are at some point connected to the flight network in our dataset. In table **??** we present the 168 MSAs, the ones that are connected at least once, and the ones that are connected in the four years.

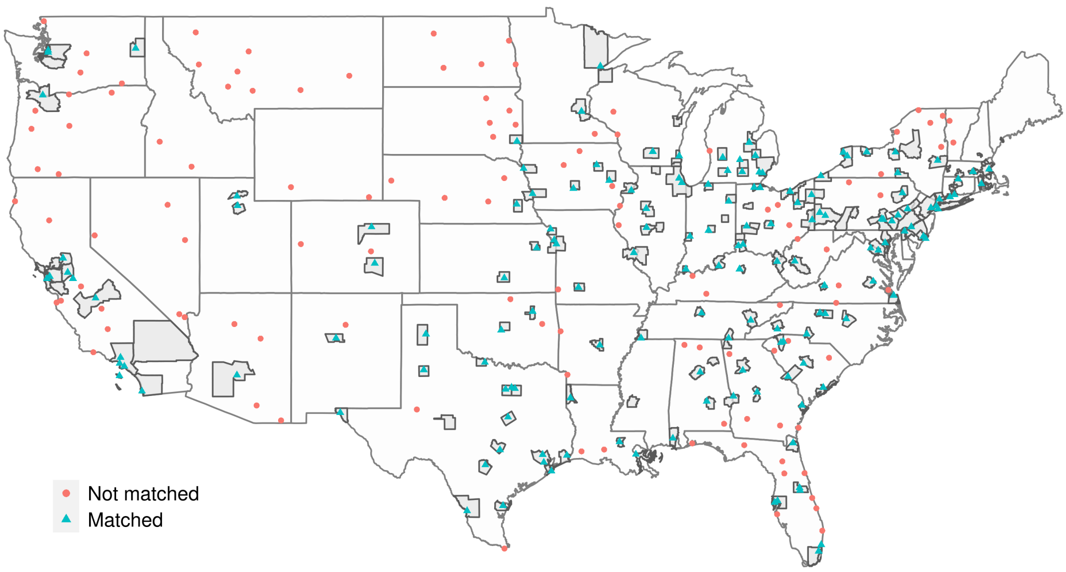

Figure 19: Airports matched to MSAs.

Next, we compute the shortest travel time for every airport pair, and then take the minimum to obtain shortest travel time at the MSA pair level. In particular, we apply Dijkstra's algorithm to compute shortest paths (Dijkstra et al., 1959). We adjust this algorithm to take into account the exact timing of the flight schedules. We consider a possible departure time t from origin city o and then compute the shortest path to destination city d at this time of the day. If getting to d requires switching flights, we account for the required time at the location of the layover. We repeat this procedure for every possible departure time t at origin city o and then take the minimum that gives us the fastest travel time from o to d, τ_{od}.

The flight schedule format requires us to make one assumption. In particular, the flight schedule for a multi-stop flight may either indicate the arrival time or the departure time for a particular stop. If the flight schedule only lists the departure time, we need to infer the arrival time and vice versa. We allow for five minutes between arrival and departure. This is relatively low, but still in the range of observed difference between departure and arrival for cases where we observe both. As correspondences may have been ensured by airlines in reality, i.e. one aircraft waiting with departure until other aircraft arrive, we opted for the lower end of the observed range of stopping times.

Finally, since the shortest travel time measure may not capture the benefits of a highly frequented hub, we also calculate the daily average of the shortest travel time. In particular, we compute the shortest travel time at every full hour of the day and take the average. This measure thus captures the benefits of being located near an airport where flights depart many times per day.

To conclude, we end up with a set of four origin-destination matrices indicating the fastest travel time (and another set with the average daily travel time) between US MSAs in 1951, 1956, 1961 and 1966.

A.2. Descriptive Statistics

Table 12 shows the number of non-stop connections between MSAs by year and airline. It underlines the dominant position of the *Big Four* (AA, EA, TW, UA) which were much bigger than their competitors (BN and NW). The growth of the airline industry is also apparent. All airlines had the lowest number of connections in 1951 and subsequently extended their network. At the same time the average distance of the connections gradually increased over time. Part of this may have been due to jet technology allowing for longer aircraft range. We thus analyze a period where more and longer flights are introduced.

Table 12: Domestic Non-Stop Connections by Airline and Year

Airline	Year	Number of connections	Jet Share (connections)	Jet Share (km)	Mean Distance (in km)
AA	1951	258	0.00	0.00	515.32
AA	1956	367	0.00	0.00	889.66
AA	1961	325	22.15	50.50	768.24
AA	1966	282	73.40	89.52	1020.36
BN	1951	96	0.00	0.00	317.90
BN	1956	210	0.00	0.00	380.60
BN	1961	176	8.52	18.84	460.41
BN	1966	150	72.00	76.64	553.09
EA	1951	345	0.00	0.00	319.87
EA	1956	479	0.00	0.00	412.60
EA	1961	595	3.70	13.28	441.42
EA	1966	492	54.47	75.46	569.01
NW	1951	77	0.00	0.00	521.70
NW	1956	95	0.00	0.00	724.77
NW	1961	127	11.02	32.43	824.59
NW	1966	136	77.94	90.86	945.81
TW	1951	210	0.00	0.00	503.69
TW	1956	253	0.00	0.00	711.78
TW	1961	240	28.75	54.63	807.72
TW	1966	265	86.42	96.05	1143.30
UA	1951	291	0.00	0.00	492.88
UA	1956	361	0.00	0.00	714.39
UA	1961	323	31.89	65.32	803.49
UA	1966	533	49.91	79.54	781.38

While these changes in the network are remarkable, airlines were constrained by the regulator in opening new routes. Accordingly, table 13 shows that the network remains relatively stable over time with more than three quarters of connections remaining intact within a five-year window. Interestingly, during the beginning of the jet age (i.e. 1956 to 1961), the network appears to have been especially stable, with only 11% of connections either disappearing or newly being added. Thus, the rise of jet aircraft did not lead to a vast reshaping of the network. Given the very different technology, this may be surprising, but may partly be due to heavy regulation.

The table also shows that newly introduced routes were over long distances whereas those discontinued were operating on shorter distances. When changes in the network took place, they thus seemed to improve the network for places further apart.

Table 13: Network Changes (weighted by frequency)

Period	Remain connected	Newly connected	Disconnected
Share of Non-stop Connections (%)			
1951 to 1956	78.47	16.79	4.74
1956 to 1961	88.96	6.43	4.6
1961 to 1966	80.64	12.37	6.99
Mean distance (km)			
1951 to 1956	411	1075	337
1956 to 1961	524	914	972
1961 to 1966	568	769	450

Table 14: Network Changes

Period	Remain connected	Newly connected	Disconnected
Connected MSAs			
1951 to 1956	119	7	8
1956 to 1961	122	0	4
1961 to 1966	114	7	8
Non-stop Connections			
1951 to 1956	721	357	124
1956 to 1961	908	231	170
1961 to 1966	912	331	227

Changes in the number of connected MSAs and connections among them. A MSA is connected if in our data it appears as having at least one incoming and one outgoing flight. A non-stop connection refers to a pair of origin MSA-destination MSA between which a non-stop flight operates.

Figure 20 shows all non-stop connections in our data weighted by the (log) frequency. Initially, the network was concentrated in the Eastern states and transcontinental routes were not yet established, due to technological limitations. In contrast, in the 1960s, after the jet is introduced, intercontinental routes quickly emerge and are operated at a high frequency. Similarly, direct connections from the Northeast to Florida intensify. The figure echos the findings from table 14 which illustrates that the overall number of MSA pairs with a direct connection increases over time.

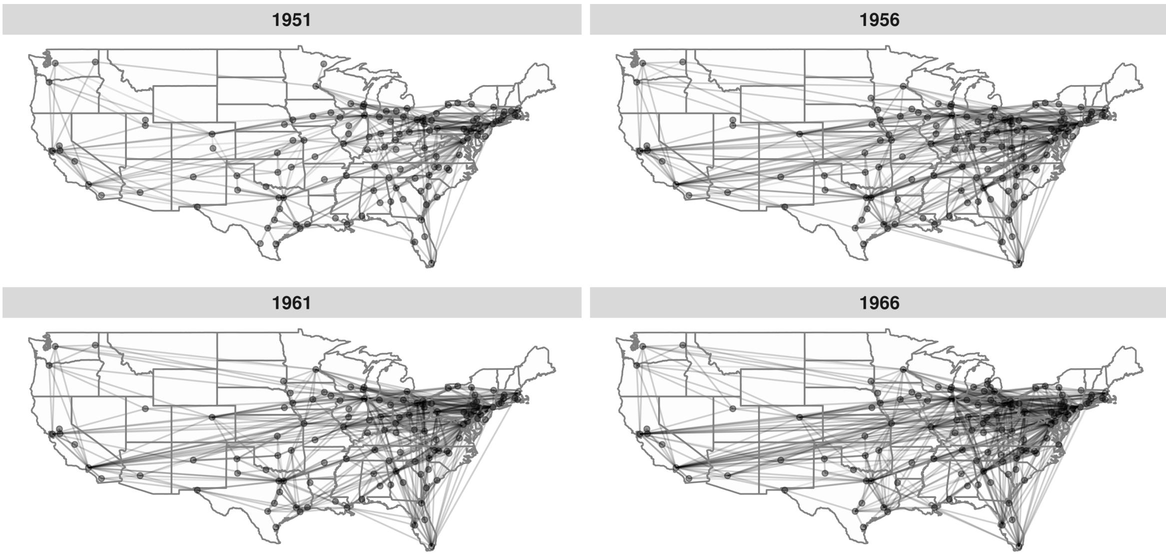

Figure 20: Flight Network by Year. Weighted by log weekly frequency.

Airlines differed in their speed of adoption of the newly arrived jet aircraft. Table 12 shows that, in 1961, 65% of UA's connections between MSAs were flown using a jet aircraft (weighted by distance), whereas this was only true for 13% of EA's connections. While adoption was heterogeneous across airlines, adoption was fast. By 1966, all airlines were operating 75% of their connections with jet aircraft (weighted by distance).

Figure 21 show the average speed of jet and propeller aircraft by distance. Generally, jet aircraft were substantially faster, but especially so on long-distance flights, where they could be up to twice as fast as propeller-driven aircraft. This particularly stark difference in speed for long-haul flights is also reflected by adoption. Figure 22 shows that jet aircraft were first introduced on long-haul flights. Only 50% of MSA pairs at around 1,500 km distance had at least one jet aircraft operating, whereas 100% of pairs above 3,000 km. Then, in the late 1960s, they were also gradually introduced on shorter distances. In fact, for all pairs above 2,000 km there was at least one jet engine-powered flight.

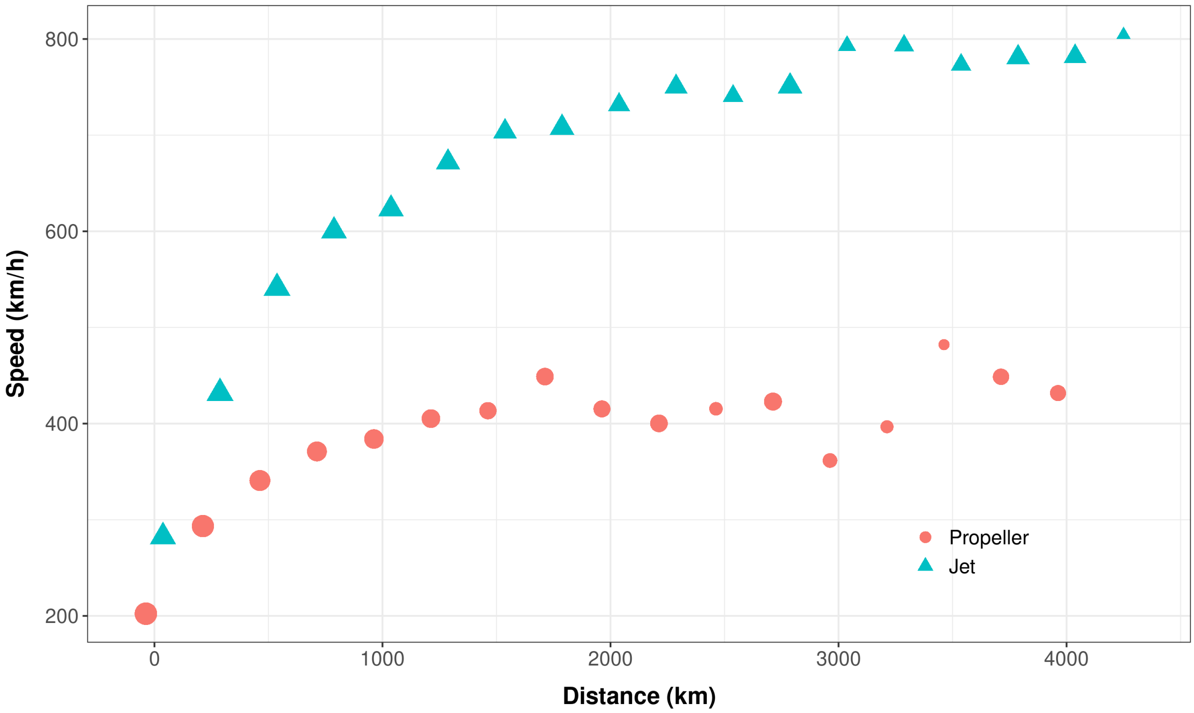

Figure 21: Speed by Aircraft Type. Pooling all Years.

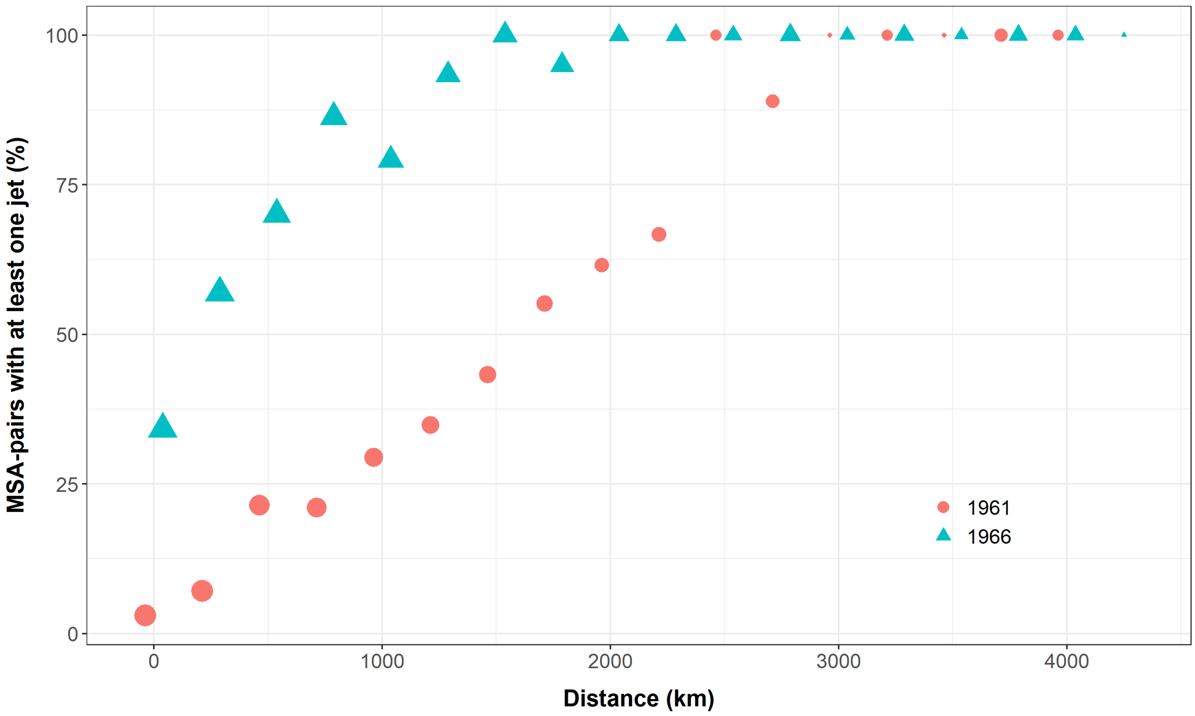

Figure 22: Jet Adoption.

Figure 23 shows on which routes jets were operating. In the early days of the jet age it was mainly the transcontinental corridor between New York and California that benefited. In 1966 propeller aircraft were already being phased out and only operating in the dense Eastern part of the US where distances between cities are relatively small.

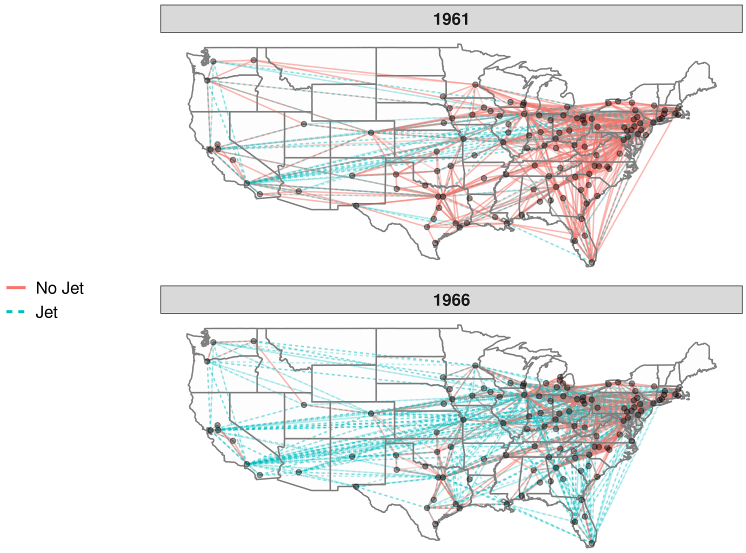

Figure 23: Jet Adoption by Year.

The increase in speed due to jet aircraft caused a dramatic reduction in travel times between US cities. When looking at the full origin-destination matrix, i.e. including indirect flights, a network-wide reduction in travel time becomes apparent. Figure 24 shows travel times between US MSAs. While the figure shows a gradual decline in travel time from 1951 to 1966, it also illustrates that conditional on distance and year a large amount of variation in travel time remains, as only a small fraction of all MSA pairs were connected via a direct flight (around 8.5% in 1966).

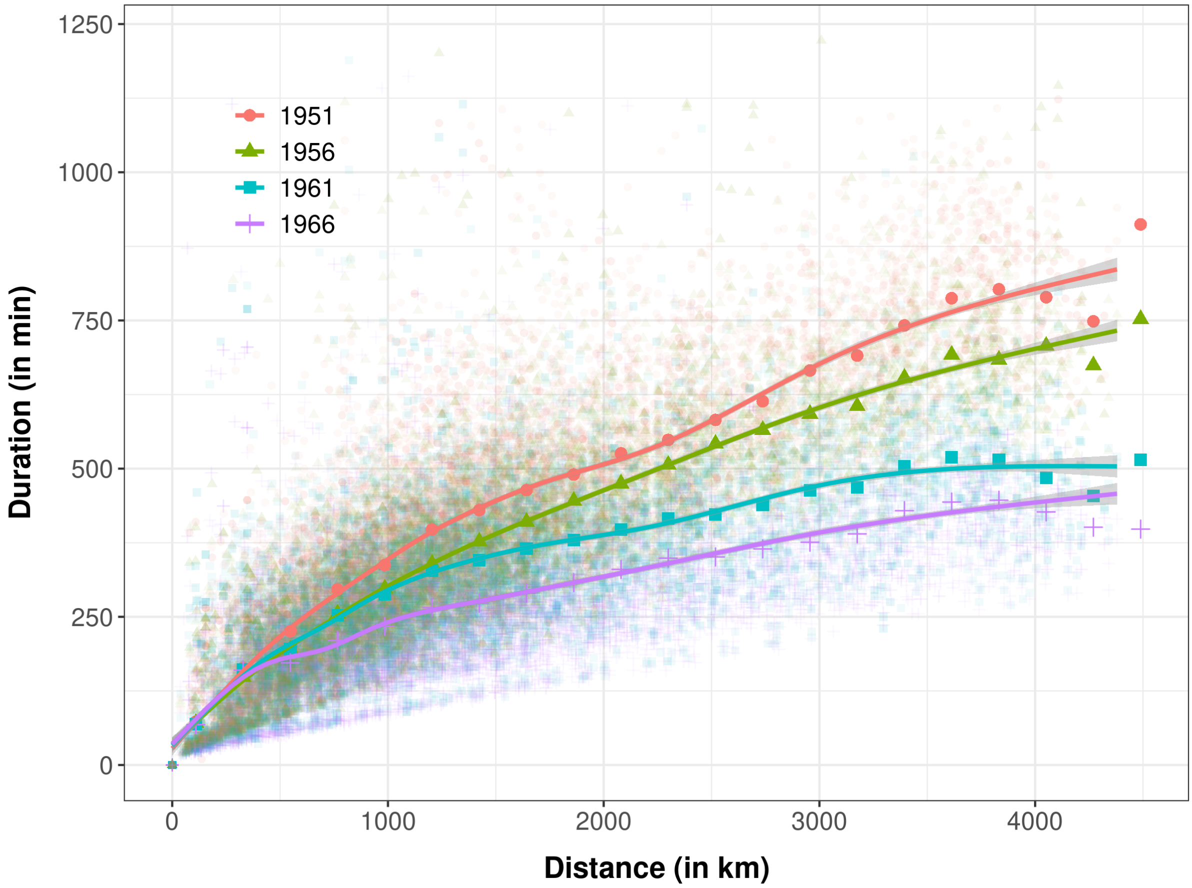

Figure 24: Travel Times between US MSAs.

Figure 25 that the change in travel time is accompanied by a reduction of the amount of legs needed to connect two MSAs at every distance. This reduction is specially marked between 1951 and 1956, and 1961 and 1966. In Figure 26 we open up the change in travel time by the way an MSA pair was connected in 1951 and 1966: either directly (non-stop flight) or indirectly (connecting flight). We observe that much of the increase in travel time for MSA pairs less than 250km apart comes from routes that were operated non-stop and then it needed a connecting flight. Interestingly, for MSA-pairs more than 2,000km apart travel time reduced on average 42% for those pairs that were connected indirectly in both periods, and 51% for those that switched from indirect to direct. This fact shows the relevance of improvements in flight technology even for MSAs not directly connected. It could be the case that a reduction in the amount of legs or an increase in frequency of flights reduces layover time. In Figure 28 we compare the change in travel time from 1951 to 1966 with a fictitious change in travel time in which we eliminate layover time in both time periods. We observe that the average change in travel time is stronger at every distance if we disregard layover time. This implies that the relative importance of layover time over total travel time increases between 1951 and 1966, preventing total travel time to decrease proportionally to the change of in-flight travel time.

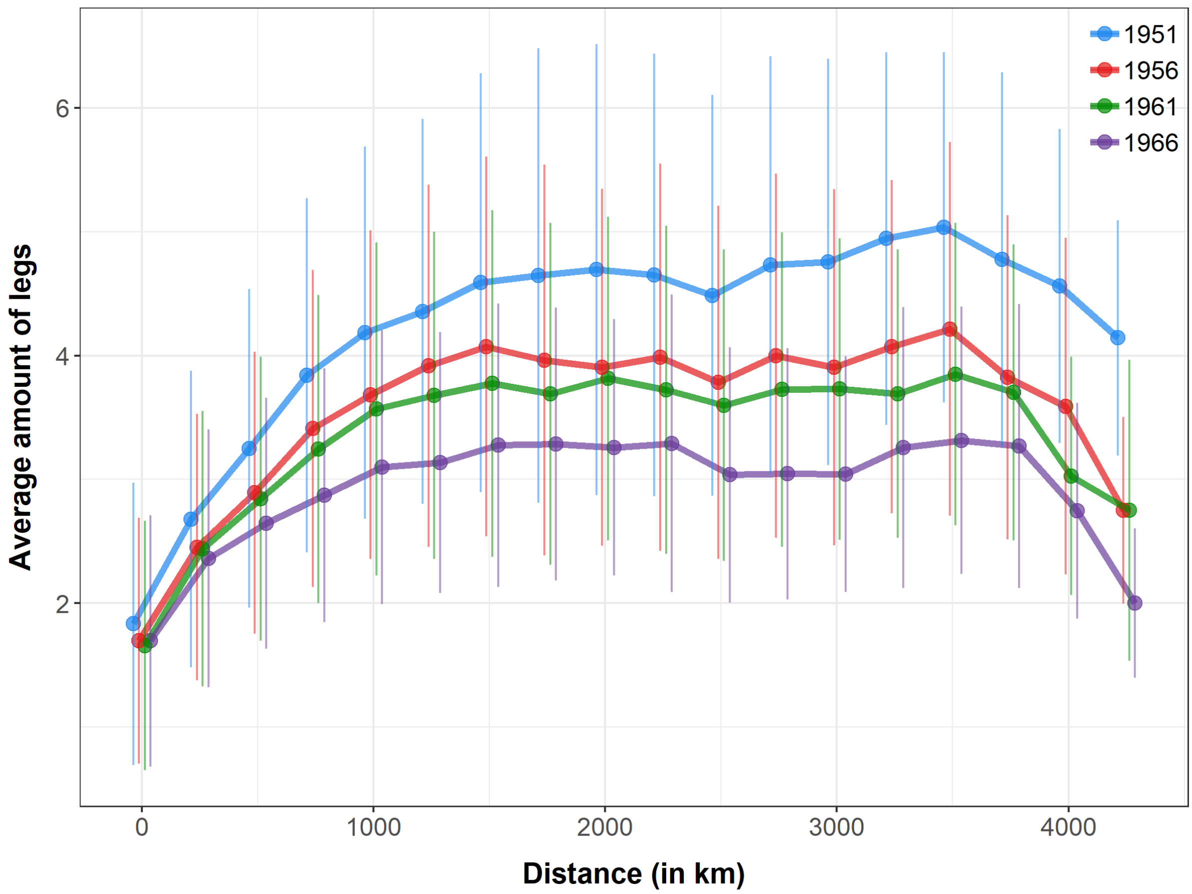

Figure 25: Average amount of legs per route

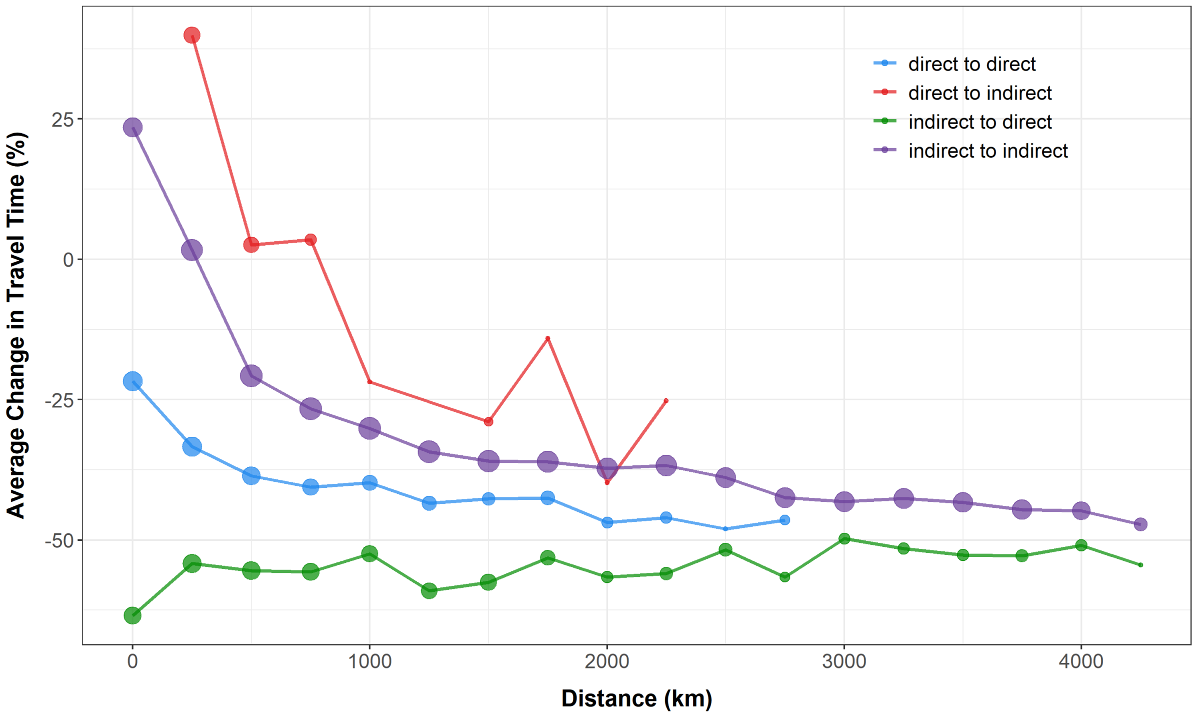

Figure 26: Change in US travel time 1951 to 1966: connections
89

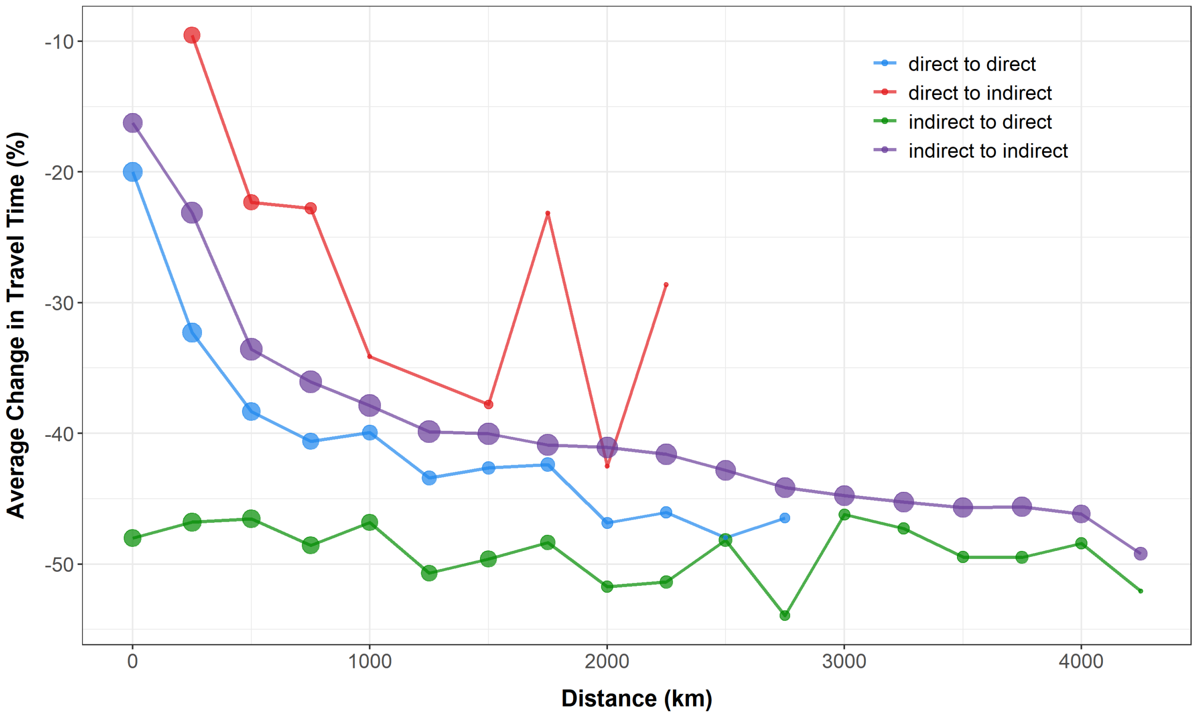

Figure 27: Change in US travel time 1951 to 1966: connections, discarding layover time

90

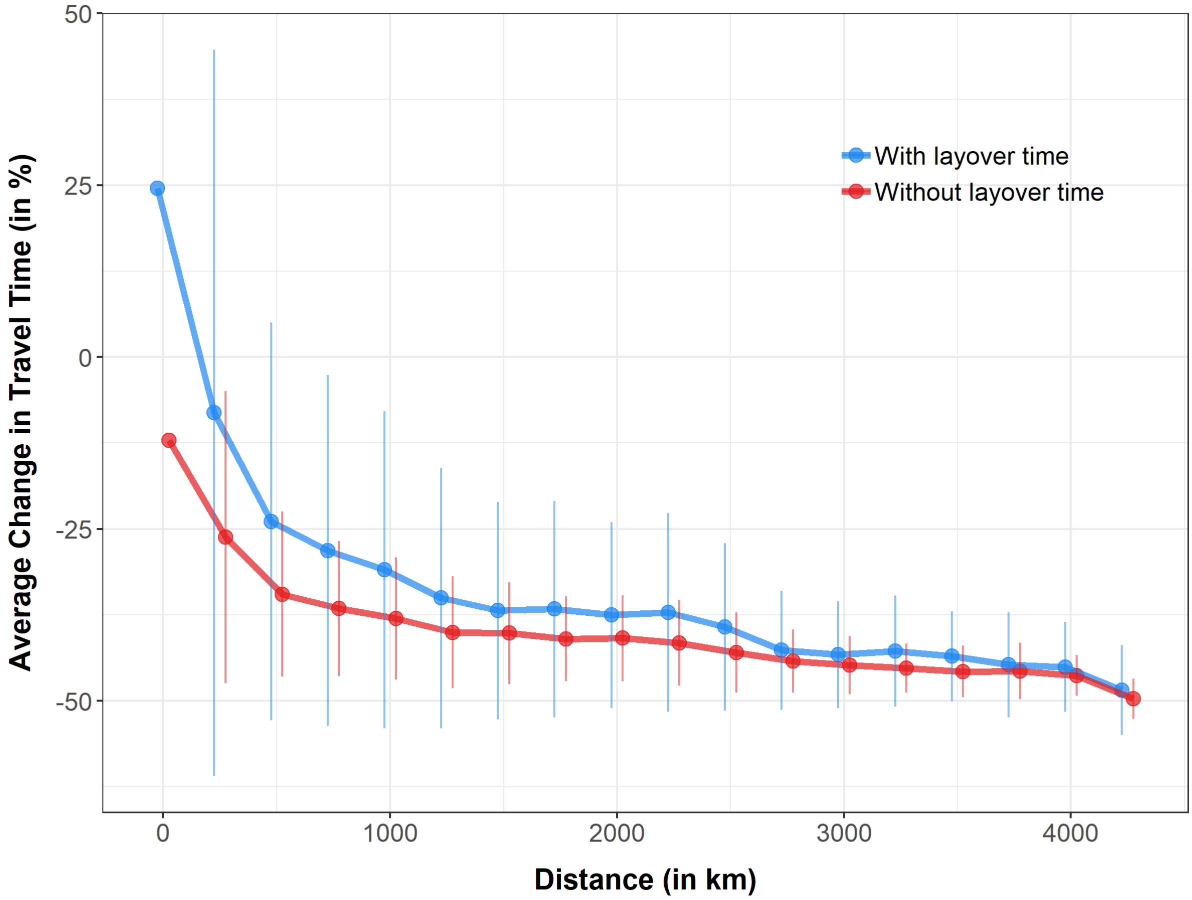

Figure 28: Change in US travel time 1951 to 1966: layover time

In figure 29 we show the average change in travel time in three counterfactual flight networks. The first counterfactual fixes the flight routes[91] and allows aircraft speed to evolve. The second counterfactual fixes aircraft speed and allows flight routes to evolve. The third counterfactual allows both flight routes and aircraft speed to evolve. We obtain that around 90% of the change in travel time is due to the change in speed of aircrafts, while around 10% of the change is due to the change in the flight routes. In the figure 30 in the appendix we show that the proportion is relatively constant for all distances. This confirms that most of the observed changes in the network are due to improvements in the flight

[91]Fixes the origin-destination airports that are connected with a non-stop flight

technology.

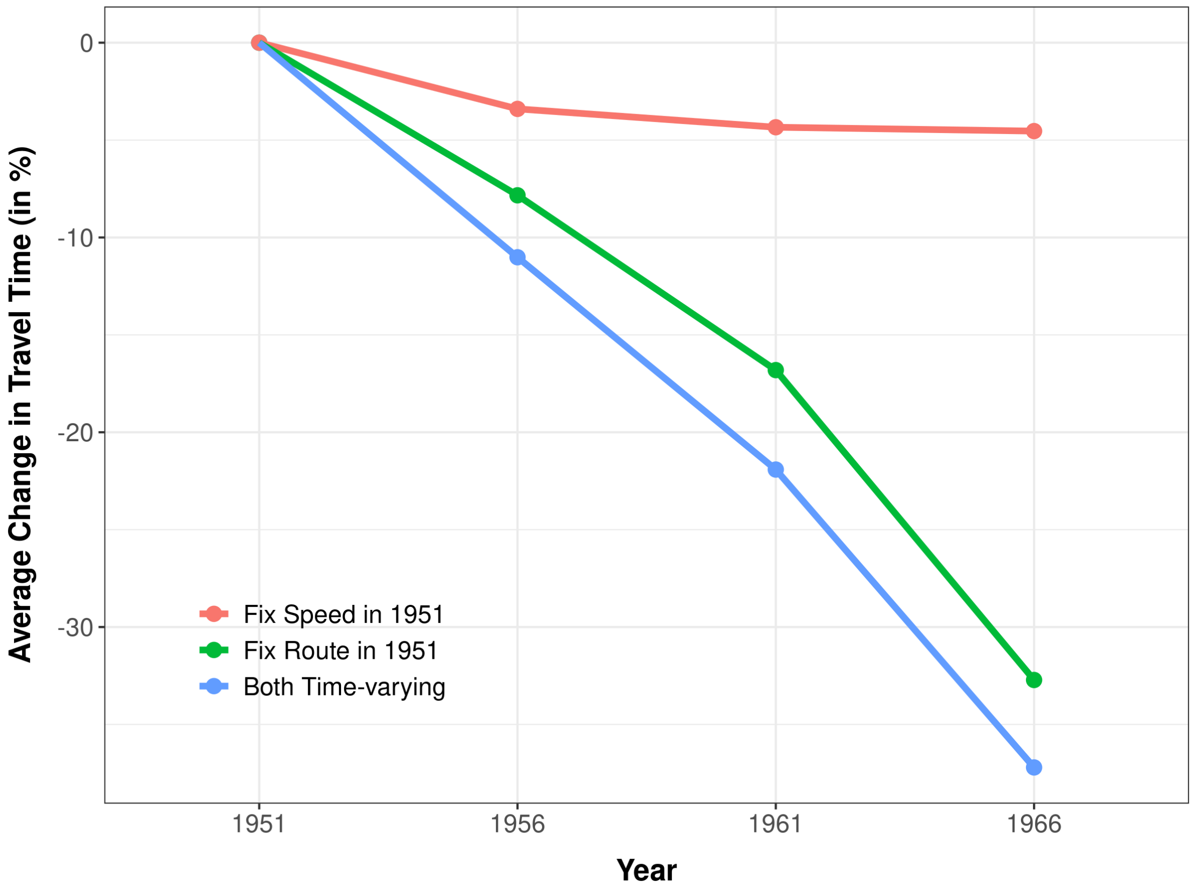

Figure 29: Counterfactual change in travel time

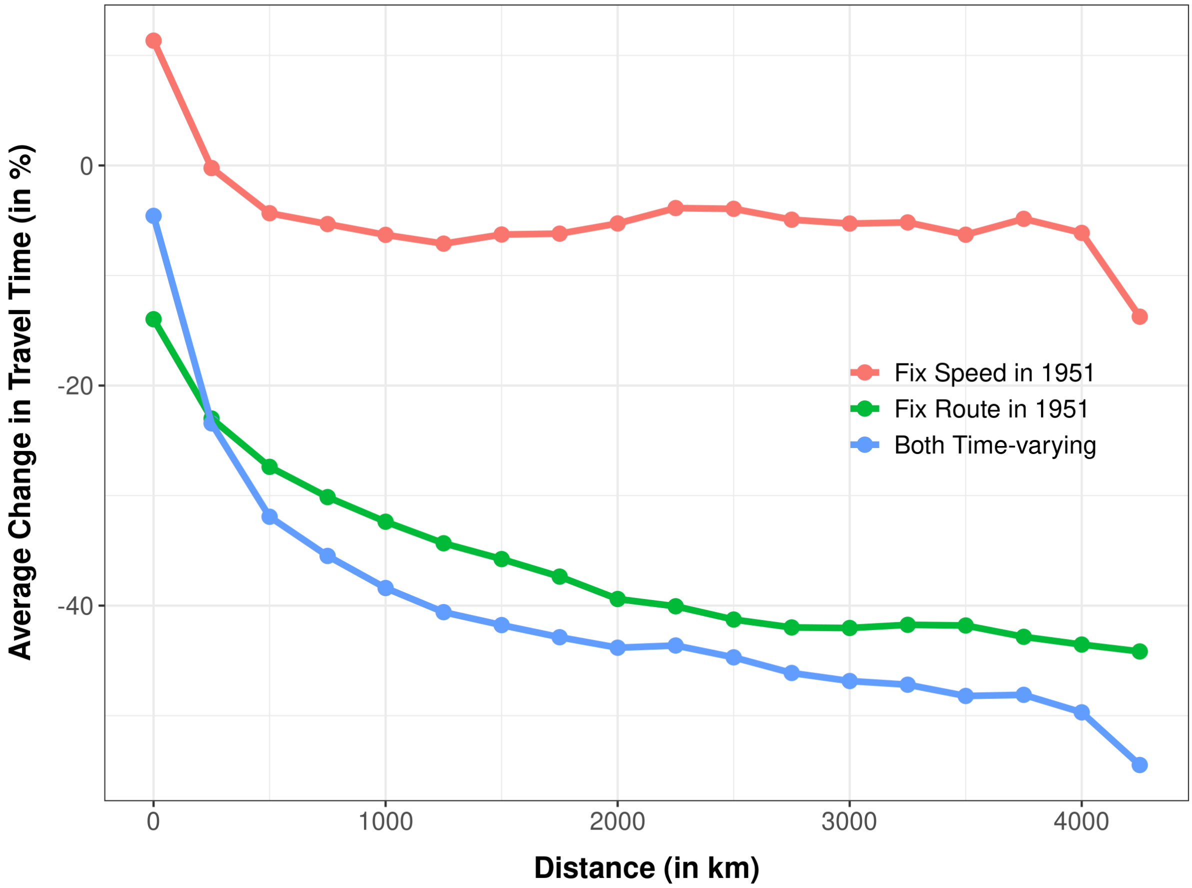

Figure 30: Counterfactual change in travel time 1951-1966

In addition to the changes over time in the network leading to faster travel times, another feature of the US airline industry becomes salient in the data: airlines' regional specialization. As figure 31 shows, while there was competition among the airlines in our dataset on the major routes (Lower West Coast to the Midwest and Upper East Coast to the Midwest), some airlines are very specialized and face no competition from any of the other five airlines on certain routes. In particular, NW controls the routes connecting Seattle to the Midwest and EA controls much of the connections from Florida to New York and surroundings.

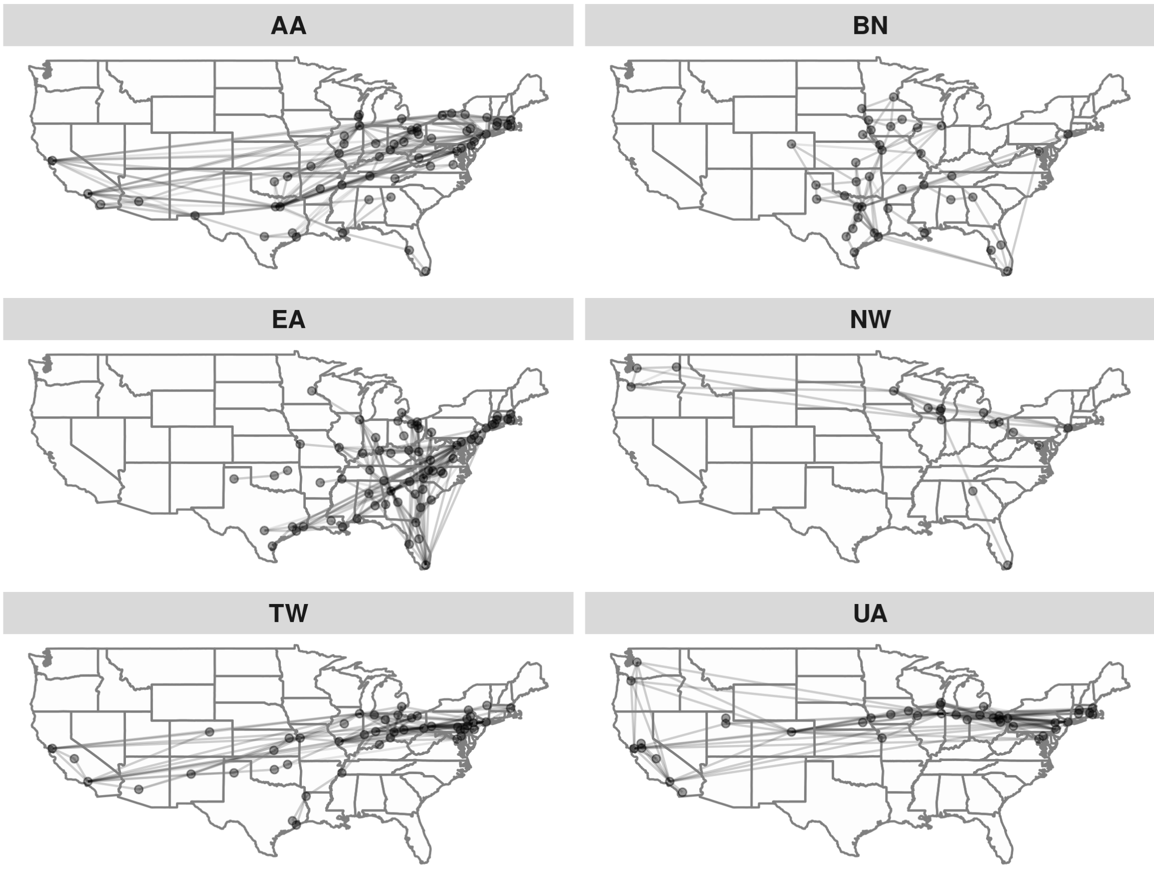

Figure 31: Flight Network in 1956 by Airline (weighted by log frequency).

B. Appendix: Descriptives patent data

In this appendix we describe facts that we observe in the US patent data, for patents filed[92] between 1945 and 1975. US patents data containing citations and

[92]Filing year, also called application year, is the closest date to the date of invention that is present in the data and it represent the date of the first administrative event in order to obtain

filing year have been downloaded from Google Patents. Then, it was merged with multiple datasets (see Appendix Patent Data Construction for more details):

- Technology classification: NBER patent database.
- Geographical localization of inventors: Histpat and Histpat International for patents published until 1975, Fung Institute for patents published after 1975. Both matched to 1950s Metropolitan Statistical Areas (MSAs).
- Ownership: Kogan et al. (2017) for patents owned by firms listed in the US stock market, Patstat for the remaining patents not matched to Kogan et al. (2017).

We highlight two details from the matching process: 1. During filing years 1971-1972 the rate of non-geocoded patents increases, possibly due to Histpat and Fung data not being a perfect continuation one of the other. 2. Kogan et al. (2017) seems to use a matching method based on the patent owner declared in the patent text, as Patstat does. Specially, Kogan et al. (2017) does not explicitly say if it takes into account firm-ownership structure to determine patent ownership, neither does Patstat.

For the analysis presented in this appendix we will use the resulting dataset from the matching procedure, where unless evident or noticed, we will use only patents that have inventors within MSAs. We discard patents that have inventors in multiple MSAs and patents that belong to government organizations or universities. We assign patents to technology categories using fractional count: if a patent is listed in two technology categories, then we assign half a patent

a patent. In the other hand, publishing or also called granting year, is the later year in which the patent is granted. The difference between filing and granting year depends on diverse non-innovation related factors (as capacity of the patent office to revise applications) and changes over time. Hence filing year is the date in our data that approximates the best to the date of invention.

to each category. We discard self citations (citations in which the citing patent owner is the same as the cited patent owner) because self-citations may be due to different incentives.

B.1. Geography of patents

In figure 32 we observe that the matching rate decreases from around 95% before 1970, to around 80% in 1971 and 1972, and then it stabilizes around 99% after 1975. Hence, geogprahical results during years 1970-1975 will contain an increased amount of measurement error.

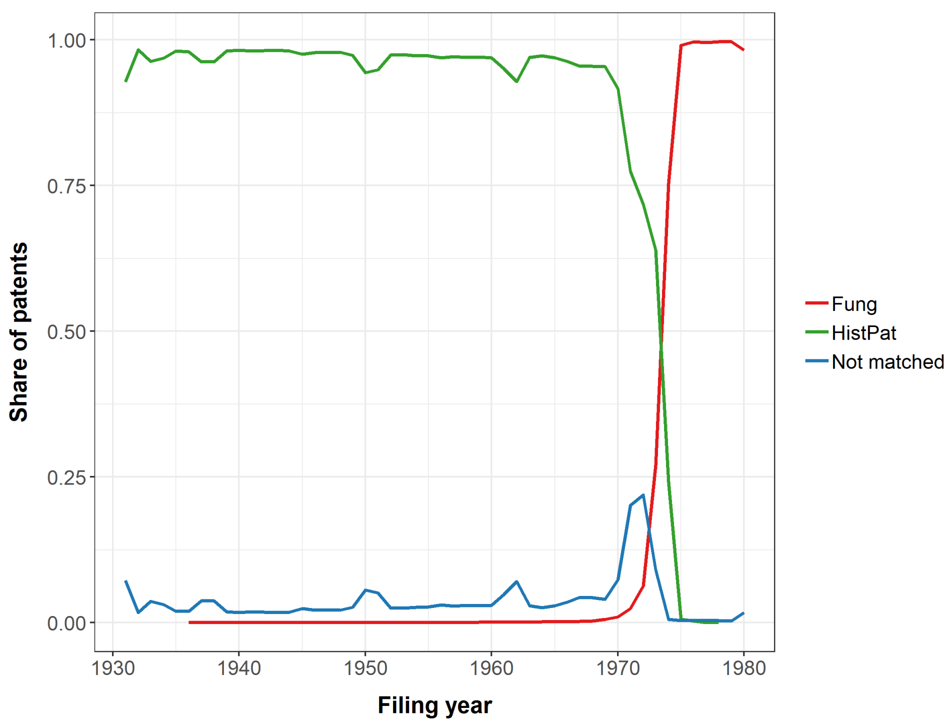

Figure 32: Non-matching rate HistPat, HistPat International and Fung

Figure 33 shows the share of patents that have inventors inside MSAs, and figure 34 displays the same by technology category.[93] We observe that the share

[93]Technologies are aggregated to six big groups, as explained in HJT 2002

of patents in MSAs decreases over time, and this is also the case for each of the technology classes. Indeed, the Herfindahl index of patenting inside-outside MSAs goes from 0.69 in 1945 to 0.55 in 1969, reflecting that patents are more equally spread within-outside MSAs.[94] We also compute the Herfindahl index across-MSAs, finding that for patents within MSAs, concentration of patenting within few MSAs also decreases: the value goes from 0.061 in 1945 to 0.044 in 1969.

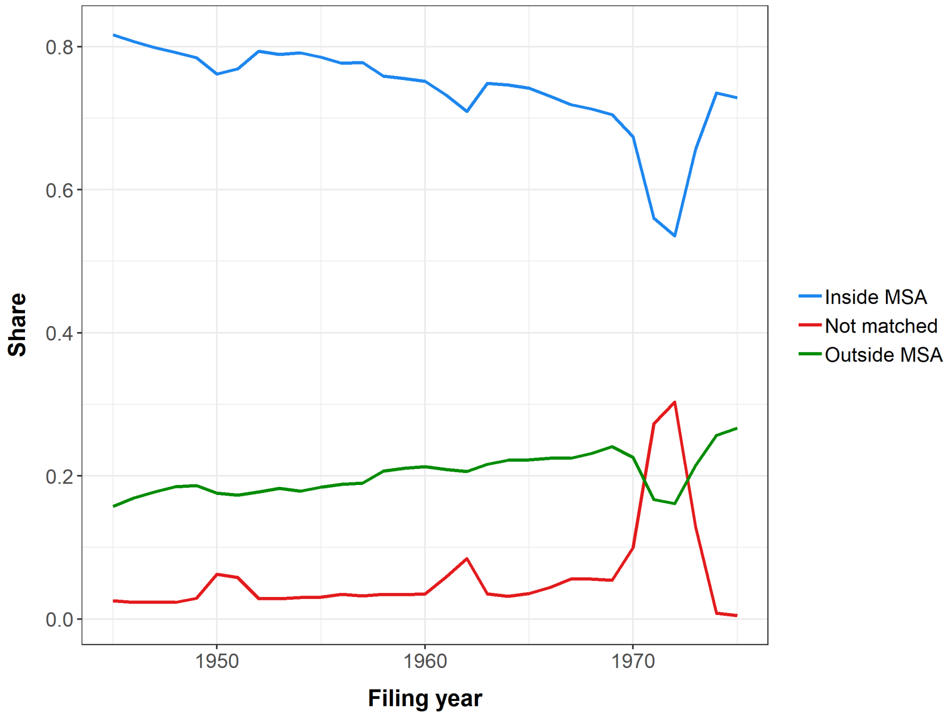

Figure 33: Share patents in Metropolitan Statistical Areas

[94] We use MSAs borders of 1950s. We have to re do the analysis with more recent MSA borders to see if we observe the same pattern or different. If MSAs have changed, probably they have expanded. If we observe different trends it prbably means that new patents had inventors just outside the existing MSAs. While MSAs have changed over the years, we believe that this trend would still be present if we let MSA borders evolve over the years.

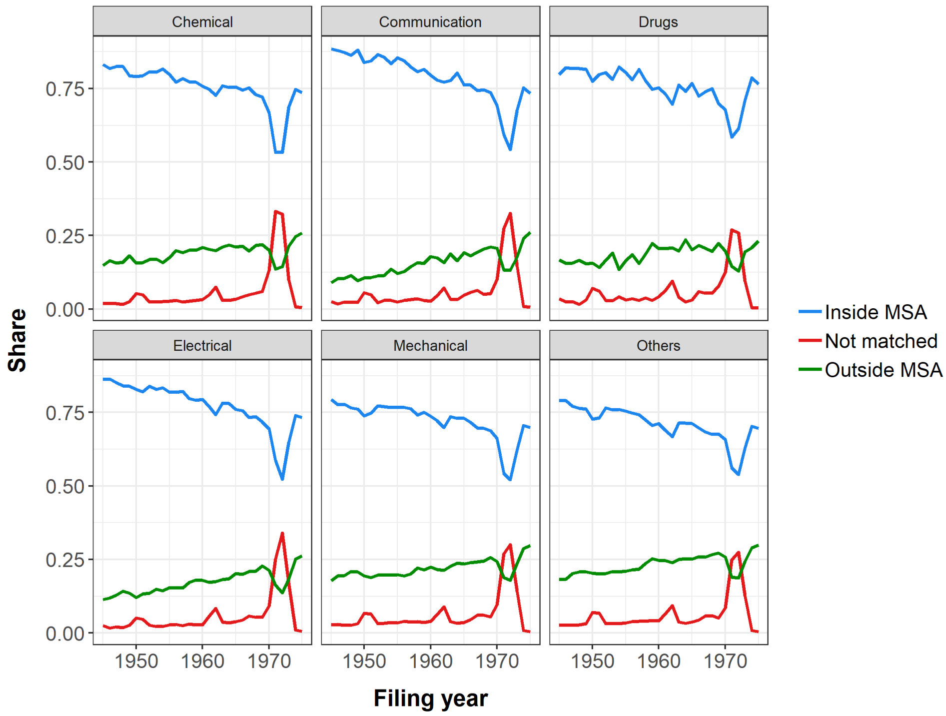

Figure 34: Share patents in Metropolitan Statistical Areas

At the same time, we observe that the amount of MSAs with at least one patent increases over time as displayed in Figure 35. Note that technology categories that have more patents will mechanically have higher probability of being filed by inventors in more MSAs. Figure 36 displays share of patents of each technology computed using only patents contained in MSAs.

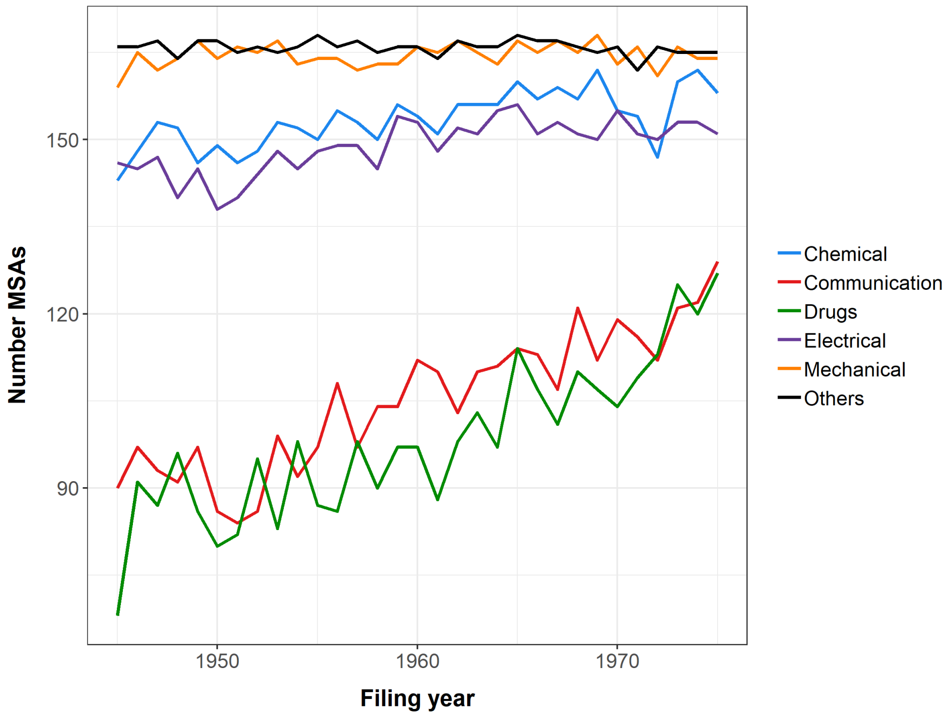

Figure 35: Number of Metropolitan Statistical Areas with at least one patent

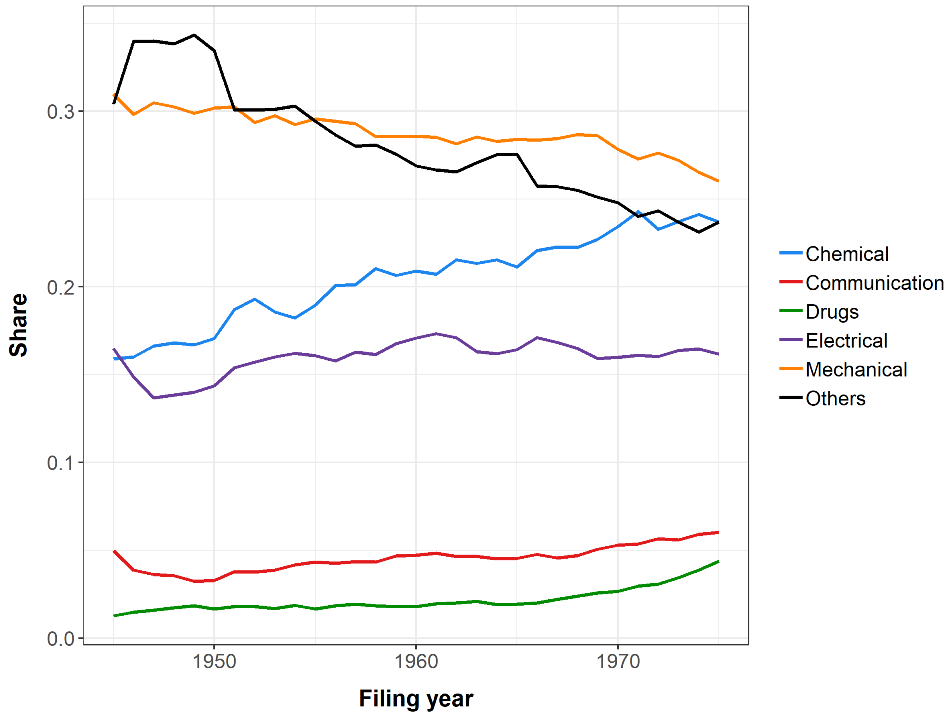

Figure 36: Share of patents by technology category

Figure 37 shows MSAs in the US colored depending on the amount of patents filed with inventors in the MSA, for all patents filed in 1949-1968.

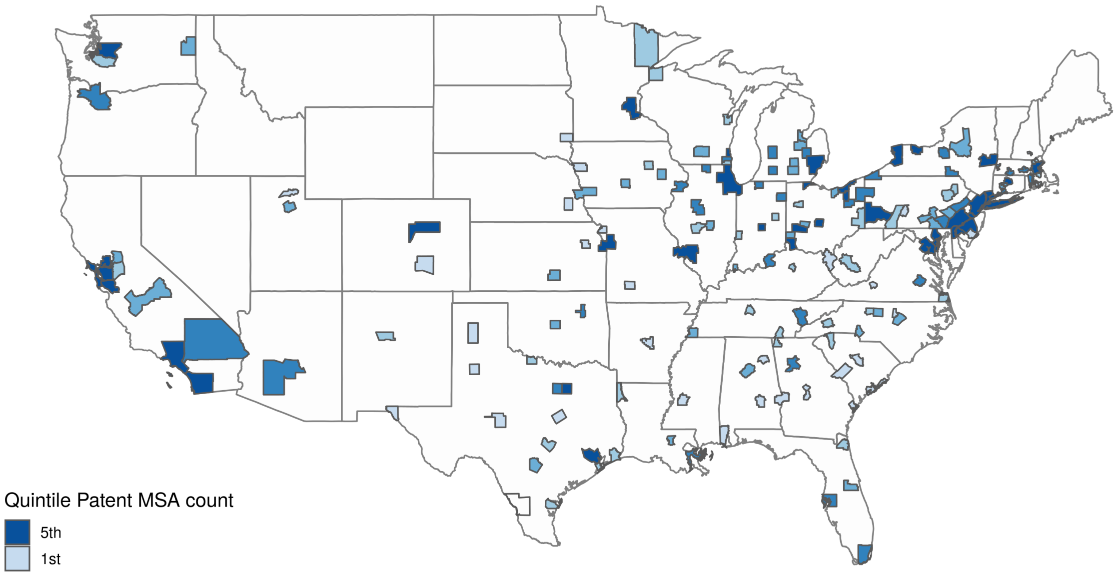

Figure 37: Geographical distribution of patents

B.2. Geography of citations

Using the location of inventor of the citing patent and inventor of the cited patent we can compute the *distance of a citation*. Figure 38 shows, by technology category, the evolution over time of different quantiles of citation distance. We observe that for all technologies 25% of citations happen between inventors located at less than 500km of each other over all the time period.[95] For the following 50% of citations we observe that over time, with the exception of Drugs patents, inventors cite other inventors located further away. This is specially the case for Communication and Electrical patents, in which their citation in the quantile 0.75 happens at around 1,000km in 1945 and increases to 2,500km in 1970: the mass of citations shifts to higher distances over time.

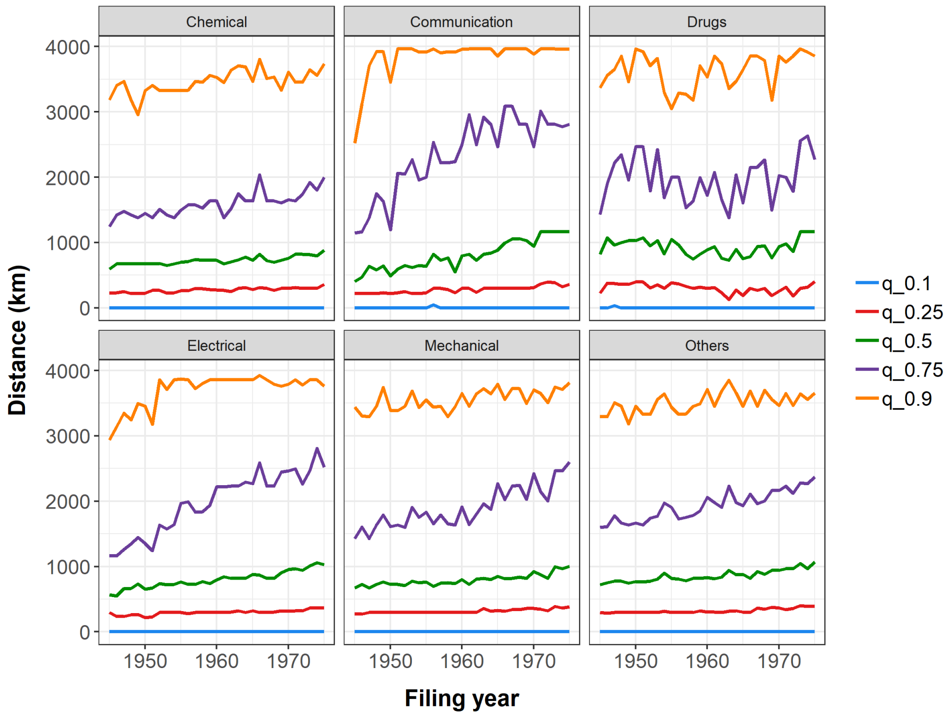

Figure 38: Citation distance by technology

[95]As a reference, the straight line distance from New York City NY to other places is: Boston MA 300km, Chicago IL 1,140km, Dallas TX 2,200km, San Francisco CA 4,130km.

In the same spirit as how Input-Output tables of industries are constructed, we can use citations as a reflection of sourced (input) knowledge. In this case, we interpret the cited patent as being a source of knowledge, and the citing patent as being a destination. In Figure 39 we aggregate citations by citing-cited technology category in the years 1949-1953. Rows represent the source technology and columns the destination technology. Columns should sum to 1 (round errors may exist). We highlight in bold those IO coefficients that are higher than 0.1. We observe that the diagonal has coefficients greater than 0.5, implying that technologies rely on themselves to create new knowledge. At the same time, we observe the importance of Electrical to create Communication technologies, and the small relevance of Drugs for every other technology.

Given that we know the location of the citing inventor and cited inventor, we can construct a *Distance Input-Output* matrix, and analyze how it changes over time. Similarly, making use of the travel time data, we can construct a *Travel Time Input-Output* matrix. In order to be consistent across *distance* and *travel time* Input-Output tables, we are going to restrict to citations for which we have travel time.[96] In Figure 40 we present the distance of the median citation for the period 1949-1953, and in Figure 41 we present the change[97] in distance from 1949-1953 to 1964-1968, exposed as a percentage point divided by 100. Figure 42 displays the change in travel time.

We observe that the distance of the median citation increased over time for most of the cases. Specially, we observe that for io coefficients greater than 0.1 the distance of median citation increases for all except Drugs-Drugs. In fact, distance of citations to Drugs from every other technology did not change much during

[96]Distance is not a constraint, given that we can calculate it for every two locations.

[97]Values represent the percentage change/100, so a value of 1 would imply an increase in median distance of citation of 100%

the period. Together with the fact that Drugs are almost only cited by Drugs (see row of Drugs in Figure 39), we could interpret this as drugs not changing geographical location.[98] We may think that Drugs may be a geographically immobile technology, maybe due to laboratory regulations.[99]

In the case of changes in travel time, we observe that it decreases for every source-destination pair. Note that travel time changes are both due to a change in travel speed but also due to a change in citation distance. Given that we observe a decrease in travel time, it means that speed increased more than distance. Specially, we may think that citations may be redirected to destinations (cited locations) that received a bigger proportional decrease in travel time, which may be at the same time farther away. Because of this mix of distance and speed, in Figure 44 we display the change in speed of the median citation. Interestingly, we observe that the technologies that observe the highest increase in citation distance Communication and Electrical are also the ones that have the highest increase in speed. This finding leads us to reinforce our belief that certain technologies are able to better accommodate their citing-cited locations when they face a change in speed: they are able to exploit better the shock to transportation. Another way to tease out the effect of the endogenous reaction of citation distance to travel time is by fixing the citation network of 1949-1953 and letting travel time evolve. Figure 43 shows the change in travel time applied to citations of 1949-1953. We observe that the decrease of travel time in Figure 43 is much stronger than in Figure 42, specially for citing technologies Communication and Electrical. The difference between the two is exactly consequence of the different

[98]Note that if Drugs technologies were to be patented in geographical locations distant to the initial ones, they would probably cite previously existing knowledge and hence we would observe a change in citation distance. In other words, patents already filed have their inventors' addresses fixed, so a change in citation distance will mainly come from a change in the location of the citing inventor.

[99]We are working to create a measure of degree of mobility of each technology.

citation distance.

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**0.74**	0.01	**0.13**	0.03	0.05	0.05
Communication	0	**0.6**	0	0.07	0.01	0.01
Drugs	0.01	0	**0.6**	0	0	0.01
Electrical	0.03	**0.28**	0.03	**0.7**	0.05	0.04
Mechanical	**0.11**	0.07	0.07	0.1	**0.72**	**0.15**
Others	**0.11**	0.05	**0.16**	0.09	**0.16**	**0.75**
Total	1	1	1	1	1	1

Figure 39: Input-Output of technologies 1949-1953

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**675**	641	**537**	672	707	709
Communication	641	**594**	520	560	655	672
Drugs	736	766	**1041**	802	742	879
Electrical	**672**	**583**	1007	**641**	672	766
Mechanical	**675**	675	819	672	**759**	**820**
Others	728	672	868	721	**785**	**820**

Figure 40: Median distance Input-Output of technologies 1949-1953

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**0.05**	0.08	**0.21**	0.31	0.08	0.16
Communication	-0.12	**0.73**	1.19	0.72	0.43	0.14
Drugs	-0.11	0.99	**-0.11**	0.27	0.29	-0.04
Electrical	**0.26**	**0.65**	-0.15	**0.36**	0.21	0.18
Mechanical	**0.14**	0.30	0.26	0.31	**0.08**	**0.00**
Others	0.09	0.22	0.15	0.30	**0.12**	**0.15**

Figure 41: Percentage change median distance Input-Output of technologies 1949-1953 to 1964-1968

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**-0.36**	-0.31	**-0.36**	-0.20	-0.35	-0.39
Communication	-0.35	**-0.11**	0.33	-0.14	-0.25	-0.30
Drugs	-0.43	0.10	**-0.39**	-0.33	-0.33	-0.39
Electrical	**-0.16**	**-0.18**	-0.39	**-0.21**	-0.32	-0.34
Mechanical	**-0.37**	-0.26	-0.16	-0.28	**-0.34**	**-0.34**
Others	-0.39	-0.33	-0.31	-0.33	**-0.31**	**-0.33**

Figure 42: Percentage change median travel time Input-Output of technologies 1949-1953 to 1964-1968

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**-0.43**	-0.47	**-0.38**	-0.44	-0.43	-0.38
Communication	-0.46	**-0.39**	-0.38	-0.39	-0.47	-0.41
Drugs	-0.37	-0.50	**-0.39**	-0.39	-0.39	-0.39
Electrical	**-0.44**	**-0.39**	-0.42	**-0.38**	-0.42	-0.39
Mechanical	**-0.43**	-0.44	-0.45	-0.43	**-0.41**	**-0.41**
Others	-0.41	-0.39	-0.41	-0.39	**-0.39**	**-0.39**

Figure 43: Percentage change median travel time Input-Output of technologies 1949-1953 to 1964-1968, fixed citation network 1949-1953

Source/Destination	Chemical	Communication	Drugs	Electrical	Mechanical	Others
Chemical	**0.67**	0.67	**0.66**	0.84	0.63	0.74
Communication	0.62	**0.85**	0.59	0.91	0.63	0.57
Drugs	0.54	0.69	**0.57**	0.66	0.72	0.62
Electrical	**0.77**	**0.83**	0.45	**0.81**	0.81	0.69
Mechanical	**0.72**	0.71	0.76	0.81	**0.68**	**0.68**
Others	0.76	0.61	0.78	0.83	**0.72**	**0.75**

Figure 44: Percentage change median speed Input-Output of technologies 1949-1953 to 1964-1968

B.3. Compustat vs. Others

The identifier for assignee (owner) of a patent comes from two sources that we match sequentially. We first match on Kogan et al. (2017) that matches patents with their owners who are listed in the stock market (Compustat firms). Second, we match with Patstat ID. We obtain around 500 Compustat IDs, and other 200,000 unique Patstat IDs. In each year, we have patents being filed by around 300 unique Compustat IDs and 12,000 Patstat IDs. We believe that Compustat firms may be bigger than those obtained from Patstat IDs, so we decide to compare the two groups.

In Figure 45 we observe that most patents are not filed by Compustat firms. Nonetheless, the amount of patents filed by Compustat firms increases over time, while the amount of patents filed by Others decreases. Given that we know the owner of the patent and the location of the inventors, we use a pair 'owner ID-inventor location' as a proxy for an establishment (associated with that ID, and placed in that location). Then, we can look at the distribution of establishments per ID: the distribution of Compustat firms is shifted towards more establishments per ID with respect to Patstat ID (not displayed).[100] However, when comparing the right tail of the distribution of size of IDs measured by amount of establishments, we find that Compustat firms are relatively similar in size to Patstat IDs.[101] Figure 46 shows the top 1, 10, and 100 ID of Compustat

[100] We obtained the distribution of amount establishments per ID both for Compustat and for Patstat IDs. We observe that for Compustat, 25% of IDs have 1 establishment over all the time period, 50% have 2 establishments or less and 75% have 4 establishments or less. The firm in the quantile 0.99 has 14 establishments in 1945, increasing to 22 establishments in 1968. Note that there are around 300 Compustat firms per year, so the top 1% includes 3 firms. Differently, there are around 12,000 Patstat IDs per year, so the top 1% includes 120 IDs. In the case of Patstat IDs, we observe that 90% of the IDs have 1 establishment, and 99% of the IDs have 3 establishments or less. The ID in the quantile 0.999 has 7 establishments in 1945 and 10 establishments in 1969.

[101] It remains to be checked who are these Patstat IDs. It is possible that they are government institutions or universities.

and Patstat (*Other*) by amount of establishments. Finally, for multi-establishment firms we can compute the average distance across establishments. In Figure 47 we observe that the average distance of establishments within ID increases over time, and this is specially the case of Other IDs. This implies that the firm is having establishments more spread out over the geographical space.

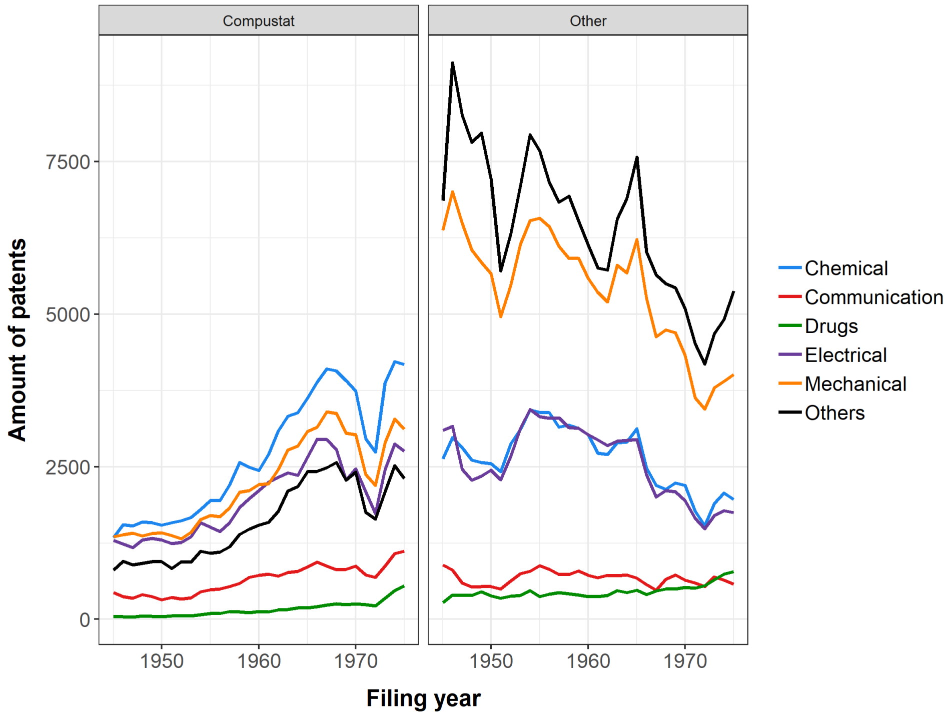

Figure 45: Amount of patents by owner type and technology

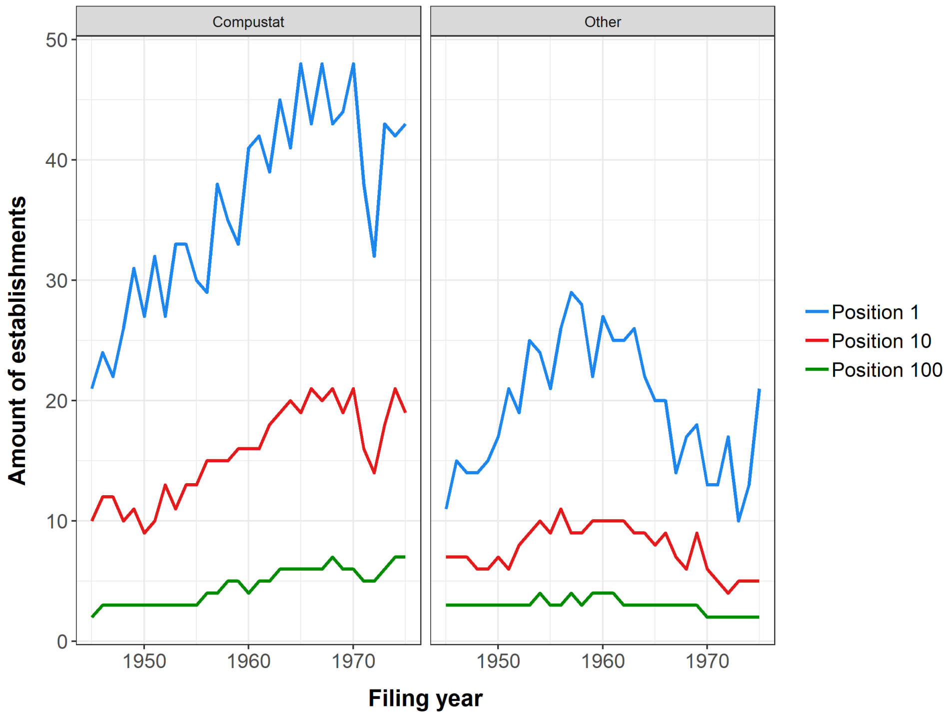

Figure 46: Top IDs by amount of establishments

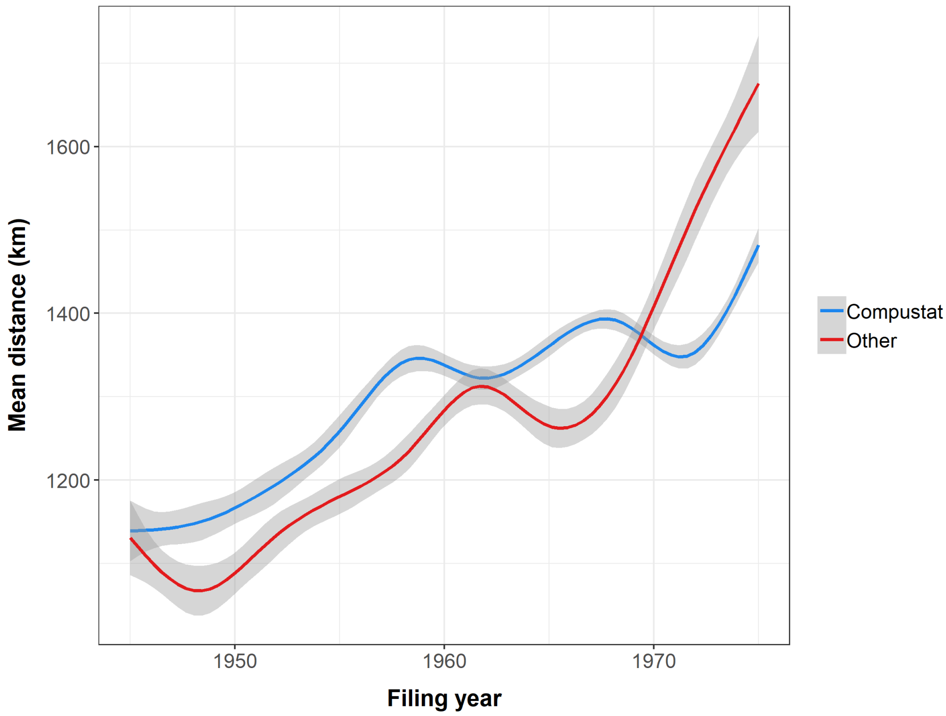

Figure 47: Mean distance of establishments within firm

B.4. General Electric research establishments

Using the patent owner identifier we can display the geographical distribution of research establishments for a selected firm. Figure 48 shows the research establishments of General Electric in the period 1945-1953. We say that a firm F had a research establishment in location i in time period t if firm F filed at least one patent in time period t with inventors located in location i. The headquarters location q of firm F is defined as the location in which the firm filed the largest amount of patents in the period 1945-1953. General Electric had research establishments in 62 MSAs in the period 1945-1953, and the MSA with the largest amount of patents was Schenectady, New York. Figure 49 shows the location of patents cited by patents filed by General Electric with inventors in Fort Wayne, Indiana, in the period 1949-1953. Figure 50 shows the research establishments of General Electric during periods 1949-1953 and 1964-1968. General Electric had research establishments in 51 MSAs in 1949-1953 and in 76 MSAs in 1964-1968. 42 out of them appear in both time periods.

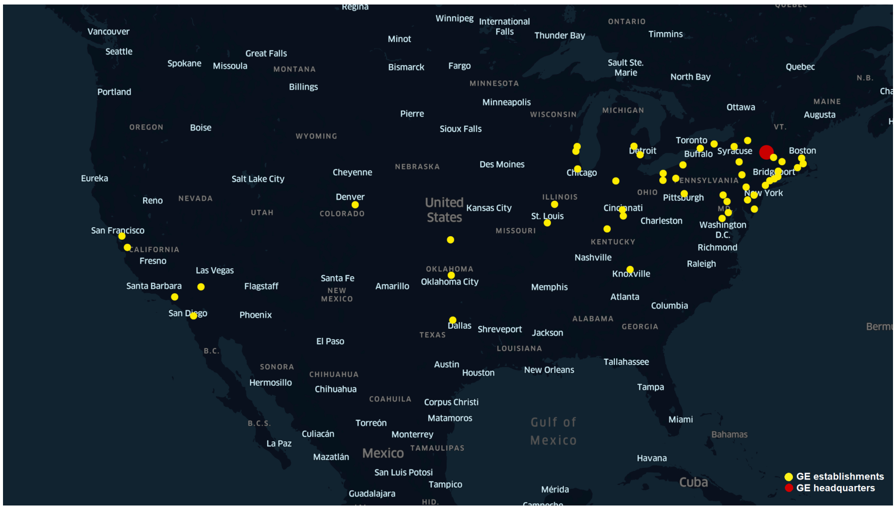

Figure 48: Research establishments of General Electric 1949-1953

Figure 49: Citations General Electric at Fort Wayne IN 1949-1953

Figure 50: Change location research establishments of General Electric between 1949-1953 and 1964-1968

C. Appendix: Additional results

C.1. Diffusion of knowledge

C.1.1. Heterogeneous effects

First, we perform an intensive margin/extensive margin decomposition of the effect of travel time on citations. We find that the effect is coming from both margins. In the instrumental variables approach, the intensive margin is only statistically different from zero for distance greater than 2,000km, while for the extensive margin it is for distance greater than 300km. Results for the baseline analysis are shown in Table 15 and for the IV estimation in Table 16.

Second, we investigate if the elasticity varies by the degree of concentration of patents across establishments in the citing technology or cited technology, we find no statistically significant heterogeneous effect. Results are shown in columns (1) and (2) of Table 18.

Third, we check if the elasticity varies by the median forward and backward citation lags of the cited and citing technologies. We find that the elasticity of citations to travel time is *more negative* both for technologies that accumulate citations during a longer time period and for technologies that cite older patents. To be able to precisely show if it is *newer* or *older* technologies that diffuse better as consequence of the jet requires an analysis with the citation level forward and backward lag, and not using the median lag in the technology. Nonetheless, the results seem to suggest that jets improved the diffusion of *older* technologies. Results are shown in columns (3) and (4) of Table 18.

Fourth, we extend the sample of patents to include patents with a patent owner identified as a government organization or university. Column (5) of Table 18 opens the elasticity of citations to travel time by whether the citing patent belongs to a government organization of university. Column (6) includes a dummy for whether the cited patent belongs to a government organization or university. We do not observe a particular change in the pattern of the elasticity of citations to travel time.

Sixth, we extend the sample to include self citations (citations in which the citing and cited patents belong to the same patent owner F). Column (7) of Table 18 shows that the elasticity is not statistically different for self citations.

Seventh, we check if the elasticity varies with the level of innovativeness of the citing firm. It may be the case that those firms that actually have the -time and monetary- budget to take a plane are only the most innovative ones. We rank firms F in technology h according to the amount of patents filed by F in technology h at the initial time period 1949-1953. We define quantile 0.00 as all those firms that did not file patents in 1949-1953, while quantile 0.01 is assigned to those that filed patents but not as many as to be in the quantile 0.25 or higher. Results are shown in Table 17. We do not find a particular pattern related to the initial innovativeness.

Eighth, we check if the elasticity varies with the citing technology, cited technology and citing-cited technology pair. Results are shown in Table 19 and Table 21. We find that the elasticity is negative and significant mainly when the citing and cited technology are the same. In Appendix B.4 we show that most citations happen within a technology, so most identification power would be when citing and cited technologies are the same.

	PPML		log-log		linear probability	
Dep. variable: *citations*	$cit_{FiGjhkt}$		$log(cit_{FiGjhkt})$		$cit_{FiGjhkt} > 0$	
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time)	-0.083^{***} (0.019)		-0.071 (0.098)		-0.013*** (0.003)	
log(travel time):0-300km		0.019 (0.036)		0.318^{**} (0.152)		-0.0045 (0.005)
log(travel time):300-1000km		-0.089^{***} (0.023)		-0.265^{*} (0.145)		-0.008^{***} (0.003)
log(travel time):1000-2000km		-0.094^{***} (0.032)		-0.231 (0.209)		-0.013^{***} (0.003)
log(travel time):+2000km		-0.169^{***} (0.039)		-0.424^{**} (0.192)		-0.024^{***} (0.005)
N obs. effective	4,703,010	4,703,010	16,412	16,412	10,106,940	10,106,940
R2	0.88	0.88	0.86	0.86	0.70	0.70

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 15: Elasticity of citations to travel time: intensive and extensive margin

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\beta\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (3) shows the result of an OLS estimation of $log(citations_{FiGjhkt}) = \alpha\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with a sample of establishment-technology pairs ($FiGjhk$) that have positive citations in all periods. Column (5) shows the result of an OLS estimation of $\mathbb{1}\{citations_{FiGjhkt} > 0\} = \gamma\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with the same sample as (1). Column (2), (4) and (6) open, respectively, the coefficients β, α, γ by distance between the citing establishment Fi and the cited establishment Gj. Standard errors are presented in parentheses. Columns (1) and (2) present coefficients and bootstrap standard errors jackknife bias corrected. Columns (3) through (6) present standard errors clustered at the non-directional location pair (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	IV PPML		IV log-log		IV linear probability	
Dep. variable: *citations*	$cit_{FiGjhkt}$		$log(cit_{FiGjhkt})$		$cit_{FiGjhkt} > 0$	
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time)	-0.152^{***} (0.029)		-0.396** (0.175)		-0.027*** (0.004)	
log(travel time):0-300km		-0.076 (0.221)		1.324 (1.680)		-0.028 (0.036)
log(travel time):300-1000km		-0.134^{***} (0.044)		-0.148 (0.378)		-0.022*** (0.007)
log(travel time):1000-2000km		-0.112^{**} (0.047)		-0.314 (0.200)		-0.021*** (0.005)
log(travel time):+2000km		-0.203^{***} (0.043)		-0.388** (0.185)		-0.032*** (0.005)
N obs. effective	4,703,010	4,703,010	16,412	16,412	10,106,940	10,106,940
R2	0.88	0.88	0.86	0.86	0.70	0.70

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 16: Elasticity of citations to travel time: IV estimation intensive and extensive margin

Column (1) shows the result of Instrumental Variables Poisson estimation of $citations_{FiGjhkt} = \exp\left[\beta\, log(\text{travel time}_{ijt}) + \lambda\, \hat{u}_{FiGjhkt} + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time_{ijt} is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. The variable $\hat{u}_{FiGjhkt}$ is constructed as $\hat{u}_{FiGjhkt} = \text{travel time}_{FiGjhkt} - \hat{\lambda}_2\, \text{travel time}^{\text{fix network}}_{FiGjhkt}$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Column (3) shows the result of an IV-2SLS estimation of $log(citations_{FiGjhkt}) = \alpha\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with a sample of establishment-technology pairs ($FiGjhk$) that have positive citations in all periods. Column (5) shows the result of an IV-2SLS estimation of $\mathbb{1}\{citations_{FiGjhkt} > 0\} = \gamma\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \varepsilon_{FiGjhkt}$, with the same sample as (1). Columns (3) and (5) use $\text{travel time}^{\text{fix network}}_{ijt}$ as an instrument for travel time_{ijt}. Column (2), (4) and (6) open, respectively, the coefficients β, α, γ by distance between the citing establishment Fi and the cited establishment Gj. Standard errors are presented in parenthesis. In Columns (1) and (2) standard errors are bootstrapped. In Columns (3) to (6) standard errors clustered at the non-directional location pair (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	Concentration citing	Concentration cited	Cited lag forward	Citing lag backward	Citing govnt & uni	Cited govnt & univ	Self citation
Dep. variable: *citations*				$cit_{FiGjhkt}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(travel time):0-300km	0.103	0.160	-0.045	0.1907	0.021	0.018	0.002
	(0.121)	(0.114)	(0.472)	(0.538)	(0.038)	(0.038)	(0.039)
log(travel time):300-1000km	-0.105	-0.039	-0.546	-0.145	-0.102^{***}	-0.099^{***}	-0.077^{***}
	(0.084)	(0.095)	(0.364)	(0.366)	(0.027)	(0.027)	(0.029)
log(travel time):1000-2000km	-0.138	-0.117	0.086	0.101	-0.094^{**}	-0.093^{**}	-0.094^{**}
	(0.105)	(0.116)	(0.480)	(0.498)	(0.042)	(0.041)	(0.040)
log(travel time):+2000km	-0.287^{***}	-0.268^{***}	0.720^{**}	0.560	-0.185^{***}	-0.188^{***}	-0.153^{***}
	(0.105)	(0.090)	(0.344)	(0.472)	(0.049)	(0.048)	(0.040)
log(travel time):0-300km $\times X$	-1.180	-2.013	0.028	-0.066	-0.125	0.481	0.038
	(1.843)	(1.712)	(0.185)	(0.211)	(0.367)	(0.543)	(0.252)
log(travel time):300-1000km $\times X$	0.079	-0.880	0.178	0.018	-0.088	-0.609^{*}	0.077
	(1.188)	(1.366)	(0.144)	(0.145)	(0.265)	(0.330)	(0.127)
log(travel time):1000-2000km $\times X$	0.634	0.341	-0.073	-0.078	-0.282	-0.370	0.082
	(1.412)	(1.606)	(0.191)	(0.197)	(0.366)	(0.385)	(0.210)
log(travel time):+2000km $\times X$	1.436	1.157	-0.366^{***}	-0.299	-0.328	0.015	-0.073
	(1.456)	(1.136)	(0.137)	(0.188)	(0.410)	(0.295)	(0.170)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010	4,800,144	4,800,144	4,835,001
R2	0.88	0.88	0.88	0.88	0.88	0.88	0.94

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 18: Elasticity of citations to travel time: Heterogeneity (part 1)

Result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \mathbb{1}\{distance_{ij} \in d\}\, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \mathbb{1}\{distance_{ij} \in d\}\, \mathbb{1}\{X_{FiGjhkt}\}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0 - 300km]$, $(300km - 1000km]$, $(1000km - 2000km]$, $(2000km - max]$. The variable X takes different value depending on the column: in column (1) it is the across-MSA Herfindahl index of the citing technology, in column (2) it is the across-MSA Herfindahl index of the cited technology, in column (3) it is median forward citation lag of the cited technology, in column (4) it is median backward citation lag of the citing technology. In column (5) and (6) the sample includes government and university patents, in column (5) X is a dummy that takes value one if the citing patent belongs to a university or government organisation, in column (6) it is a dummy that takes value one if the cited patent belongs to a university or government organisation. In column (7) the sample includes self citations, the variable X is a dummy that takes value one if the citing firm F cited firm G are the same. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	Citing quantile	Cited quantile
Dep. variable: *citations*	$cit_{FiGjhkt}$	
	(1)	(2)
log(travel time) × quantile 0.00	-0.151***	-0.111***
	(0.058)	(0.039)
log(travel time) × quantile 0.01	-0.078	-0.084
	(0.114)	(0.101)
log(travel time) × quantile 0.25	-0.081	-0.159*
	(0.103)	(0.093)
log(travel time) × quantile 0.50	-0.139	-0.063
	(0.091)	(0.083)
log(travel time) × quantile 0.75	-0.262***	-0.033
	(0.079)	(0.068)
log(travel time) × quantile 0.90	-0.029	-0.127**
	(0.066)	(0.057)
log(travel time) × quantile 0.95	-0.001	-0.123***
	(0.037)	(0.038)
log(travel time) × quantile 0.99	-0.130***	-0.066*
	(0.035)	(0.039)
log(travel time) × quantile 0.999	-0.070	-0.070
	(0.045)	(0.045)
N obs. effective	4,703,010	4,703,010
R2	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 17: Elasticity of citations to travel time: Heterogeneity (part 2)

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_q \boldsymbol{\beta}_q \, log(\text{travel time}_{ijt}) \, \mathbb{1}\{quantile_{Fh} \in q\} + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. $quantile_{Fh}$ is the quantile of firm F in the distribution of firms within technology h, using patents applied by F in h in the time period 1949-1953. Column (2) repeats the analysis using the quantile of the cited firm G in technology k. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location in parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

	PPML	
	Citing technology	Cited technology
Dep. variable: *citations*	$cit_{FiGjhkt}$	
	(1)	(2)
log(travel time) × Chemical	-0.066 (0.045)	-0.093^{**} (0.045)
log(travel time) × Computers & Communications	-0.100 (0.079)	-0.140^{*} (0.077)
log(travel time) × Drugs & Medical	-0.053 (0.162)	-0.005 (0.181)
log(travel time) × Electrical & Electronic	-0.070 (0.048)	-0.054 (0.046)
log(travel time) × Mechanical	-0.080^{**} (0.031)	-0.087^{***} (0.032)
log(travel time) × Others	-0.147^{***} (0.045)	-0.113^{**} (0.044)
N obs. effective	4,703,010	4,703,010
R2	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 19: Elasticity of citations to travel time by citing and cited technology
Part 1

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_{tech} \beta_h \, \mathbb{1}\{tech = h\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G located in j, in technology k. $\mathbb{1}\{tech = h\}$ is a dummy variable that takes value 1 when the citing technology h is equal to technology $tech$. In column (2) the dummy is modified to $\mathbb{1}\{tech = k\}$ such that it takes value 1 when the cited technology k is equal to technology $tech$. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

Cited \ Citing	Chemical	Computers & Communications	Drugs & Medical	Electrical & Electronic	Mechanical	Others
Chemical	-0.092^{**} (0.052)	0.219 (0.262)	0.113 (0.199)	-0.299^{***} (0.094)	-0.025 (0.071)	-0.070 (0.068)
Computers & Communications	-0.089 (0.259)	-0.306^{***} (0.095)	-0.657 (0.976)	0.107 (0.090)	0.122 (0.149)	0.095 (0.169)
Drugs & Medical	0.224 (0.239)	0.567 (1.205)	-0.278 (0.268)	-0.230 (0.561)	-0.334 (0.362)	0.358 (0.323)
Electrical & Electronic	0.233^{**} (0.093)	0.171^{*} (0.096)	-0.224 (0.634)	-0.102^{**} (0.056)	0.087 (0.070)	-0.063 (0.079)
Mechanical	-0.060 (0.076)	0.151 (0.145)	-0.152 (0.402)	0.106 (0.082)	-0.129^{***} (0.035)	-0.032 (0.056)
Others	0.042 (0.074)	0.173 (0.169)	0.204 (0.274)	0.052 (0.072)	0.019 (0.053)	-0.209^{***} (0.054)

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 21: Elasticity of citations to travel time by citing and cited technology
Part 2

Column (1) shows the result of one single Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_{\text{tech pair}} \beta_{hk}\, \mathbb{1}\{\text{tech pair} = hk\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G located in j, in technology k. $\mathbb{1}\{\text{tech pair} = hk\}$ is a dummy variable that takes value 1 when the citing technology h is equal to technology $tech$. In column (2) the dummy is modified to $\mathbb{1}\{tech = k\}$ such that it takes value 1 when the cited technology k is equal to technology $tech$. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values. The amount of observation in the effective sample is 4,703,010.

Sample of establishments

During the time period there was entry and exit of research establishments that was not uniform across locations. We may then think that the change in diffusion of knowledge is only consequence of the change in the geographical location of innovation. To test this possibility, in Table 22 we estimate the baseline regression 3 with different samples. In column (1) we include the baseline results.[102] In column (2) we use only citing establishments Fi that filed patents during the initial time period 1949-1953. In column (3) we further restrict the sample to both citing

[102] Coefficients are not bias corrected.

establishments Fi and cited establishments Gj that filed patents in 1949-1953.[103] We find that the coefficient at more than 2,000km remains comparable to the one in the baseline regression, statistically significant at the 1%.

	All	Citing establishment	Citing & Cited establishment
Dep. variable: *citations*		$cit_{FiGjhkt}$	
	(1)	(2)	(3)
log(travel time) × 0-300km	0.021	0.020	0.028
	(0.039)	(0.043)	(0.043)
log(travel time) × 300-1,000km	-0.099^{***}	-0.095^{***}	-0.095^{***}
	(0.027)	(0.029)	(0.030)
log(travel time) × 1,000-2,000km	-0.093^{**}	-0.092^{**}	-0.062
	(0.042)	(0.047)	(0.050)
log(travel time) × +2,000km	-0.185^{***}	-0.155^{***}	-0.179^{***}
	(0.049)	(0.052)	(0.052)
N obs. effective	4,703,010	3,109,285	1,960,851
R2	0.88	0.88	0.89

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 22: Elasticity of citations to travel time:
Fix sample of establishments

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \times \mathbb{1}\{distance_{ij} \in d\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0 - 300km], (300km - 1000km], (1000km - 2000km], (2000km - max]$. Column (2) truncates the sample keeping only citing establishments Fi that where present in the initial time period $1949 - 1953$. Column (3) truncates the sample keeping only citing establishments Fi and cited establishments Gj that where present in the initial time period. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location pair are presented in parenthesis (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

Ticket prices

During the period of analysis ticket prices were set by the Civil Aeronautics Board, so airlines could not set prices of their own tickets. Some airlines included

[103]We require Fi and Gj to have positive amount of patents applied during 1949-1953. However, those establishments need not to have cited each other.

a sample of prices in the last page of their booklet of flight schedules a sample of prices, which we digitized. We document multiple facts about prices.

First, prices were set in the form of an intercept plus a variable increment depending on distance between origin and destination. Second, all airlines operating within the same route charged exactly the same price.

Third, ticket prices of flights operated by jet airplanes had a surcharge of around 6% on top of the one operated by propeller airplanes.

Fourth, prices were relatively constant over time (with a growth rate approximately equal to the one of the consumer price index) until 1962-1963, years in which we observe a drop in prices of around 20% for routes of more than 1,000km distance, breaking the linearity of prices on distance previously observed.

	PPML							
Dep. variable: *citations*	$cit_{FiGjhkt}$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(travel time) × 0-300km	0.0213	0.0276	0.0198	0.0318	0.0252	0.0349	0.0283	0.0313
	(0.0388)	(0.0385)	(0.0391)	(0.0393)	(0.0389)	(0.0391)	(0.0396)	(0.0393)
log(travel time) × 300-1,000km	-0.0990***	-0.1040***	-0.0935***	-0.0745**	-0.1014***	-0.0857***	-0.0748**	-0.0861***
	(0.0269)	(0.0292)	(0.0265)	(0.0303)	(0.0290)	(0.0312)	(0.0303)	(0.0312)
log(travel time) × 1000-2,000km	-0.0928**	-0.1155**	-0.0710*	-0.0395	-0.0948*	-0.0498	-0.0318	-0.0435
	(0.0418)	(0.0485)	(0.0423)	(0.0523)	(0.0502)	(0.0573)	(0.0520)	(0.0576)
log(travel time) × +2,000km	-0.1848***	-0.1761***	-0.1724***	-0.1238**	-0.1658***	-0.1052*	-0.1236**	-0.1041*
	(0.0492)	(0.0531)	(0.0498)	(0.0587)	(0.0542)	(0.0607)	(0.0590)	(0.0609)
log(highway time) × 0-300km		-0.1306			-0.1060	-0.0422		-0.0374
		(0.1210)			(0.1231)	(0.1415)		(0.1426)
log(highway time) × 300-1,000km		0.0020			0.0309	0.0808		0.0867
		(0.1134)			(0.1137)	(0.1495)		(0.1491)
log(highway time) × 1,000-2,000km		0.0530			0.0695	0.0578		0.0681
		(0.1017)			(0.1090)	(0.1569)		(0.1582)
log(highway time) × +2,000km		-0.0650			-0.0486	-0.0712		-0.0707
		(0.1134)			(0.1162)	(0.1780)		(0.1779)
log(mean share telephone) × year 1956			10.58**		10.43**		4.855	4.811
			(4.689)		(4.671)		(4.587)	(4.584)
log(mean share telephone) × year 1961			13.47**		13.13**		7.539	7.471
			(6.243)		(6.251)		(6.066)	(6.085)
log(mean share telephone) × year 1966			16.39**		16.50**		12.02*	12.23*
			(6.761)		(6.752)		(6.686)	(6.691)
log(distance) × year 1956				0.0119***		0.0119***	0.0111***	0.0111***
				(0.0025)		(0.0025)	(0.0026)	(0.0026)
log(distance) × year 1961				0.0144***		0.0147***	0.0133***	0.0136***
				(0.0044)		(0.0044)	(0.0045)	(0.0044)
log(distance) × year 1966				0.0131**		0.0137**	0.0112**	0.0120*
				(0.0054)		(0.0067)	(0.0055)	(0.0068)
N obs. effective	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010	4,703,010
R2	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 23: Elasticity of citations to travel time: additional controls

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \, \mathbb{1}\{distance_{ij} \in d\} \, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \, \mathbb{1}\{distance_{ij} \in d\} \, \mathbb{1}\{X_{FiGjhkt}\} \, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, for citations of patents filed by establishment of firm F in location i, technology h and time period t, to patents filed by establishment of firm G in location j and technology k. travel time$_{ijt}$ is the travel time in minutes between location i and j at time period t, and it is set to 1 when $i = j$. d are distance intervals: $[0 - 300km]$, $(300km - 1000km]$, $(1000km - 2000km]$, $(2000km - max]$. Relative to (1), columns (2) to (8) contain additional controls. Log highway time between i and j changes in every time period t. The log mean share of households with telephone line in ij pair interacted in 1960 is interacted with a time dummy. Log distance ij is interacted with a time dummy. When $FiGjhk$ has positive citations in at least one period and no citations in another, we attribute zero citations in the missing period. Standard errors clustered at the non-directional location in parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

Highway travel time

C.2. Creation of knowledge

C.2.1. Heterogeneous effects

C.2.2. Robustness

	Baseline	Quartile absolute	Quartile per capita
Dependent Variable: *Patents*		$Patents_{Fiht}$	
	(1)	(2)	(3)
log(knowledge access)	10.14^{***} (3.66)	9.36^{**} (3.69)	7.77^{**} (3.70)
log(knowledge access) × quartile 0.50		2.05^{***} (0.58)	0.75^{**} (0.34)
log(knowledge access) × quartile 0.25		3.80^{***} (0.90)	1.58^{***} (0.50)
log(knowledge access) × quartile 0.00		5.00^{***} (1.30)	4.03^{***} (0.77)
N obs. effective	991,480	991,480	991,480
R2	0.85	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 24: Elasticity of new patents to knowledge access: absolute and per capita MSA innovativeness

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed within technology using the absolute level of patents in the MSA-technology in 1949-1953. Column (3) computes the quartile of innovativeness using patents per capita in the MSA-technology in 1949-1953 using 1950 population. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

C.3. Mechanism

	PPML		β by distance		+300km		+1,000km		+2,000km	
Dependent Variable: *Patents*					$Patents_{Fiht}$					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(knowledge access)	10.14***	9.36**	18.17***	16.50**	10.09**	8.70*	18.82***	19.08***	12.70	10.26
	(3.66)	(3.69)	(4.63)	(4.76)	(4.66)	(4.67)	(5.82)	(5.74)	(8.18)	(7.92)
log(knowledge access) × quartile 0.50		2.05***		2.70***		2.12***		2.08***		1.94***
		(0.58)		(0.84)		(0.58)		(0.53)		(0.49)
log(knowledge access) × quartile 0.25		3.80***		5.96***		4.19***		3.97***		3.64***
		(0.90)		(1.42)		(0.88)		(0.81)		(0.73)
log(knowledge access) × quartile 0.00		5.00***		8.94***		5.49***		5.28***		4.68***
		(1.30)		(1.97)		(1.25)		(1.23)		(1.07)
N obs. effective	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480
R2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 25: Elasticity of new patents to knowledge access, varying beta or distance.

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) opens the coefficient ρ by the quartile of innovativeness of location i within technology h, computed using patents in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Relative to columns (1) and (2), columns (3) and (4) compute Knowledge Access using four distance-specific β parameter according to distance bins between i and j. The bins are [0km, 300km], (300km, 1000km], (1000km, 2000km], +2,000km. Columns (5) to (10) use the same β as column (1) and (2), but computing Knowledge Access with a truncated sample of j that are further than a certain distance threshold from i. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

	PPML		OLS	
Dependent Variable: *Patents*	$Patents_{Fiht}$		$log(Patents_{Fiht})$	
	(1)	(2)	(3)	(4)
log(knowledge access)	10.14^{***} (3.66)	9.36^{**} (3.69)	6.83^{*} (3.19)	6.27^{*} (3.20)
log(knowledge access) × quartile 0.50		2.05^{***} (0.58)		0.92^{*} (0.51)
log(knowledge access) × quartile 0.25		3.80^{***} (0.90)		2.64^{**} (1.03)
log(knowledge access) × quartile 0.00		5.00^{***} (1.30)		3.82^{**} (1.79)
N obs. effective	991,480	991,480	300,539	300,539
R2	0.85	0.85	0.87	0.87

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 26: Elasticity of new patents to knowledge access: PPML and OLS

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (3) estimates $log(\text{Patents})_{Fiht} = \boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht} + \xi_{Fiht}$. Columns (2) and (4) open the coefficient ρ by the quartile of innovativeness of location i within technology h, computed within technology using the absolute level of patents in the MSA-technology in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Difference in amount of observations is due to dropping zeros in columns (3) and (4). Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

Dependent Variable: *Patents*	$Patents_{Fiht}$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(knowledge access)	10.14^{***}					9.96^{**}	11.29^{***}	10.67^{**}	12.90^{***}
	(3.66)					(4.50)	(4.32)	(4.70)	(4.43)
log(finance access hq)		0.54^{**}				0.02			
		(0.26)				(0.30)			
log(finance access hq rel)			0.40				-0.14		
			(0.25)				(0.28)		
log(finance access est)				0.56^{*}				-0.07	
				(0.31)				(0.39)	
log(finance access est rel)					0.31				-0.39
					(0.30)				(0.38)
N obs. effective	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480	991,480
R2	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 27: Elasticity of new patents to knowledge access and finance access

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. Column (2) to (5) use as regressor the finance access of establishments in location i technology h and time period t, where the measure of finance access changes across columns. Columns (6) to (9) estimate the regression using both knowledge access and finance access. Standard errors clustered at the location-technology ih are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

	IV-PPML	
Dependent Variable:	$Patents_{Fiht}$	
	(1)	(2)
log(knowledge access)	11.24*	10.26
	(6.35)	(6.383)
log(knowledge access) × quartile 0.50		2.317***
		(0.6554)
log(knowledge access) × quartile 0.25		4.212***
		(0.8381)
log(knowledge access) × quartile 0.00		5.770***
		(1.108)
residual	-2.31	-2.249
	(7.20)	(7.268)
residual × quartile 0.50		-2.553
		(1.594)
residual × quartile 0.25		-4.341**
		(1.972)
residual × quartile 0.00		-8.267**
		(3.277)
N obs. effective	991,480	991,480
R2	0.85	0.85

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.10$

Table 28: Elasticity of new patents to knowledge access: IV-PPML

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + \lambda\, \hat{u}_{Fiht} + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, for patents filed by establishment of firm F in location i, technology h and time period t. KA_{iht} is knowledge access of establishments in location i technology h and time period t. $\hat{u}_{Fiht}$ is the estimated residual of $log(KA_{Fiht}) = \lambda_2 log(\widetilde{KA}_{Fiht}) + u_{Fiht}$, where the subindex F in KA_{Fiht} is used to denote that there are multiple observations per iht. Column (2) open the coefficient ρ and λ by the quartile of innovativeness of location i within technology h, computed within technology using the absolute level of patents in the MSA-technology in 1949-1953. Higher quartile indicates higher initial level of innovativeness. The fourth quartile is used as reference category. Bootstrap standard errors are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

β	ρ	$\beta \times \rho$	Predicted yearly growth p.p.	Share yearly growth explained	Predicted yearly growth differential p.p.	Share yearly growth differential explained
-0.186	10.14	-1.89	3.47	0.78	1.1	0.21
-0.1	19.35	-1.94	3.5	0.78	1.07	0.2
-0.2	9.4	-1.88	3.47	0.78	1.1	0.21
-0.3	6.1	-1.83	3.45	0.77	1.14	0.22
-0.4	4.48	-1.79	3.44	0.77	1.16	0.22
-0.5	3.52	-1.76	3.44	0.77	1.19	0.23
-0.6	2.91	-1.74	3.45	0.77	1.2	0.23
-0.7	2.48	-1.73	3.47	0.78	1.22	0.23
-0.8	2.17	-1.73	3.5	0.78	1.22	0.23
-0.9	1.93	-1.73	3.52	0.79	1.24	0.24
-1	1.72	-1.72	3.51	0.79	1.28	0.24
-2	0.58	-1.16	2.8	0.63	1.55	0.3
-5	0.04	-0.19	1.19	0.27	3.65	0.7
-8	0.09	-0.76	8.22	1.84	6.96	1.33
-10	0.11	-1.08	15.16	3.4	8.19	1.56
-20	0.13	-2.63	69.8	15.65	21.66	4.14
-50	0.16	-8.22	531.34	119.16	219.49	41.94
-100	0.12	-12.33	5428.85	1217.49	2971.74	567.91

Table 29: Effect of knowledge access on new patents: varying the value of elasticity of knowledge diffusion