ONLINE APPENDIX
The Creation and Diffusion of Knowledge: Evidence from the Jet Age

Stefan Pauly* Fernando Stipanicic†

December 23, 2024

A. 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:

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

*Email: *paulystefan@gmail.com*
†University of Oslo. Corresponding author, email: *fernando.stipanicic@econ.uio.no*

where New Knowledge_{Fi} is the knowledge created by firm F located in i. The 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.[1] The concept of knowledge spillovers goes back at least to **?** who explains it as one of the agglomeration forces. **?** 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

[1]The chapters of **?** 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.

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.

This theoretical framework bears resemblance to the concept of *Market Access* presented in Donaldson and Hornbeck (2016) and **?**. 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.

One assumption of the theoretical framework is that New Knowledge$_{Fi}$ is multiplicative-separable on Knowledge Access$_i$.[2] To the extent that firm's productivity z_{Fi} and choice of inputs $inputs_{Fi}$ are relatively time invariant, this assumption is not restrictive.[3] However, if for example $inputs_{Fi}$ changes with Knowledge Access$_i$, then the estimated value of the elasticity would be the sum of the direct effect of Knowledge Access$_i$ on New Knowledge$_{Fi}$ (ρ) and the indirect effect through changes in $f(\cdot)$.

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

[2]The implicit assumption is that $\frac{\partial log(\text{New Knowledge}_{Fi})}{\partial log(\text{Knowledge Access}_i)} = \frac{\partial log(f(z_{Fi}, inputs_{Fi}))}{\partial log(\text{Knowledge Access}_i)} + \rho = \rho$, meaning that $\frac{\partial log(f(z_{Fi}, inputs_{Fi}))}{\partial log(\text{Knowledge Access}_i)} = 0$.

[3]In the empirical analysis we will include a firm-location fixed effect Fi that would absorb time-invariant characteristics.

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 them. 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.

B. Historical context

B.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 speed of 500 kmh. By 1956, New York City and Los Angeles were connected with a non-stop flight in 8 hours and

[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

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 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.

B.2. Air transport: moving people, not goods

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

Figure 1 shows shipments in ton-miles for the period 1939 to 1964 by means of transportation: Airways, Pipelines, Inland Waterways, Motor, Railroads. Interestingly, we observe that air transport of goods, even if it increased, it accounted for less than

[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 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 the 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.

0.1% of transport of goods in 1964.

The massive increase in air transportation thus mainly affected the mobilility of people. To better understand changing travel patterns of people, in 1957, the US Census Bureau conducted its first-ever nationwide survey of travelers.[7] The results reveal the substantial importance of air travel for businesses at the time. 11% of all business trips rely on airplanes, while only 2% are bus and 6% railway trips, respectively.[8] Given the generally longer distances covered by airplanes, the share of air travel in terms of passenger miles is likely to be even higher. Of all flights recorded, more than 60% is due to business rather than leisure trips. In its early days, air travel was thus dominated by business travelers.

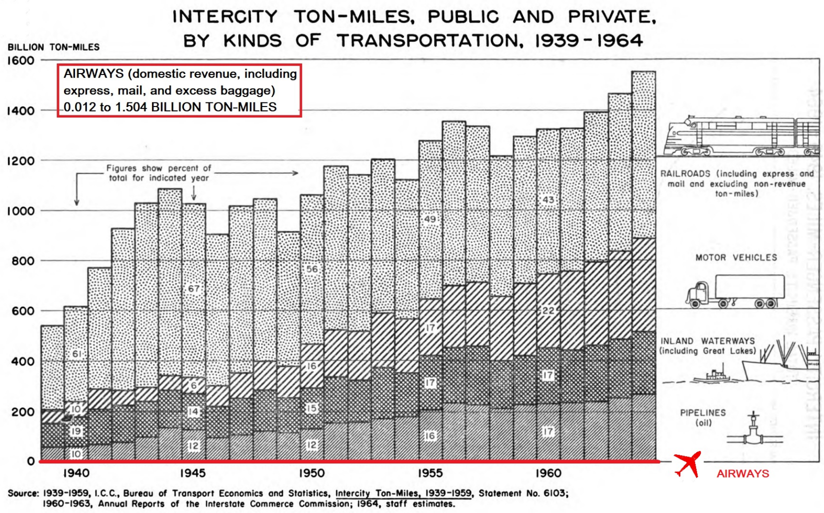


Figure 1: Ton Miles
Source: Interstate Commerce Commission, Annual Report 1965. Edited by the authors

[7] The resulting report can be accessed here: https://babel.hathitrust.org/cgi/pt?id=uc1.b5221822&seq=7

[8] A trip is considered as part of the sample if it required either on overnight stay or if the distance to the destination exceeded 100 miles.

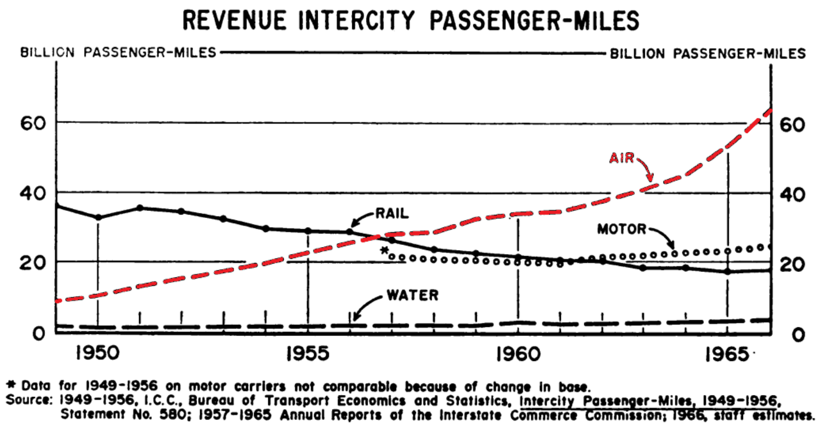


Figure 2: Passenger Miles
Source: Interstate Commerce Commission, Annual Report 1967. Edited by the authors

B.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.[9] It was empowered to control entry, fares, subsidies and mergers.[10] 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

[9] 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.

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

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.[11] Importantly, Borenstein and Rose (2014) in pages 68-69 explain that *"the regulatory route award process largely prevented airlines from reoptimizing their 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.

C. Travel Time Data

C.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).

[11] Borenstein and Rose (2014) in page 68, based on **?**, 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."*

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.

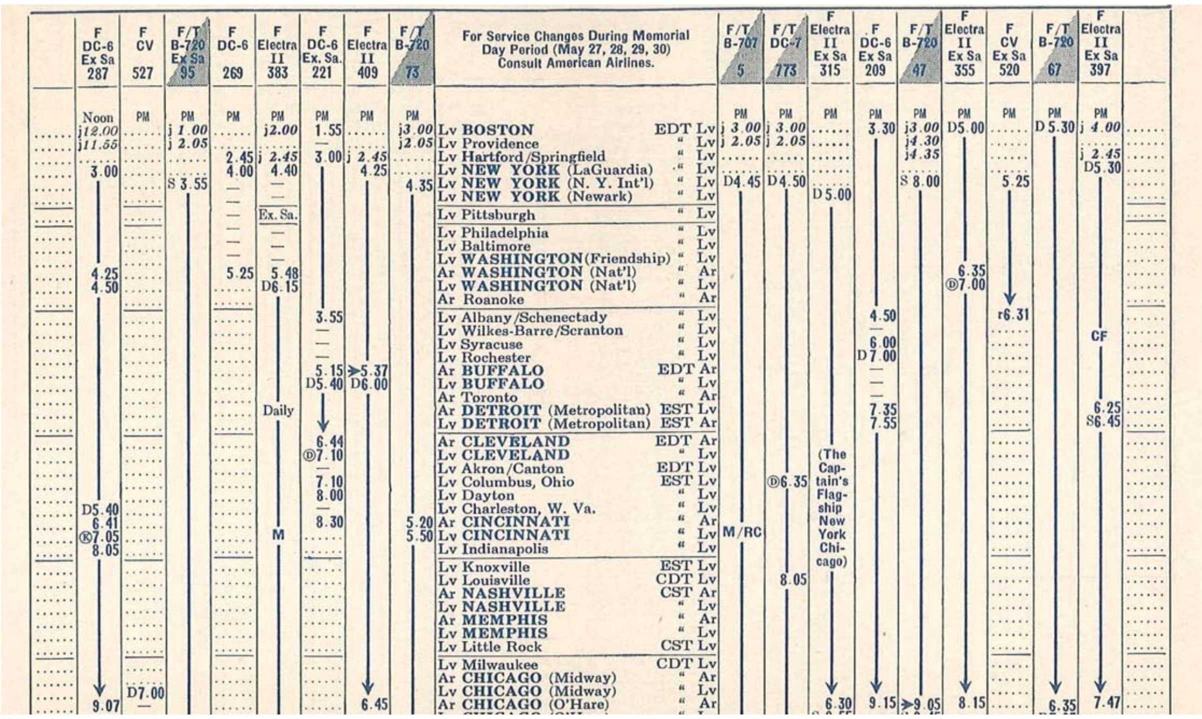
In total we obtain a sample of 5,910 flights. These flights often have multiple stops. If we count each origin-destination pair of these flights separately, our sample contains 17,469 legs.

Table 1 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 1: 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 3 shows a fragment of a page 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.



For Service Changes During Memorial Day Period (May 27, 28, 29, 30) Consult American Airlines.	F DC-6 Ex Sa 287	F CV 527	F/T B-720 Ex Sa 95	F DC-6 269	F Electra II 383	F DC-6 Ex. Sa. 221	F Electra II 409	F/T B-720 73
	Noon	PM	PM	PM	PM	PM	PM	PM
Lv BOSTON EDT Lv	j12.00		j 1.00		j2.00	1.55		j3.00
Lv Providence " Lv	j11.55		j 2.05			—		j2.05
Lv Hartford/Springfield " Lv				2.45	j 2.45	3.00	j 2.45	
Lv NEW YORK (LaGuardia) " Lv	3.00			4.00	4.40		4.25	
Lv NEW YORK (N. Y. Int'l) " Lv			S 3.55	—	—			4.35
Lv NEW YORK (Newark) " Lv				—	—			
Lv Pittsburgh " Lv				—	Ex. Sa.			
Lv Philadelphia " Lv				—	—			
Lv Baltimore " Lv				—	—			
Lv WASHINGTON (Friendship) " Lv				—	—			
Ar WASHINGTON (Nat'l) " Ar	4.25			5.25	5.48			
Lv WASHINGTON (Nat'l) " Lv	4.50				D6.15			
Ar Roanoke " Ar								
Lv Albany/Schenectady " Lv						3.55		
Lv Wilkes-Barre/Scranton " Lv						—		
Lv Syracuse " Lv						—		
Lv Rochester " Lv						—		
Ar BUFFALO EDT Ar						5.15	➤5.37	
Lv BUFFALO " Lv						D5.40	D6.00	
Ar Toronto " Ar								
Ar DETROIT (Metropolitan) EST Lv					Daily			
Lv DETROIT (Metropolitan) EST Ar								
Ar CLEVELAND EDT Ar						6.44		
Lv CLEVELAND " Lv						ⓓ7.10		
Lv Akron/Canton EDT Lv						—		
Lv Columbus, Ohio EST Lv						7.10		
Lv Dayton " Lv						8.00		
Lv Charleston, W. Va. " Lv	D5.40					—		
Ar CINCINNATI " Ar	6.41					8.30		5.20
Lv CINCINNATI " Lv	Ⓚ7.05				M			5.50
Lv Indianapolis " Lv	8.05							
Lv Knoxville EST Lv								
Lv Louisville CDT Lv								
Ar NASHVILLE CST Ar								
Lv NASHVILLE " Lv								
Ar MEMPHIS " Ar								
Lv MEMPHIS " Lv								
Lv Little Rock CST Lv								
Lv Milwaukee CDT Lv								
Ar CHICAGO (Midway) " Ar								
Lv CHICAGO (Midway) " Lv		D7.00						
Ar CHICAGO (O'Hare) " Ar	9.07	—					6.45	

For Service Changes During Memorial Day Period (May 27, 28, 29, 30) Consult American Airlines.	F/T B-707 5	F/T DC-7 773	F Electra II Ex Sa 315	F DC-6 Ex Sa 209	F/T B-720 47	F Electra II Ex Sa 355	F CV Ex Sa 520	F/T B-720 67	F Electra II Ex Sa 397
	PM	PM	PM	PM	PM	PM	PM	PM	PM
Lv BOSTON EDT Lv	j 3.00	j 3.00		3.30	j3.00	D5.00		D 5.30	j 4.00
Lv Providence " Lv	j 2.05	j 2.05			j4.30				
Lv Hartford/Springfield " Lv					j4.35				j 2.45
Lv NEW YORK (LaGuardia) " Lv									D5.30
Lv NEW YORK (N. Y. Int'l) " Lv	D4.45	D4.50			S 8.00		5.25		
Lv NEW YORK (Newark) " Lv			D5.00						
Lv Pittsburgh " Lv									
Lv Philadelphia " Lv									
Lv Baltimore " Lv									
Lv WASHINGTON (Friendship) " Lv									
Ar WASHINGTON (Nat'l) " Ar						6.35			
Lv WASHINGTON (Nat'l) " Lv						ⓓ7.00			
Ar Roanoke " Ar									
Lv Albany/Schenectady " Lv				4.50			r6.31		
Lv Wilkes-Barre/Scranton " Lv				—					CF
Lv Syracuse " Lv				6.00					
Lv Rochester " Lv				D7.00					
Ar BUFFALO EDT Ar				—					
Lv BUFFALO " Lv				—					
Ar Toronto " Ar				—					
Ar DETROIT (Metropolitan) EST Lv				7.35					6.25
Lv DETROIT (Metropolitan) EST Ar				7.55					S6.45
Ar CLEVELAND EDT Ar									
Lv CLEVELAND " Lv			(The Captain's Flagship New York Chicago)						
Lv Akron/Canton EDT Lv									
Lv Columbus, Ohio EST Lv		ⓓ6.35							
Lv Dayton " Lv									
Lv Charleston, W. Va. " Lv									
Ar CINCINNATI " Ar									
Lv CINCINNATI " Lv	M/RC								
Lv Indianapolis " Lv									
Lv Knoxville EST Lv									
Lv Louisville CDT Lv		8.05							
Ar NASHVILLE CST Ar									
Lv NASHVILLE " Lv									
Ar MEMPHIS " Ar									
Lv MEMPHIS " Lv									
Lv Little Rock CST Lv									
Lv Milwaukee CDT Lv									
Ar CHICAGO (Midway) " Ar									
Lv CHICAGO (Midway) " Lv									
Ar CHICAGO (O'Hare) " Ar			6.30	9.15	➤9.05	8.15		6.35	7.47

Figure 3: Fragment of 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.

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 3 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 2 lists all 15 cases.

Table 2: 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, 176 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 matched at least in one year to an airport ,and 108 MSAs are matched to one or more airports in the four years. In table 3 we present the 168 MSAs, the ones that are connected at least once, and the ones that are connected in the four years.

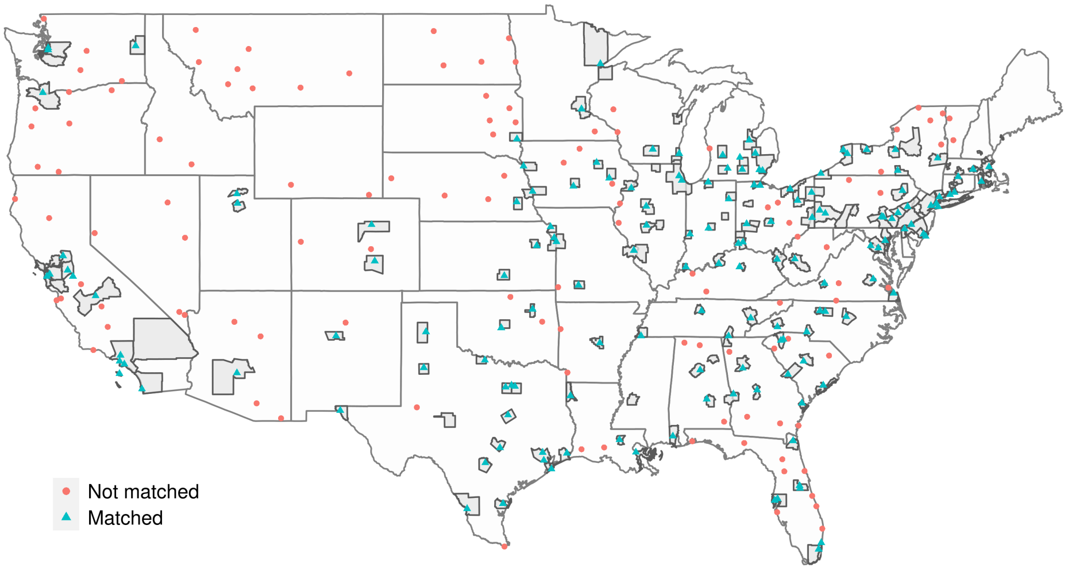


Figure 4: Airports matched to MSAs.

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 (**?**). 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.

C.2. Descriptive Statistics

Table 4 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 4: 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 5 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 5: 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 6: 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 5 shows all non-stop connections pooling all years in our data.

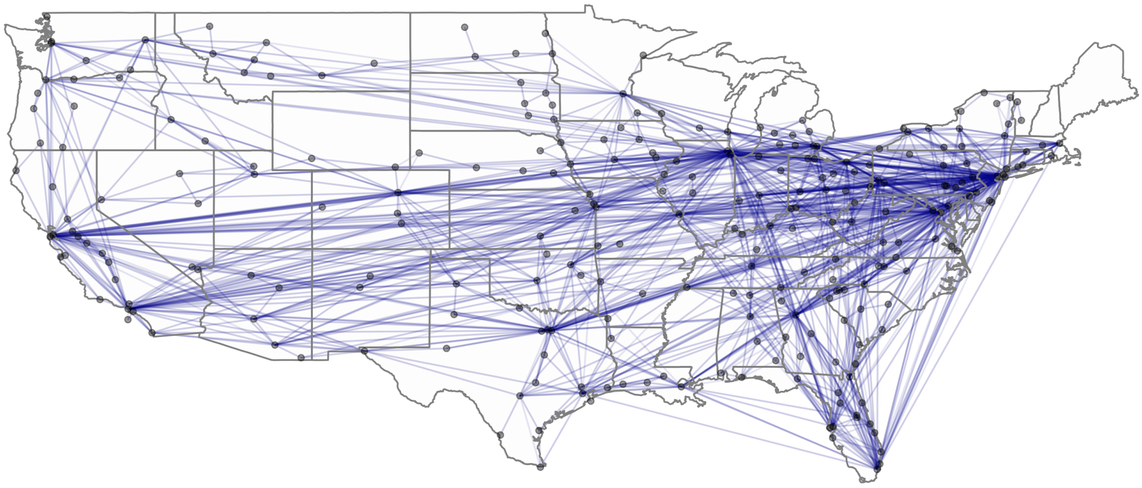


Figure 5: United States flight network 1951-1966

Figure 6 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 6 which illustrates that the overall number of MSA pairs with a direct connection increases over time.

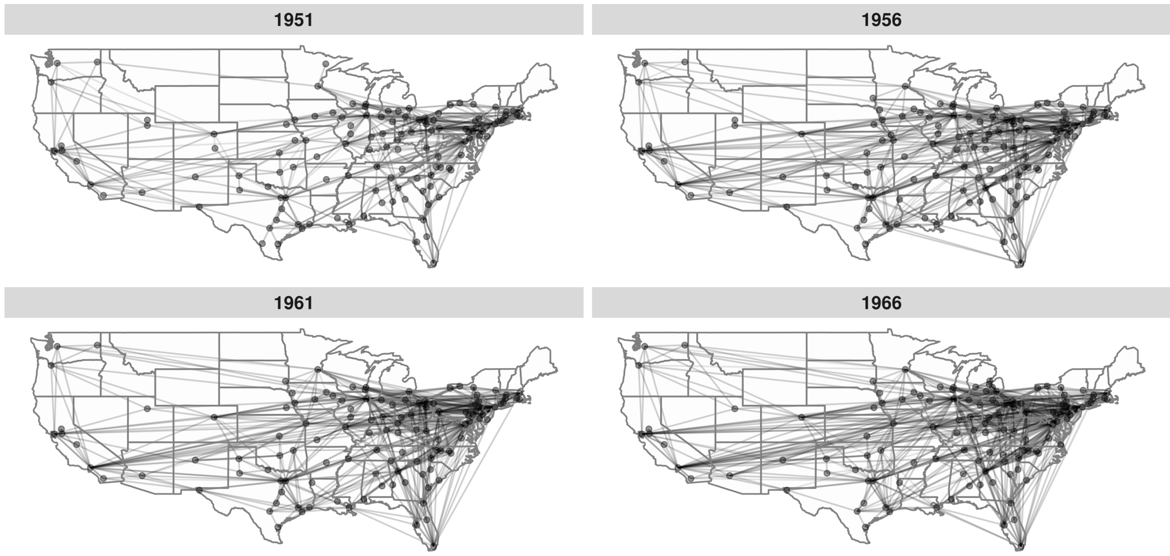


Figure 6: Flight Network by Year. Weighted by log weekly frequency.

Airlines differed in their speed of adoption of the newly arrived jet aircraft. Table 4 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 7 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 8 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.

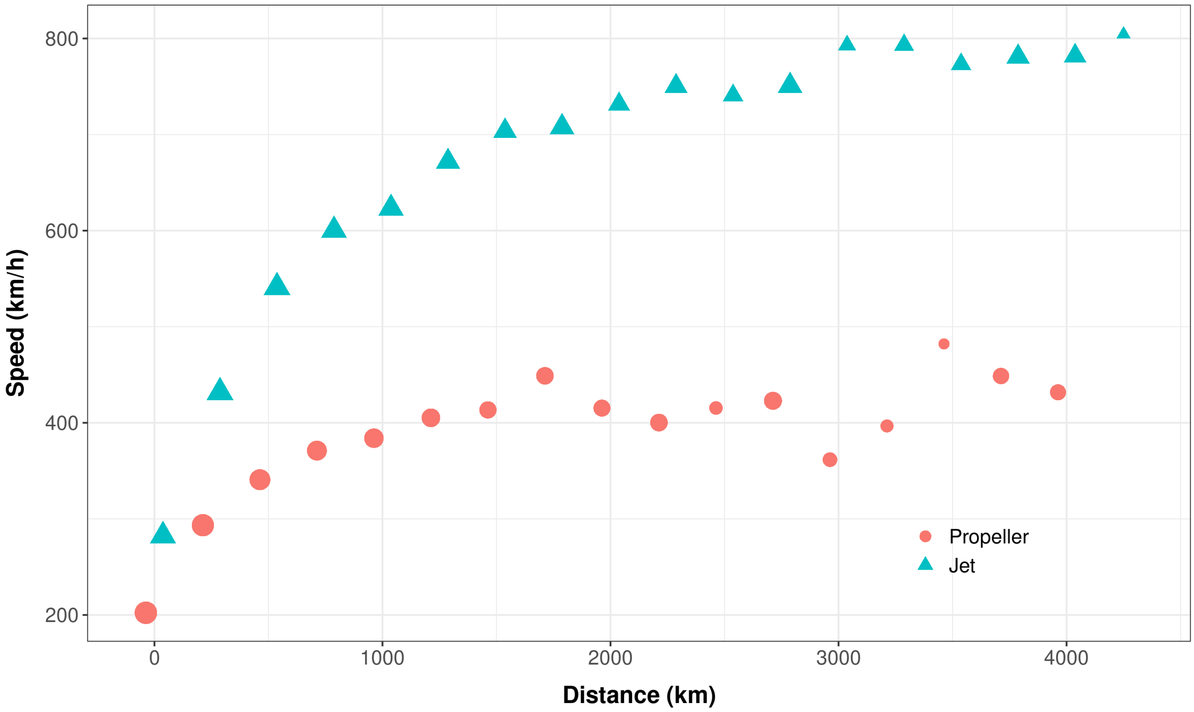


Figure 7: Speed by Aircraft Type. Pooling all Years.

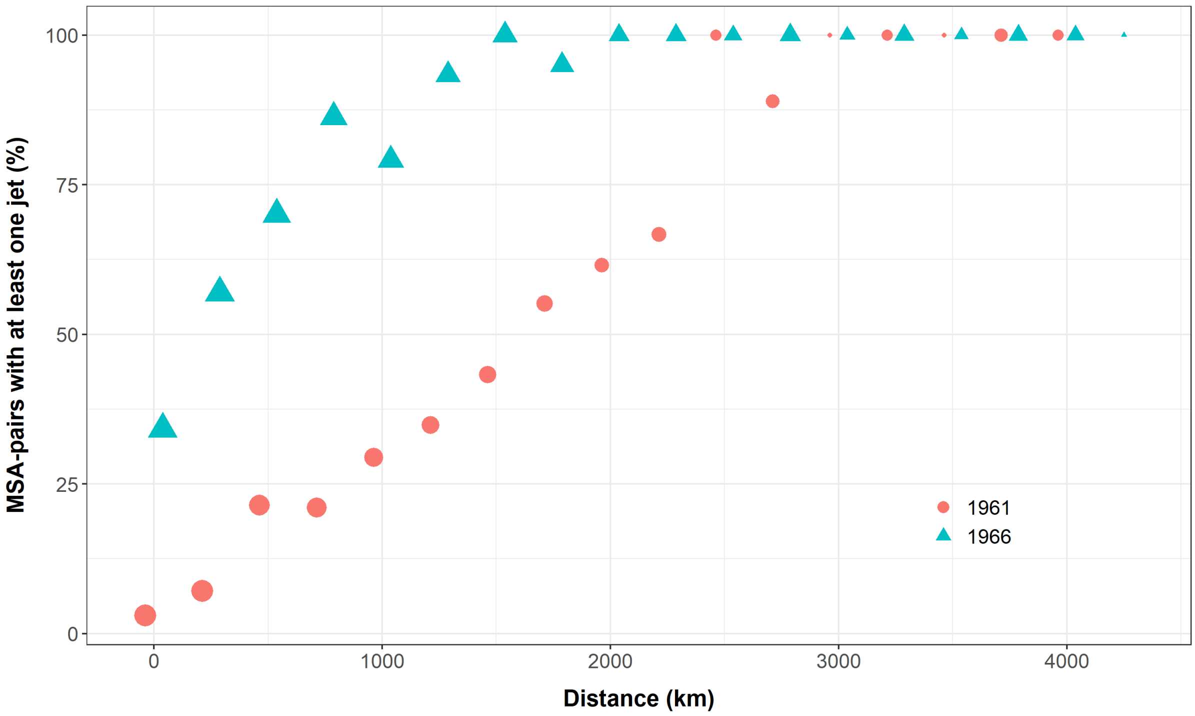


Figure 8: Jet Adoption

Figure 9 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.

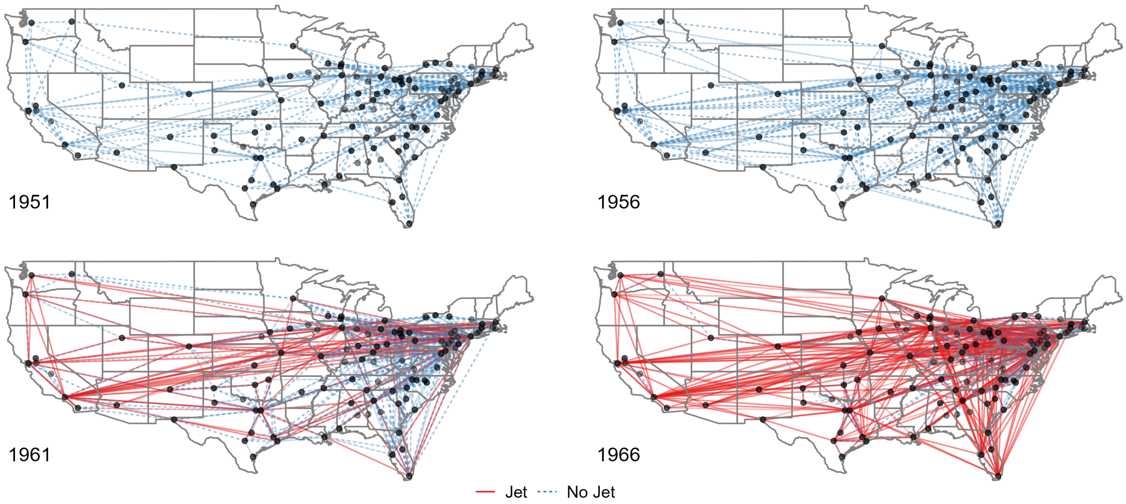


Figure 9: Jet Adoption by Year

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 11 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).

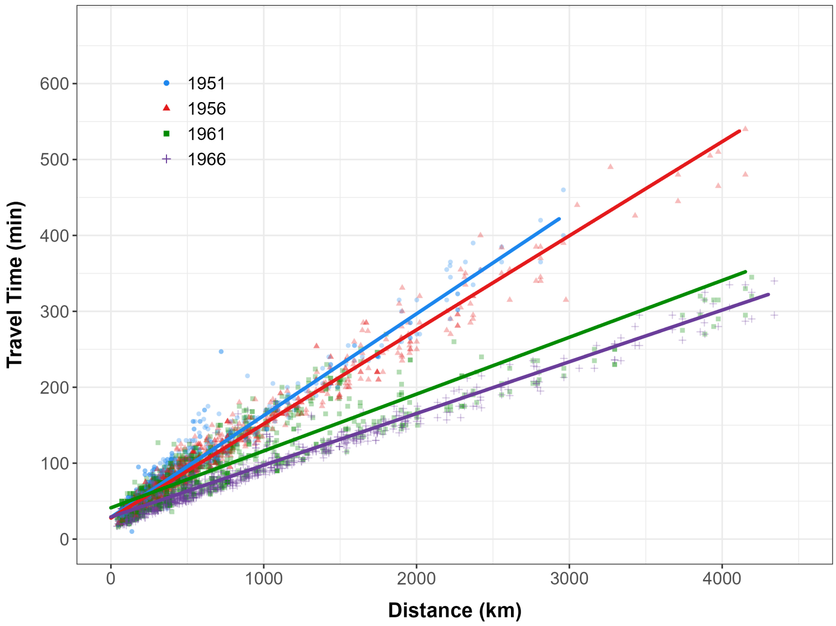


Figure 10: Non-stop fastest flights United States MSAs

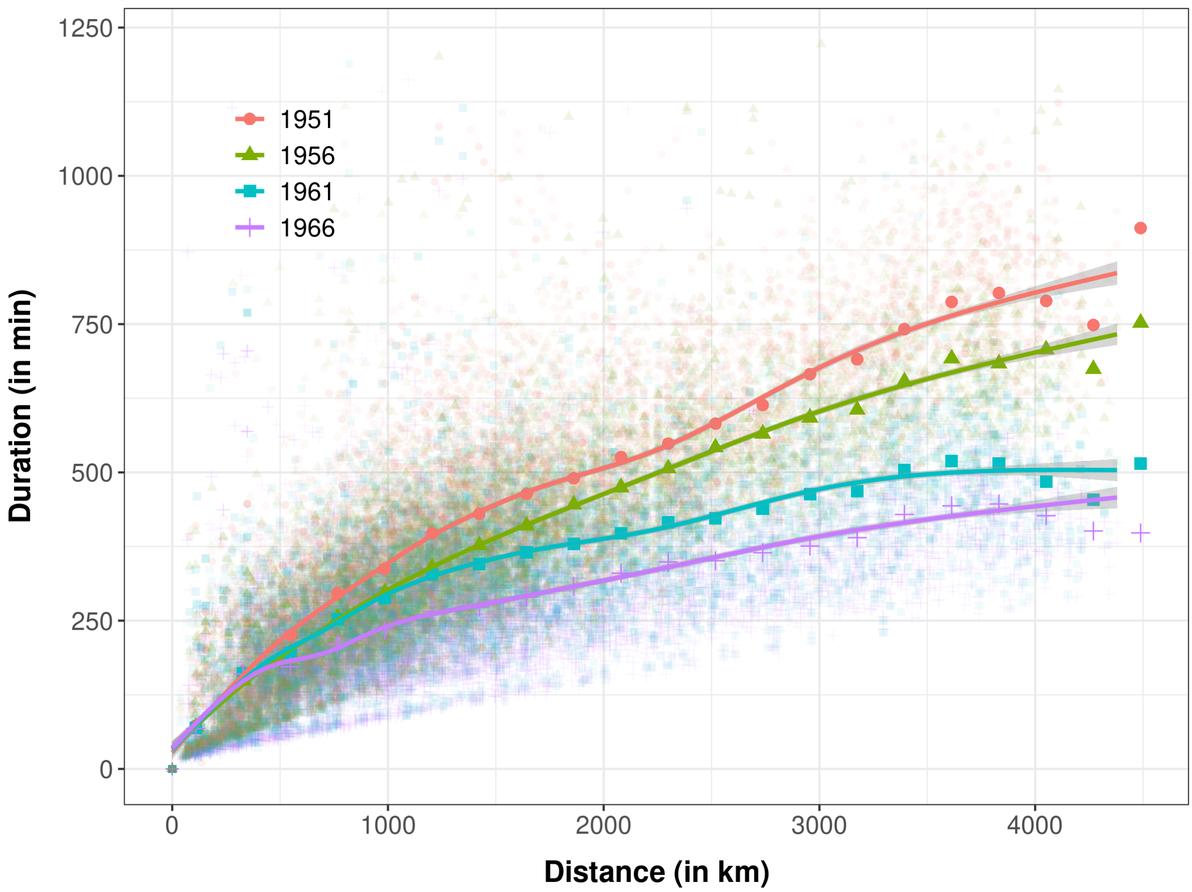


Figure 11: Travel Times between US MSAs.

Figure 12 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 13 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 15 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.

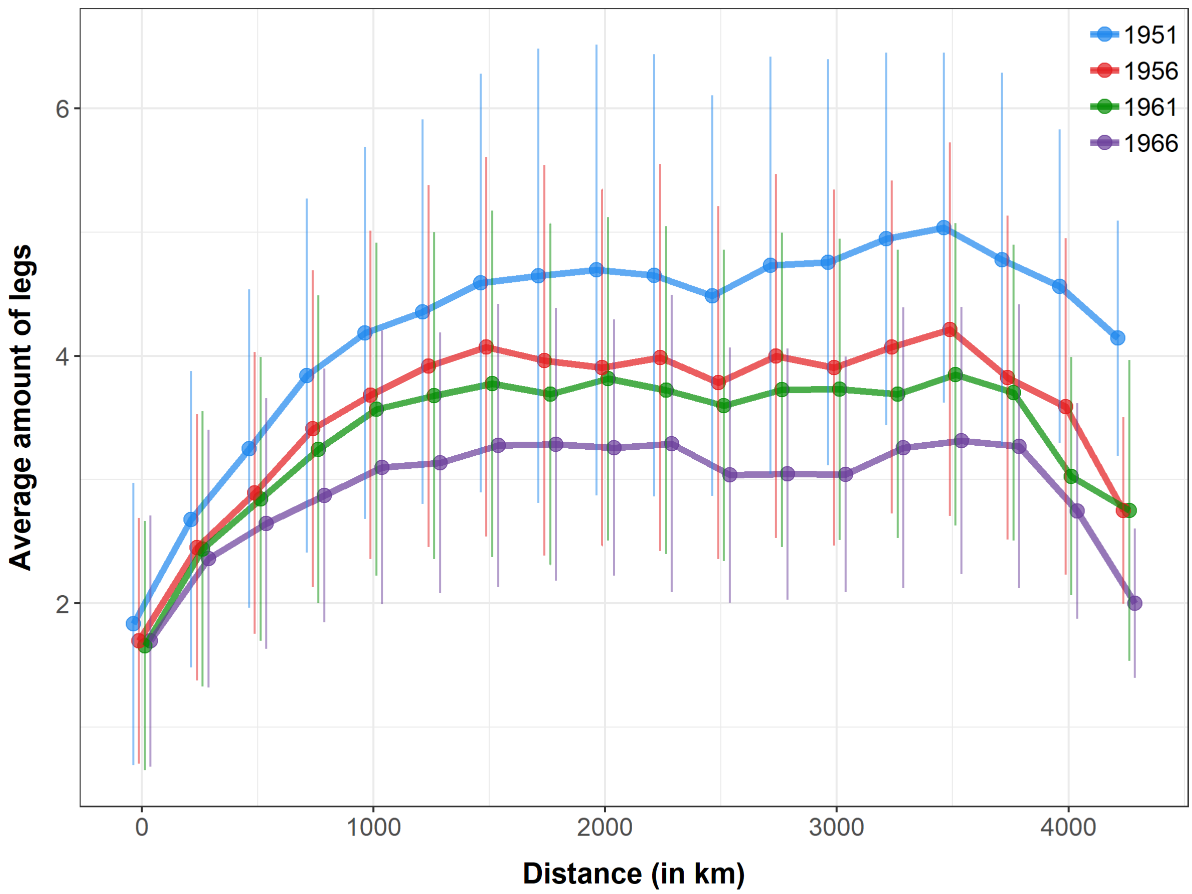


Figure 12: Average amount of legs per route

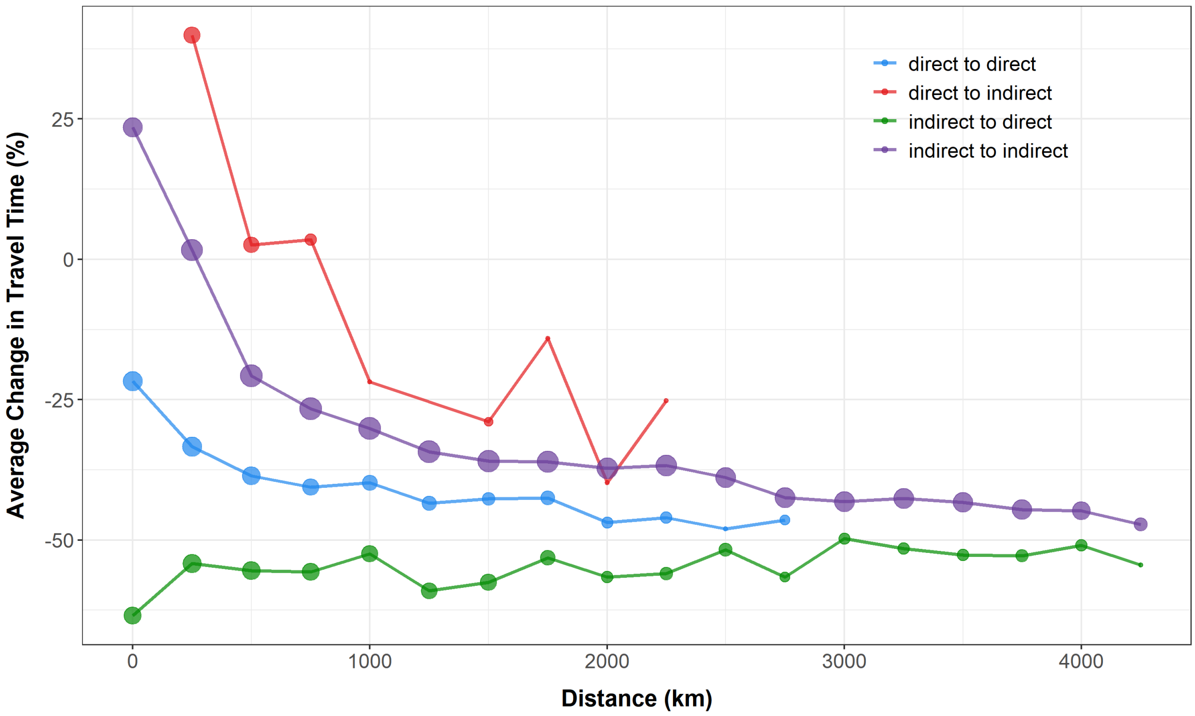


Figure 13: Change in US travel time 1951 to 1966: connections
12

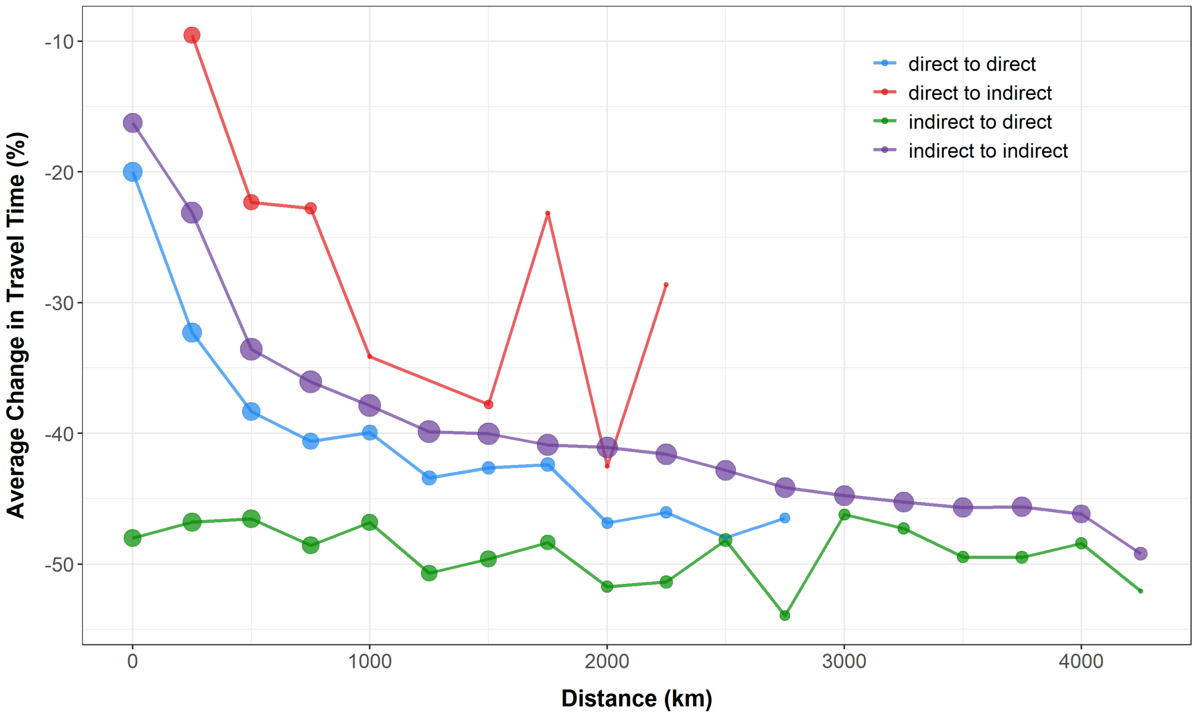


Figure 14: Change in US travel time 1951 to 1966: connections, discarding layover time
13

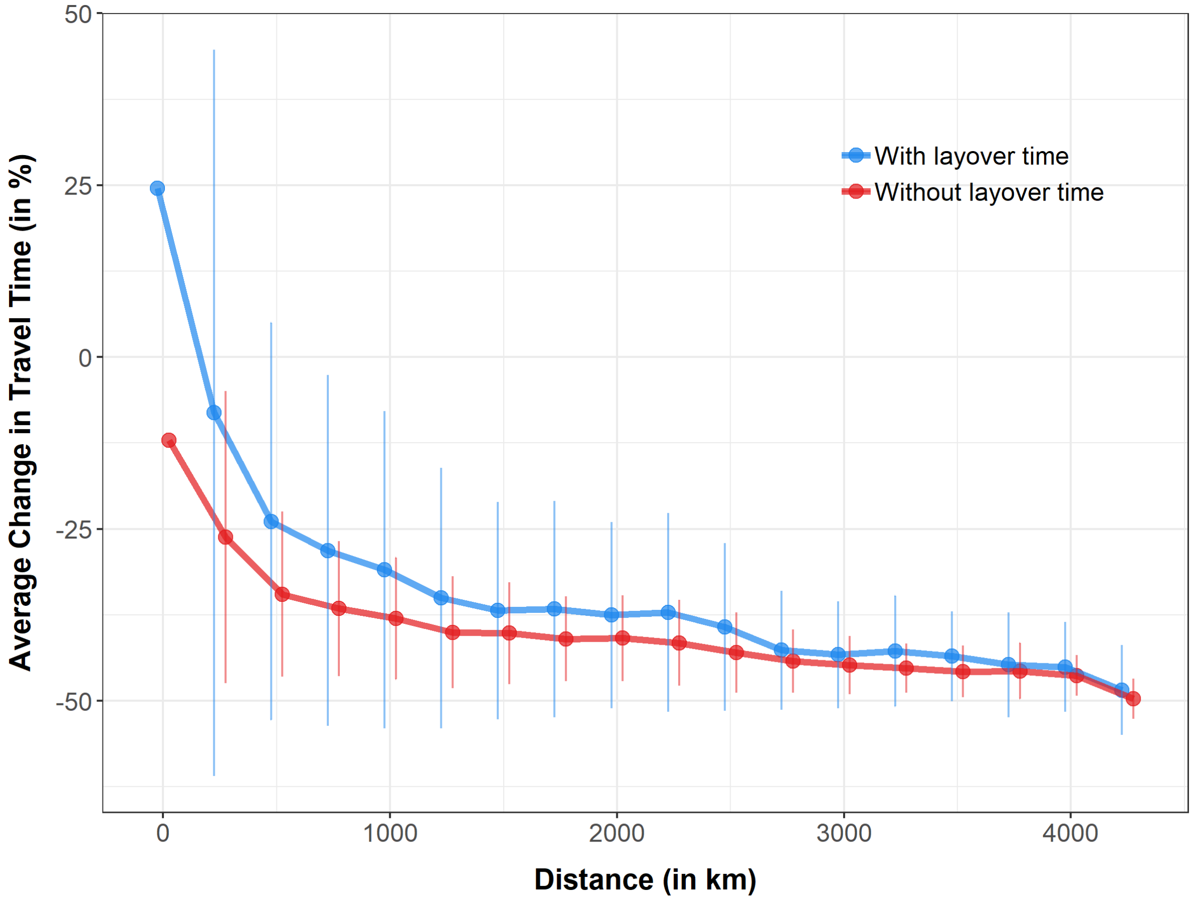


Figure 15: Change in US travel time 1951 to 1966: layover time

In figure 16 we show the average change in travel time in three counterfactual flight networks. The first counterfactual fixes the flight routes and allows aircraft speed to evolve.[14] 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 17 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 technology.

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

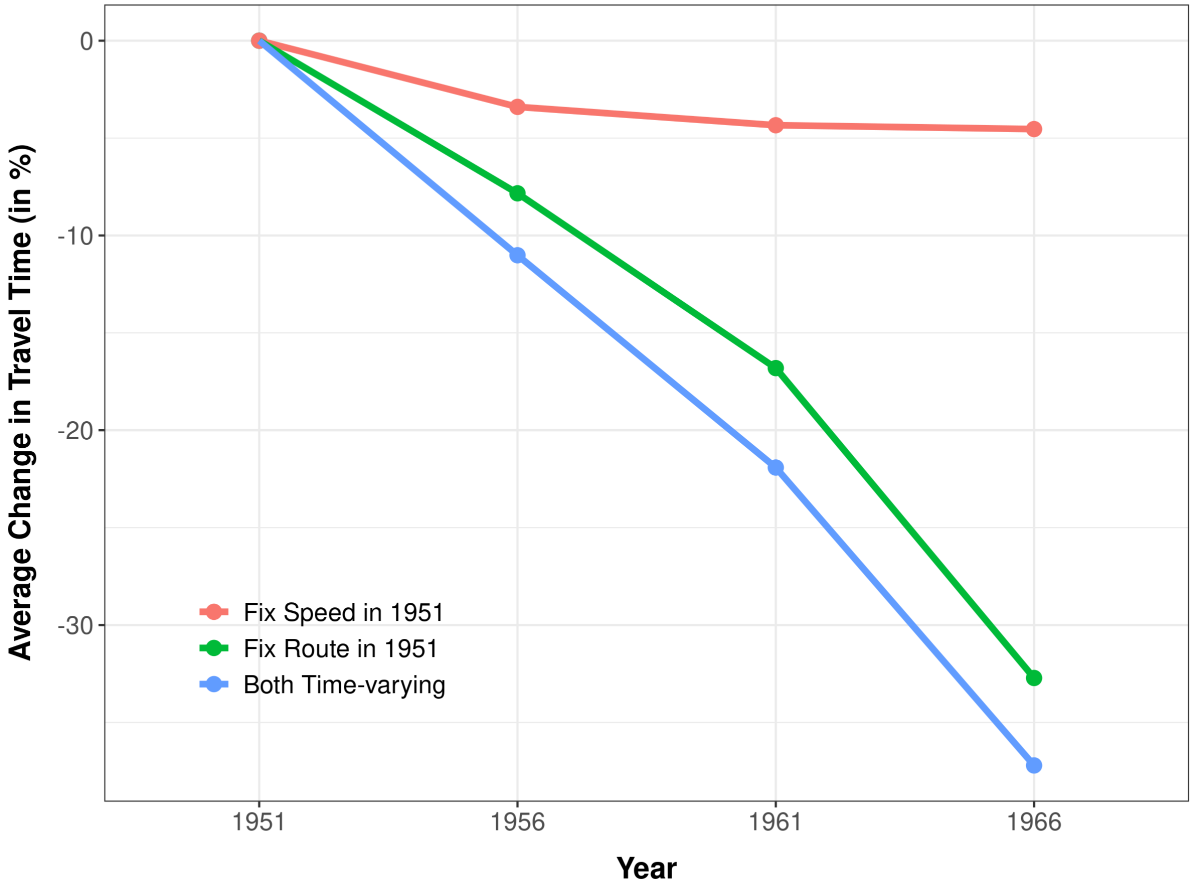


Figure 16: Counterfactual change in travel time

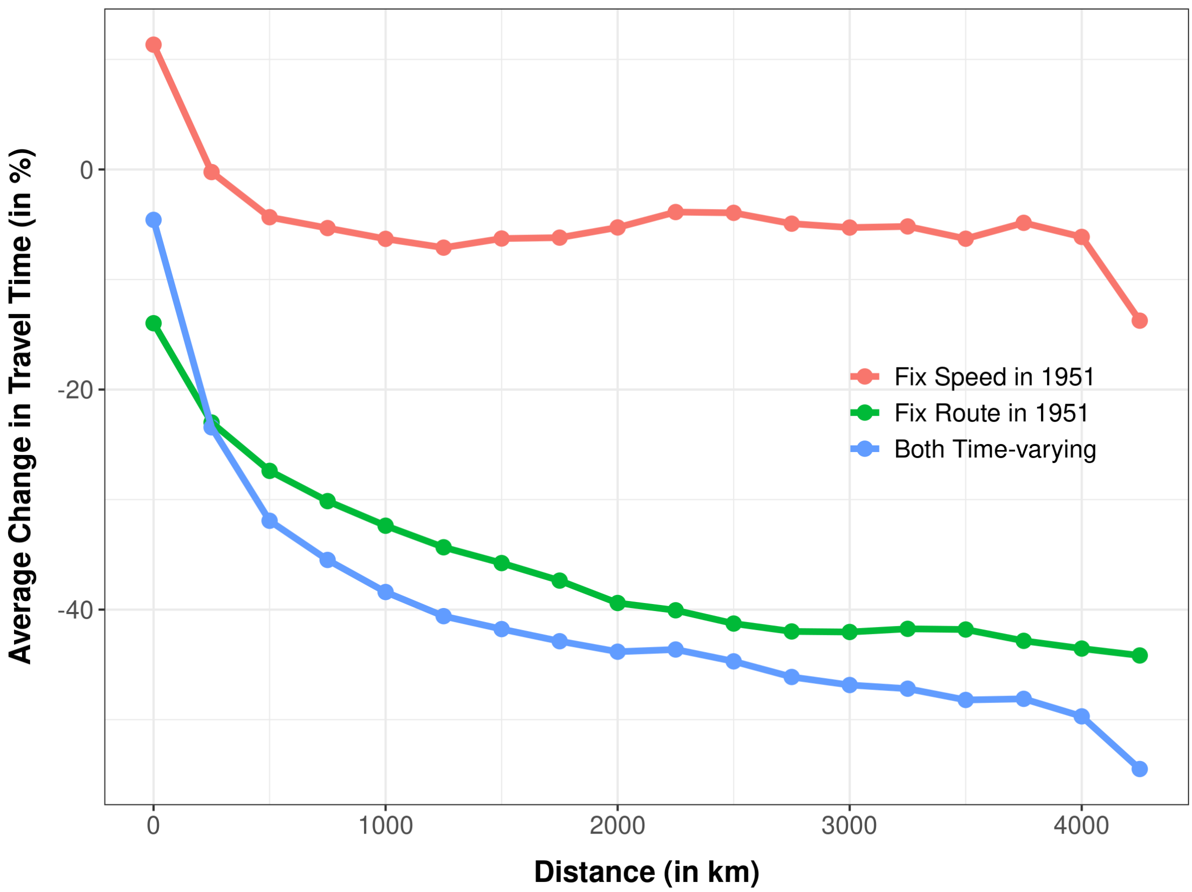


Figure 17: Counterfactual change in travel time 1951-1966

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 18 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.

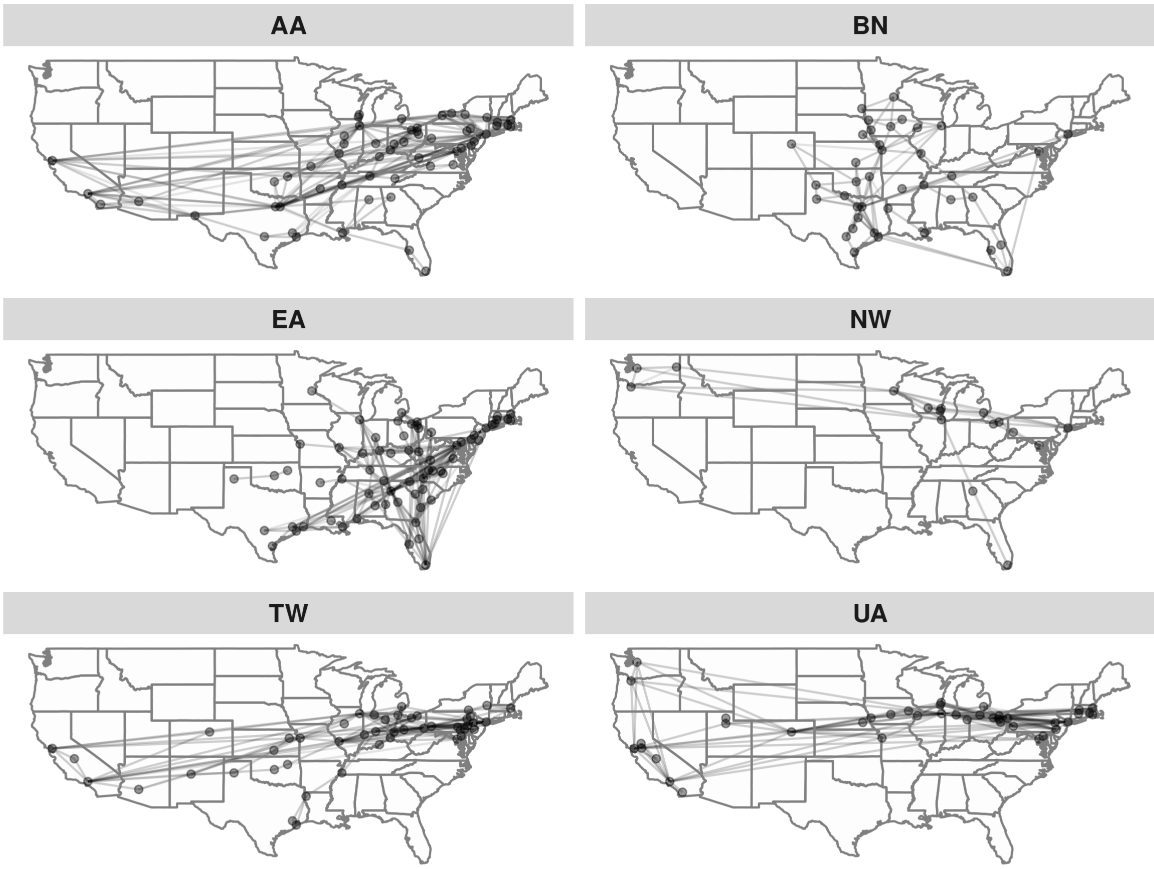


Figure 18: Flight Network in 1956 by Airline (weighted by log frequency).

C.3. Instrumental travel time

In order 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. 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. Results on this estimation are provided in Table 7.

Table 7: Regression of travel time on distance fastest non-stop flights

	Travel time (min)			
Year	1951	1956	1961	1966
	(1)	(2)	(3)	(4)
Constant	25.3***	24.1***	39.5***	29.9***
	(0.809)	(0.656)	(0.921)	(0.678)
Distance (km)	0.146***	0.132***	0.079***	0.068***
	(0.001)	(0.0007)	(0.0010)	(0.0006)
Observations	1,137	1,479	1,438	1,490
R^2	0.93	0.96	0.82	0.90
Implied speed (km/h)	412	453	758	876

The table presents the results of estimating by OLS: travel $\text{time}_{ijt} = \alpha_0 + \alpha_1 \times \textit{distance}_{ij} + \varepsilon_{ijt}$ separately for each year $t \in \{1951, 1956, 1961, 1966\}$. The sample consist of all airport pairs that are connected with a non-stop flight in the respective year. Travel time is the fastest non-stop flight between the airports measured in minutes. The implied speed is calculated as the inverse of the coefficient on distance multiplied by 60.

Table 3: Connected MSAs

MSA fips	MSA name	<=3 periods	4 periods	MSA fips	MSA name	<=3 periods	4 periods
80	Akron, OH SMA	X	X	4680	Macon, GA SMA	X	X
160	Albany-Schenectady-Troy, NY SMA	X	X	4720	Madison, WI SMA	X	X
200	Albuquerque, NM SMA	X	X	4760	Manchester, NH SMA		
240	Allentown-Bethlehem-Easton, PA-NJ SMA	X	X	4920	Memphis, TN SMA	X	X
280	Altoona, PA SMA			5000	Miami, FL SMA	X	X
320	Amarillo, TX SMA	X	X	5080	Milwaukee, WI SMA	X	X
480	Asheville, NC SMA	X		5120	Minneapolis-St. Paul, MN SMA	X	X
520	Atlanta, GA SMA	X	X	5160	Mobile, AL SMA	X	X
560	Atlantic City, NJ SMA	X		5240	Montgomery, AL SMA	X	X
600	Augusta, GA-SC SMA	X	X	5280	Muncie, IN SMA		
640	Austin, TX SMA	X	X	5360	Nashville, TN SMA	X	X
720	Baltimore, MD SMA	X	X	5400	New Bedford, MA SMA		
760	Baton Rouge, LA SMA	X		5440	New Britain-Bristol, CT SMA		
800	Bay City, MI SMA	X		5480	New Haven, CT SMA	X	X
840	Beaumont-Port Arthur, TX SMA	X		5560	New Orleans, LA SMA	X	X
960	Binghamton, NY SMA	X		5600	New York-Northeastern NJ, NY-NJ SMA	X	X
1000	Birmingham, AL SMA	X	X	5720	Norfolk-Portsmouth, VA SMA	X	
1120	Boston, MA SMA	X	X	5840	Ogden, UT SMA	X	
1160	Bridgeport, CT SMA	X	X	5880	Oklahoma City, OK SMA	X	X
1200	Brockton, MA SMA			5920	Omaha, NE-IA SMA	X	X
1280	Buffalo, NY SMA	X	X	5960	Orlando, FL SMA	X	X
1320	Canton, OH SMA	X	X	6120	Peoria, IL SMA	X	
1360	Cedar Rapids, IA SMA	X	X	6160	Philadelphia, PA-NJ SMA	X	X
1440	Charleston, SC SMA	X	X	6200	Phoenix, AZ SMA	X	X
1480	Charleston, WV SMA	X	X	6280	Pittsburgh, PA SMA	X	X
1520	Charlotte, NC SMA	X	X	6320	Pittsfield, MA SMA		
1560	Chattanooga, TN-GA SMA	X	X	6400	Portland, ME SMA		
1600	Chicago, IL-IN SMA	X	X	6440	Portland, OR-WA SMA	X	X
1640	Cincinnati, OH-KY SMA	X	X	6480	Providence, RI SMA	X	X
1680	Cleveland, OH SMA	X	X	6560	Pueblo, CO SMA	X	
1760	Columbia, SC SMA	X	X	6600	Racine, WI SMA	X	X
1800	Columbus, GA-AL SMA	X	X	6640	Raleigh, NC SMA	X	X
1840	Columbus, OH SMA	X	X	6680	Reading, PA SMA	X	X
1880	Corpus Christi, TX SMA	X	X	6760	Richmond, VA SMA	X	X
1920	Dallas, TX SMA	X	X	6800	Roanoke, VA SMA	X	X
1960	Davenport-Rock Island-Moline, IA-IL SMA	X	X	6840	Rochester, NY SMA	X	X
2000	Dayton, OH SMA	X	X	6880	Rockford, IL SMA		
2040	Decatur, IL SMA			6920	Sacramento, CA SMA	X	X
2080	Denver, CO SMA	X	X	6960	Saginaw, MI SMA	X	
2120	Des Moines, IA SMA	X	X	7000	St. Joseph, MO SMA	X	
2160	Detroit, MI SMA	X	X	7040	St. Louis, MO-IL SMA	X	X
2240	Duluth-Superior, MN-WI SMA	X		7160	Salt Lake City, UT SMA	X	X
2280	Durham, NC SMA	X	X	7200	San Angelo, TX SMA		
2320	El Paso, TX SMA	X	X	7240	San Antonio, TX SMA	X	X
2360	Erie, PA SMA	X		7280	San Bernardino, CA SMA		
2440	Evansville, IN SMA	X	X	7320	San Diego, CA SMA	X	X
2480	Fall River, MA-RI SMA	X	X	7360	San Francisco-Oakland, CA SMA	X	X
2640	Flint, MI SMA	X		7400	San Jose, CA SMA		
2760	Fort Wayne, IN SMA	X	X	7520	Savannah, GA SMA	X	
2800	Fort Worth, TX SMA	X	X	7560	Scranton, PA SMA	X	X
2840	Fresno, CA SMA	X	X	7600	Seattle, WA SMA	X	X
2880	Gadsden, AL SMA			7680	Shreveport, LA SMA	X	
2920	Galveston, TX SMA	X	X	7720	Sioux City, IA SMA	X	
3000	Grand Rapids, MI SMA	X		7760	Sioux Falls, SD SMA	X	
3080	Green Bay, WI SMA			7800	South Bend, IN SMA	X	X
3120	Greensboro-High Point, NC SMA	X	X	7840	Spokane, WA SMA	X	X
3160	Greenville, SC SMA	X	X	7880	Springfield, IL SMA	X	
3200	Hamilton-Middletown, OH SMA			7920	Springfield, MO SMA	X	
3240	Harrisburg, PA SMA	X	X	7960	Springfield, OH SMA		
3280	Hartford, CT SMA	X	X	8000	Springfield-Holyoke, MA-CT SMA	X	X
3360	Houston, TX SMA	X	X	8040	Stamford-Norwalk, CT SMA	X	
3400	Huntington-Ashland, WV-KY-OH SMA	X		8120	Stockton, CA SMA	X	X
3480	Indianapolis, IN SMA	X	X	8160	Syracuse, NY SMA	X	X
3520	Jackson, MI SMA	X		8200	Tacoma, WA SMA		
3560	Jackson, MS SMA			8280	Tampa-St. Petersburg, FL SMA	X	X
3600	Jacksonville, FL SMA	X	X	8320	Terre Haute, IN SMA	X	X
3680	Johnstown, PA SMA			8400	Toledo, OH-MI SMA	X	X
3720	Kalamazoo, MI SMA	X		8440	Topeka, KS SMA	X	
3760	Kansas City, MO-KS SMA	X	X	8480	Trenton, NJ SMA		
3800	Kenosha, WI SMA			8560	Tulsa, OK SMA	X	X
3840	Knoxville, TN SMA	X	X	8680	Utica-Rome, NY SMA		
4000	Lancaster, PA SMA	X	X	8800	Waco, TX SMA	X	
4040	Lansing, MI SMA	X		8840	Washington, DC-MD-VA SMA	X	X
4080	Laredo, TX SMA	X		8880	Waterbury, CT SMA		
4160	Lawrence, MA SMA			8920	Waterloo, IA SMA	X	
4280	Lexington, KY SMA	X	X	9000	Wheeling-Steubenville, WV-OH SMA	X	
4320	Lima, OH SMA			9040	Wichita, KS SMA	X	X
4360	Lincoln, NE SMA	X	X	9080	Wichita Falls, TX SMA	X	X
4400	Little Rock-North Little Rock, AR SMA	X	X	9120	Wilkes-Barre–Hazleton, PA SMA	X	X
4440	Lorain-Elyria, OH SMA	X	X	9160	Wilmington, DE-NJ SMA	X	X
4480	Los Angeles, CA SMA	X	X	9220	Winston-Salem, NC	X	X
4520	Louisville, KY-IN SMA	X	X	9240	Worcester, MA SMA	X	
4560	Lowell, MA SMA			9280	York, PA SMA	X	X
4600	Lubbock, TX SMA	X	X	9320	Youngstown, OH-PA SMA	X	X

D. Patent data

In this appendix we describe facts that we observe in the US patent data, for patents filed between 1945 and 1975.[15]

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 contains patent number, filing year and citations.[16,17] Based on the patent number we merge it with multiple datasets. First, we obtained technology class from the USPTO Master Classification File (**?**) and we aggregated them to the six technology categories of **?**. Second, we obtained geographic location of inventors from three datasets: HistPat (**?**), HistPat International (**?**) and Fung Institute (**?**). 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: **?** for patents owned by firms listed in the US stock market and Patstat (**?**) for the remaining unmatched patents.[18]

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. **?** seems to use a matching method based on the patent owner declared in the patent text, as Patstat does. Specially, **?** does not explicitly say if it takes into account firm-ownership structure to determine patent

[15]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, 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.

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

[17]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)

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

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 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.

D.1. Matching patents to locations

In figure 19 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.

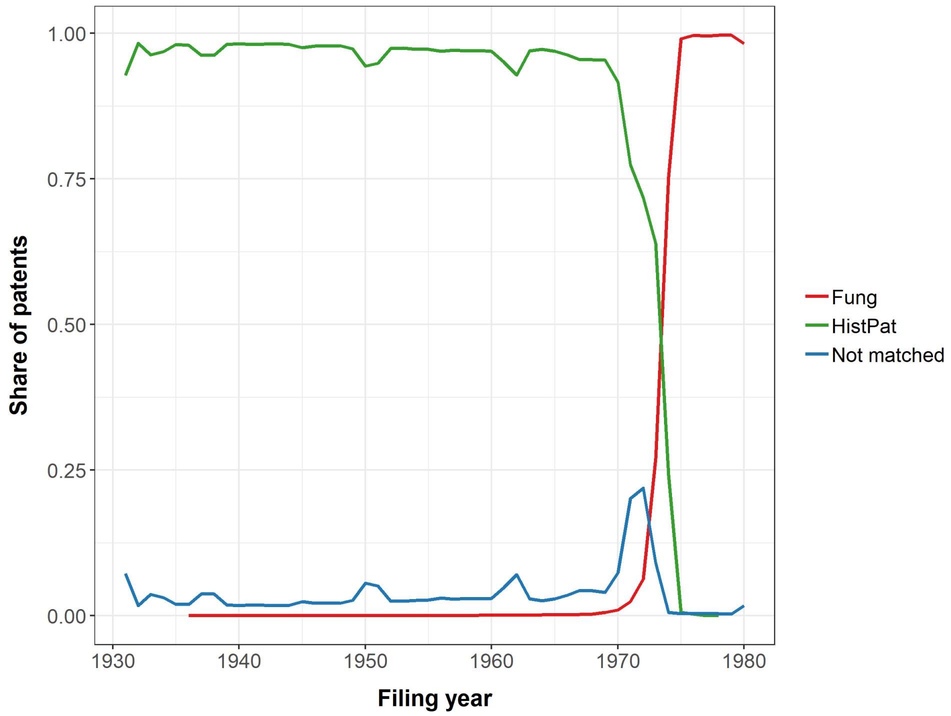


Figure 19: Non-matching rate HistPat, HistPat International and Fung

Figure 20 shows the share of patents that have inventors inside MSAs, and figure 21 displays the same by technology category.[19]

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

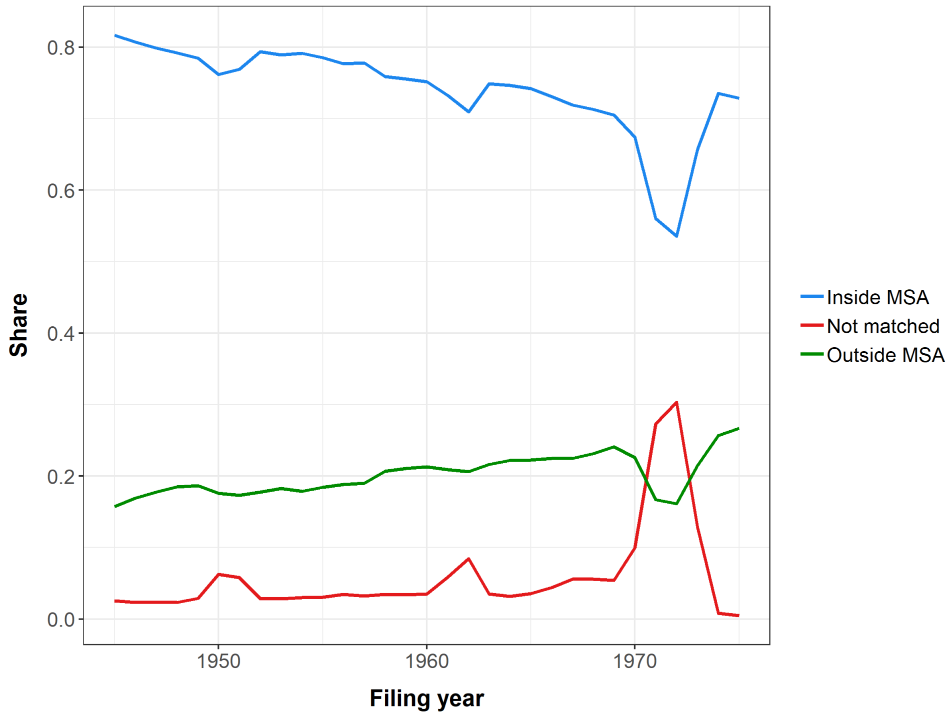


Figure 20: Share patents in Metropolitan Statistical Areas

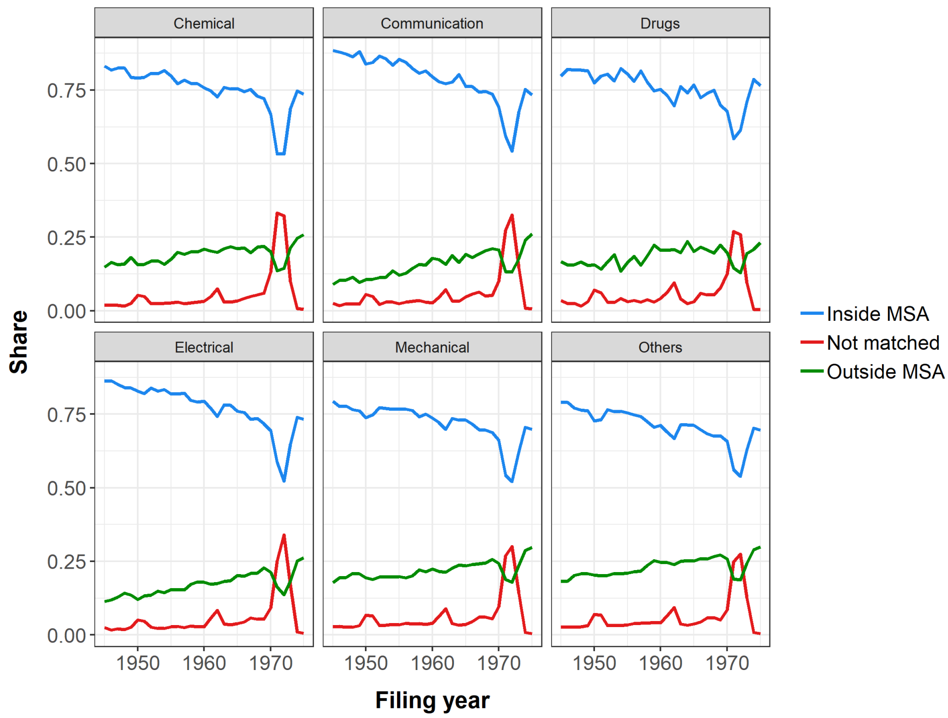


Figure 21: Share patents in Metropolitan Statistical Areas

D.2. Input-Output of patents

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 22 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.

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 22: Input-Output of technologies 1949-1953

D.3. Descriptive statistics

Table 8 shows descriptive statistics along each step of the patent data matching and sample selection. The final dataset contains 515,089 patents and 1,639,326 citations.

Table 8: Patent data sample selection

Sample	N patents	N citations	1st quartile cit dist (km)	2nd quartile cit dist (km)	3rd quartile cit dist (km)
Google patents	964,582	4,392,725			
With location	923,150	4,191,886			
US	749,410	3,569,578			
MSA	589,870	2,354,844			
Single location	571,969	2,237,095	213	730	1,682
With owner id	571,824	1,963,644	199	696	1,673
Non gov/univ	565,372	1,932,297	199	696	1,664
With travel time (final sample)	515,089	1,639,326	184	689	1,645

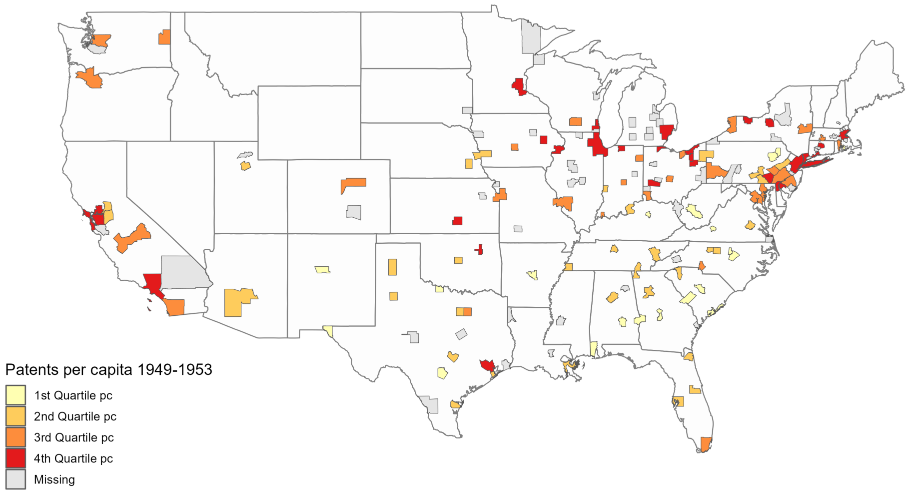


Figure 23: Patents per capita in 1951
Quantiles of patents per capita are computed in each technology and then averaged across technologies. Population is from 1950 Census.

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

Figure 24 shows the MSA's ranking of innovativeness in 1951 and its subsequent patenting growth rate in 1951-1966. MSAs that were initially more innovative (lower values in the ranking) are those that saw lower values of subsequent patenting growth.[20] We

[20]Each dot in Figure 24 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

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 25 shows that MSAs that were initially less innovative and had high subsequent growth were located in all four regions.

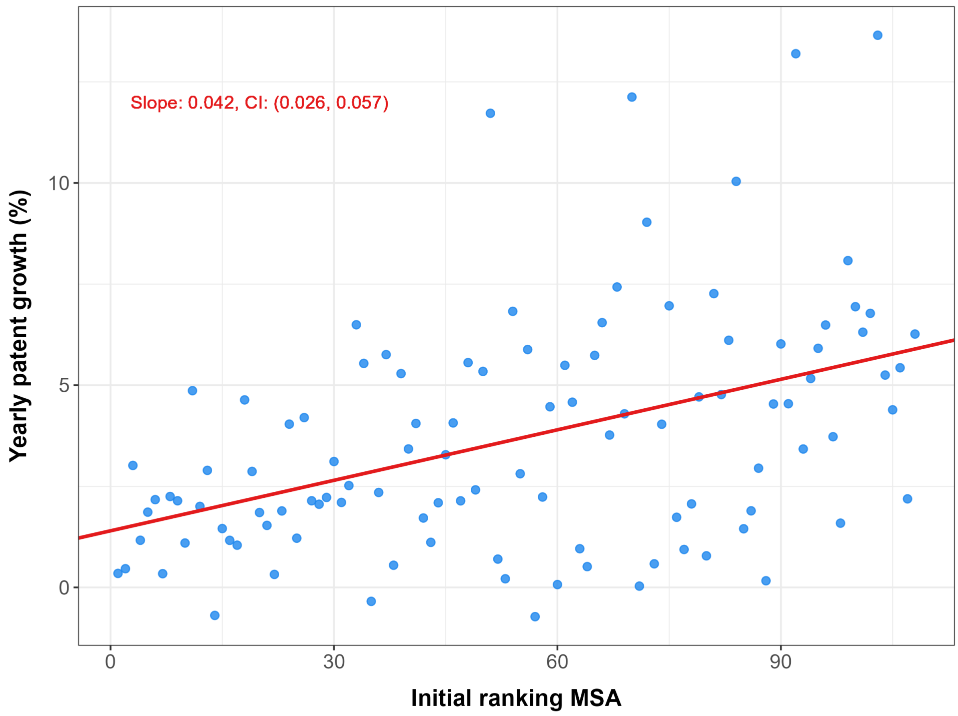


Figure 24: Patent growth rate by initial innovativeness ranking of MSA

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).

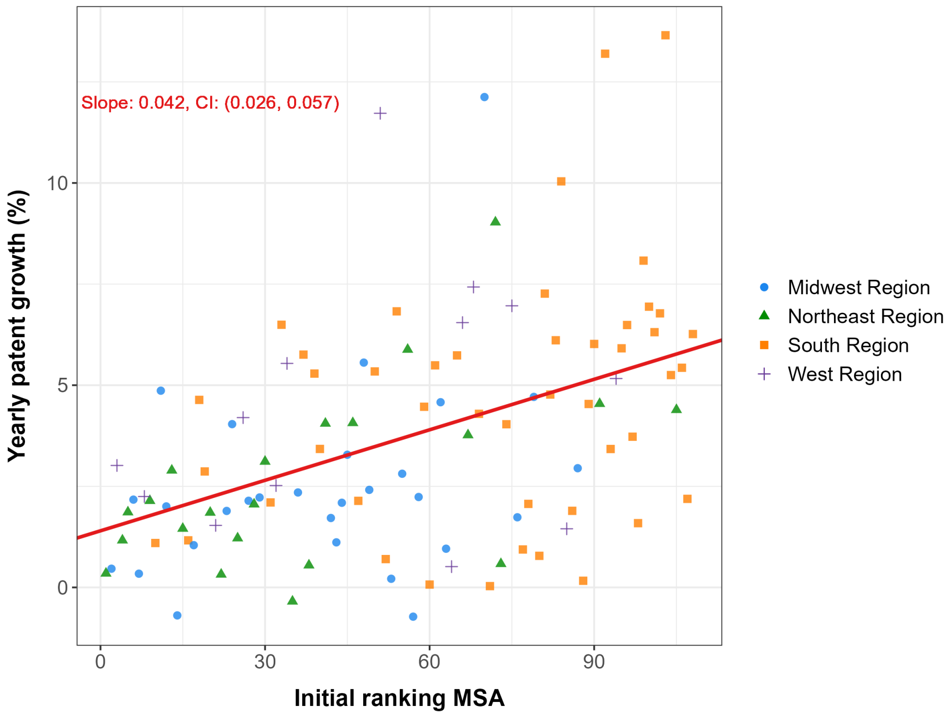


Figure 25: Patent growth by initial innovativeness ranking of MSA

Fact 2: The South and the West of the US had a higher patenting growth rate

Figures 26 and 27 present averages across technologies within a region-year. Figure 26 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 change in the shares implies a higher growth rate of patenting in the South and the West relative to the Midwest and the Northeast.

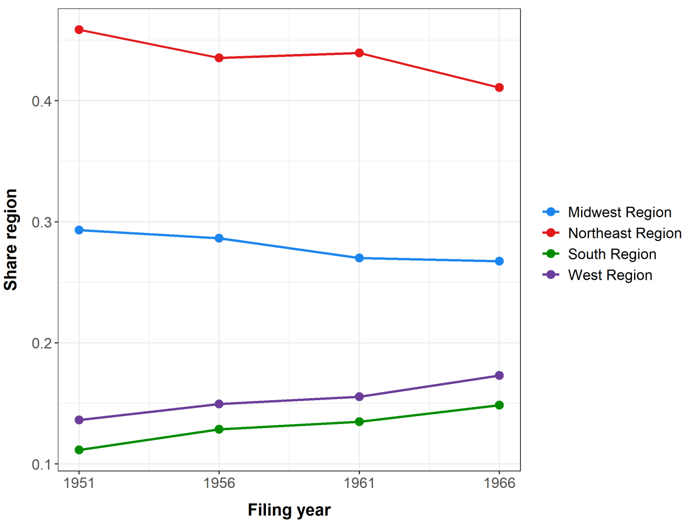


Figure 26: Share of patents by region

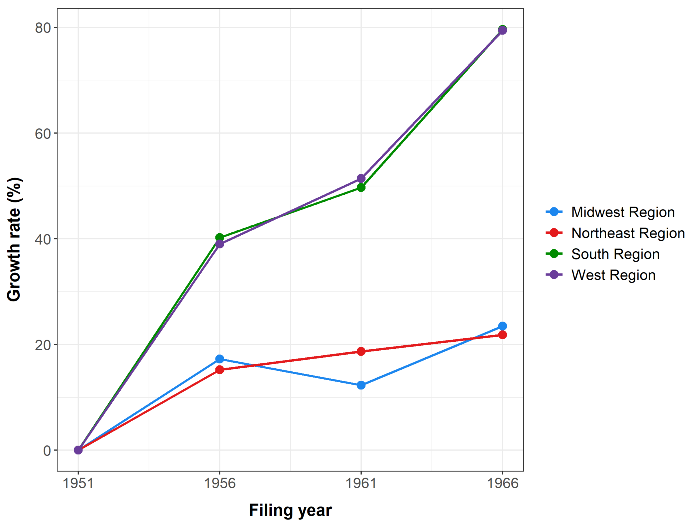


Figure 27: Patent growth by region

D.3.1. Descriptive statistics by technology

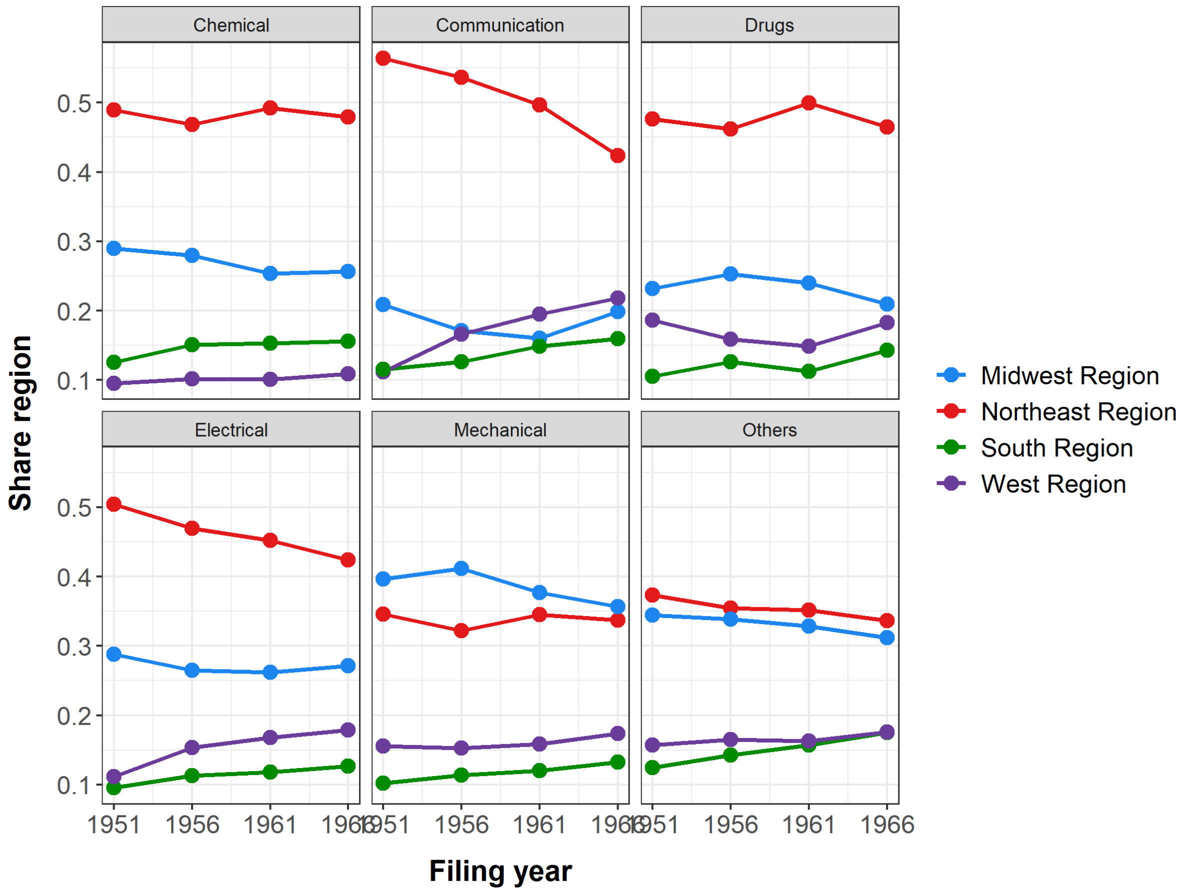


Figure 28: Share of patents by region

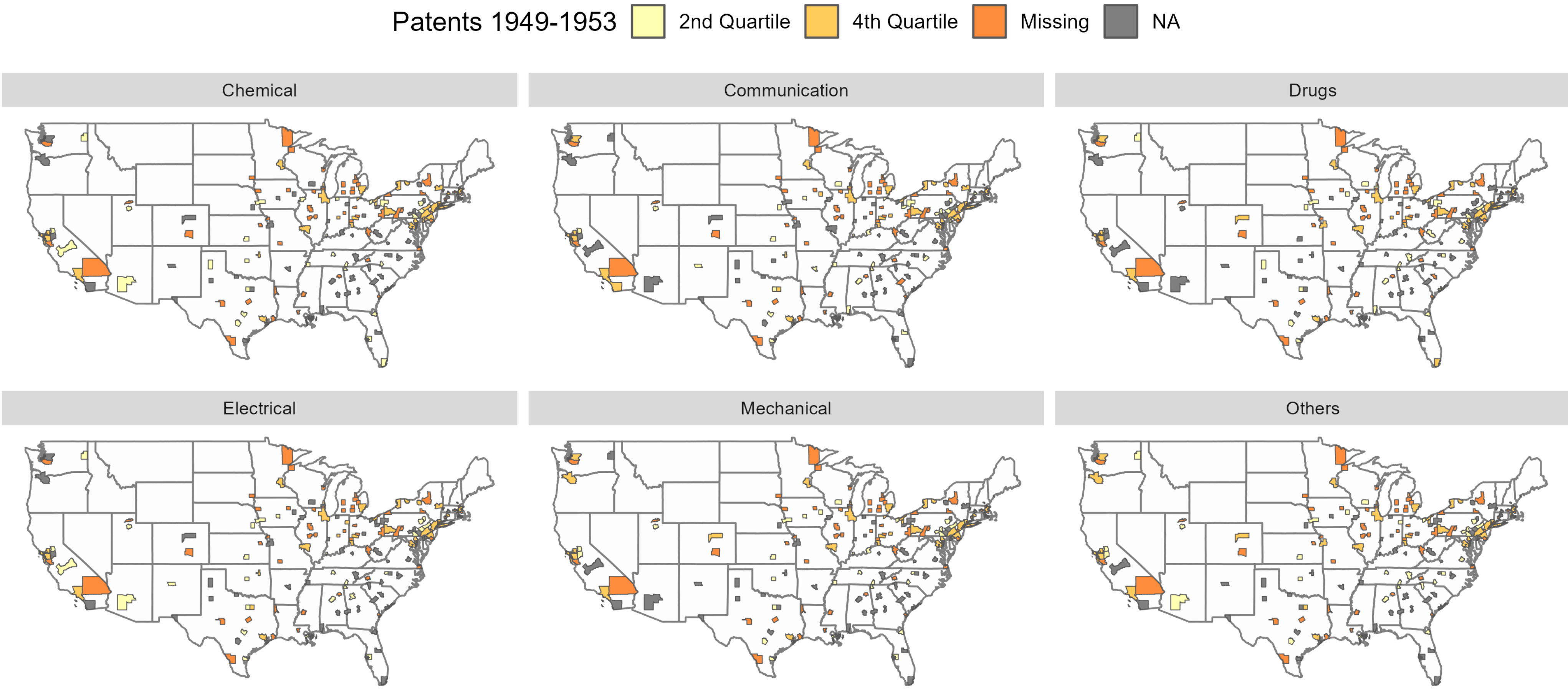


Figure 30: Geography of patenting 1951

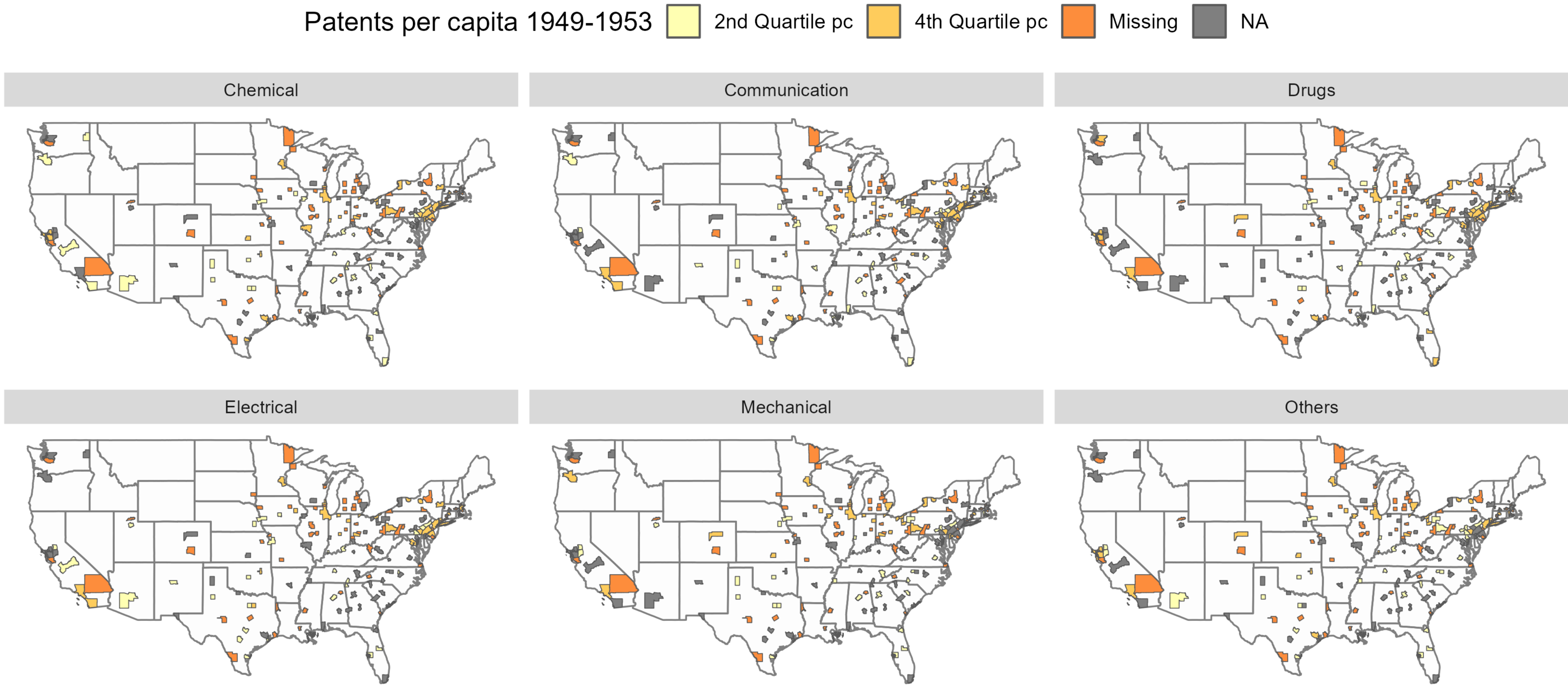


Figure 31: Patents per capita in 1951

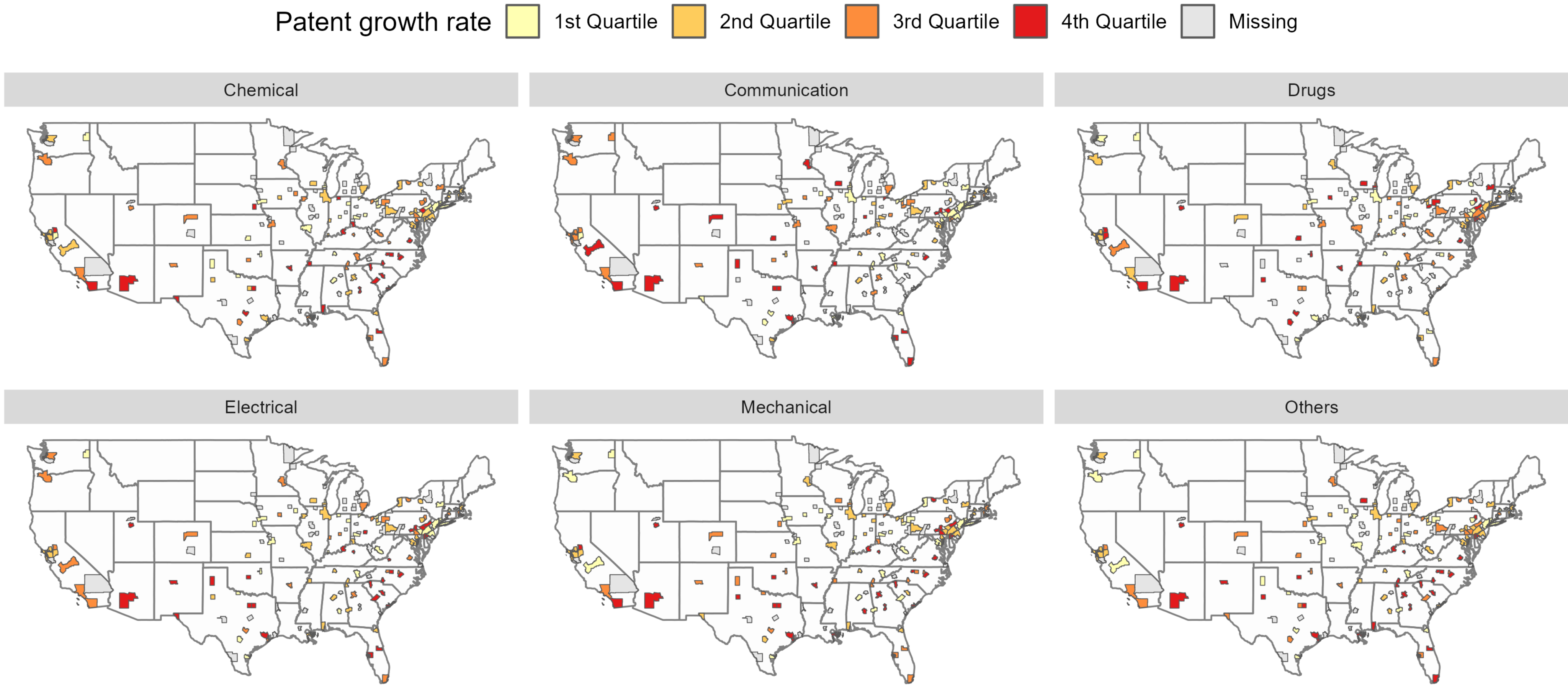


Figure 32: Patent growth rate 1951-1966

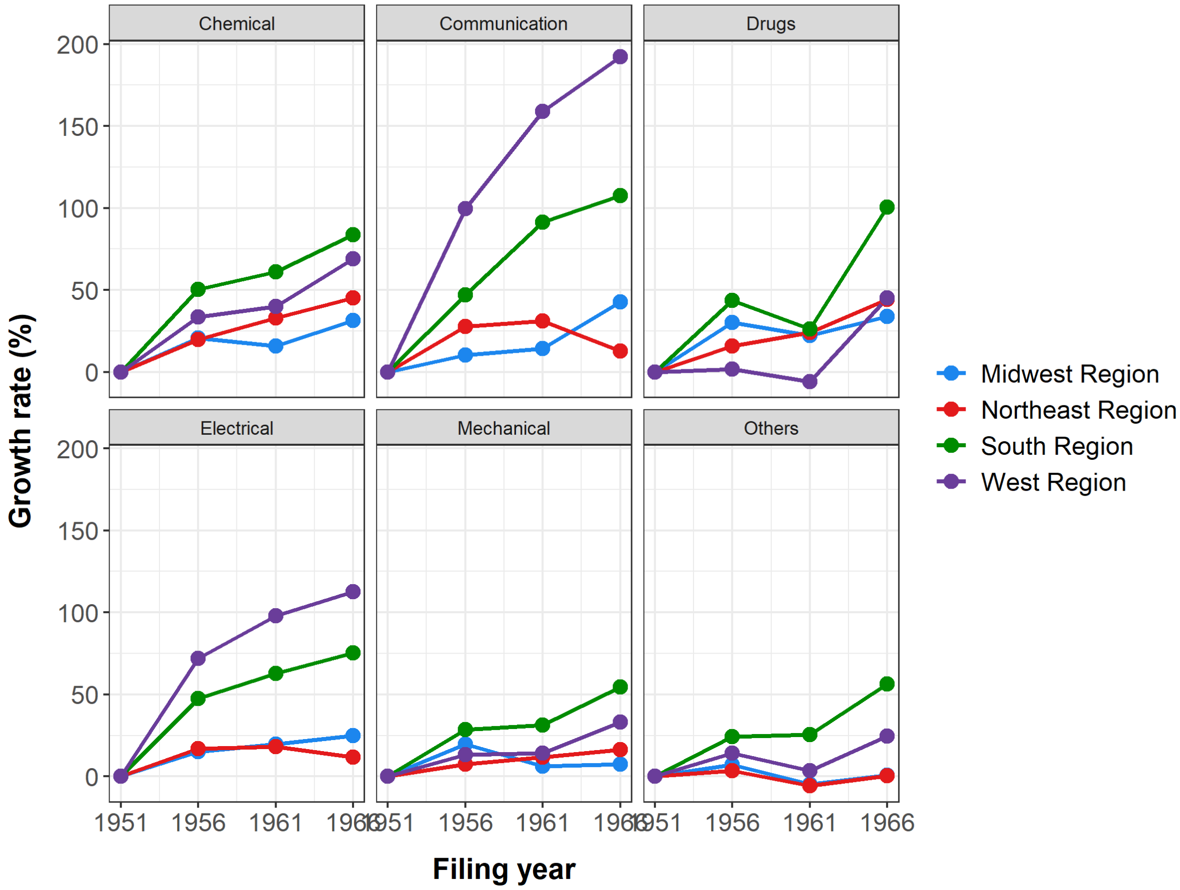


Figure 29: Patent growth rate by region

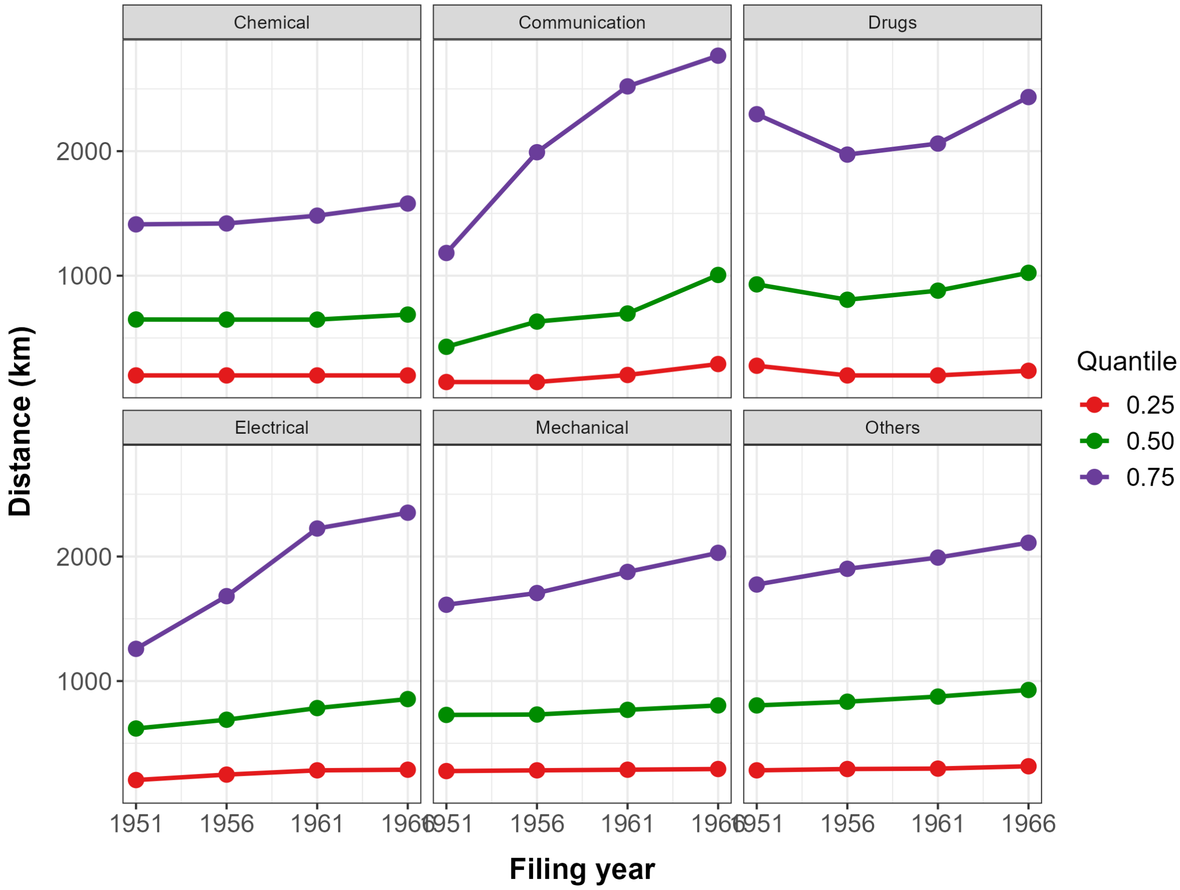


Figure 33: Quantiles of citation distance

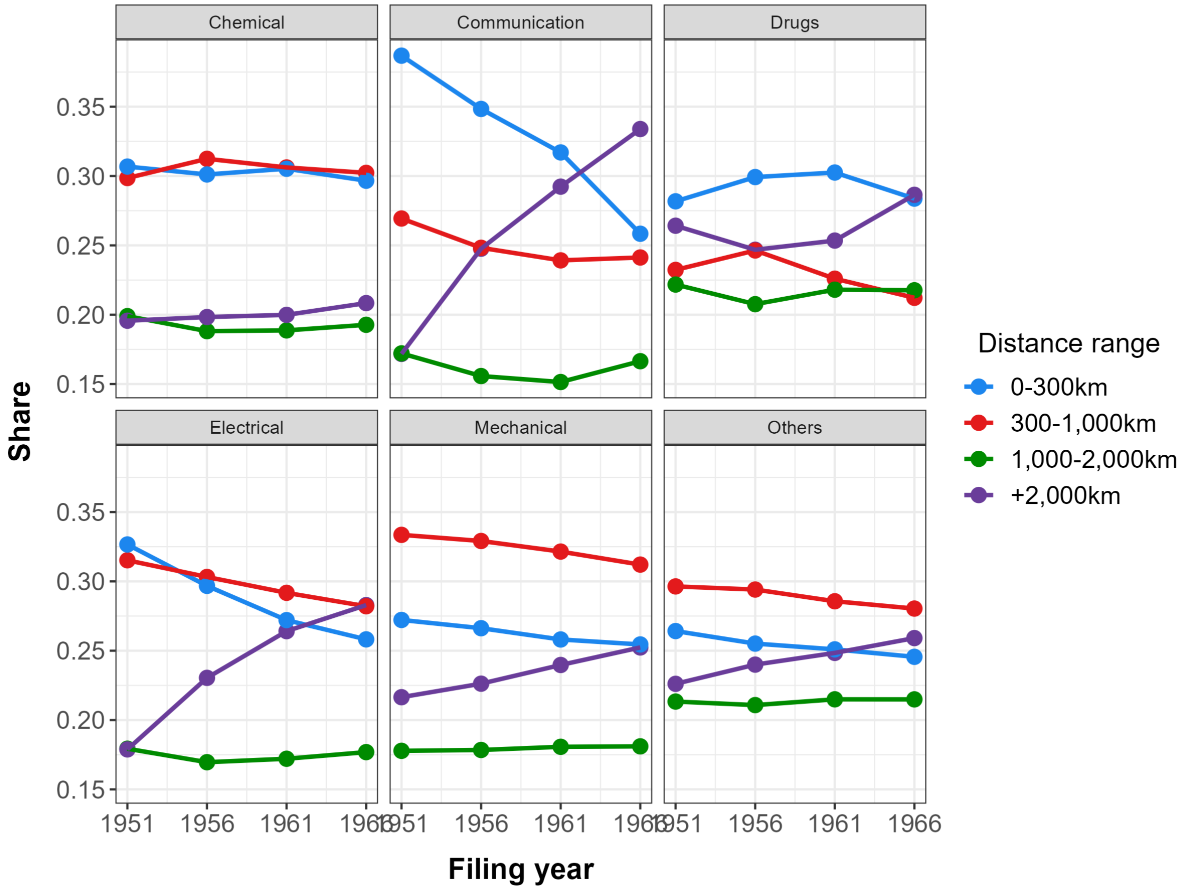


Figure 34: Share of citations by distance

E. Appendix: US Census Regions

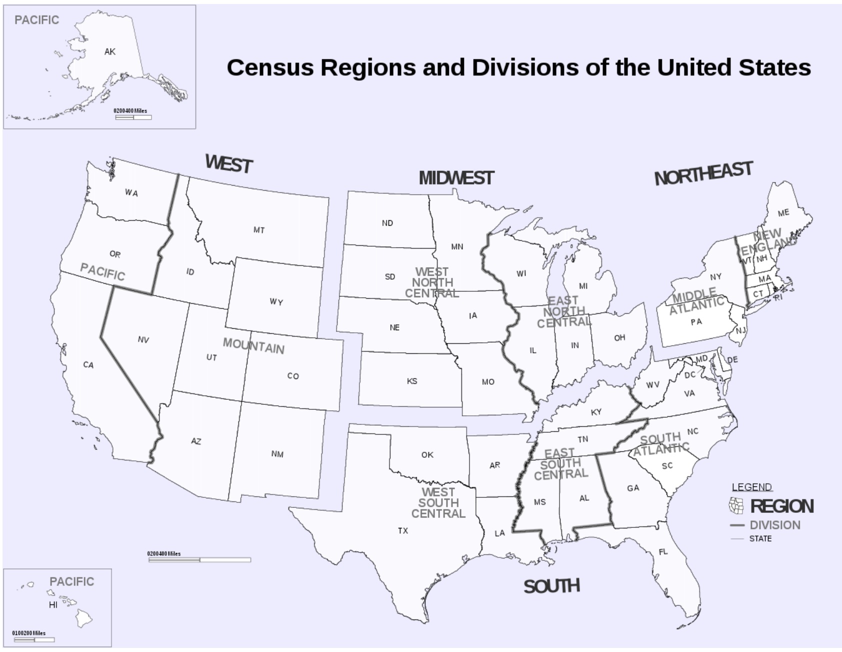


Figure 35: US Census Regions
Source: US Census Bureau

F. Bias Correction and IV estimation

F.1. Split-panel jackknife bias correction

Weidner and Zylkin (2021) show that PPML estimation of gravity equations with three-way fixed effects (origin-time, destination-time, origin-destination) is consistent but asymptotically biased. In their words: "*the asymptotic distribution of the estimates is not centered at the truth as* $N \to \infty$" (page 2). The asymptotic bias concerns both point estimates and standard errors. In order to correct the bias we apply their suggested split-panel jackknife bias correction of section 3.4.1 to both point estimates and bootstrap standard errors. The idea of the jackknife bias correction is to estimate the model in many subsamples and then subtract the average coefficients of the subsamples from (twice) the original coefficient.

As suggested in Weidner and Zylkin (2021) when using real world data (as opposite to simulated data), we estimate the bias correction repeatedly. We modify equation (14) in Weidner and Zylkin (2021) to define the bias corrected coefficient as:

$$\tilde{\beta}_N^J := 2 \times \hat{\beta} - \frac{1}{Z} \sum_z \sum_p \frac{\hat{\beta}_{(p,z)}}{4} \tag{3}$$

where p is a random subsample of size 1/4th of the original sample, and Z is the amount of times to subsample.

The *procedure to estimate bias corrected point estimate* $\tilde{\beta}_N^J$ is as follows:

1. Estimate $\hat{\beta}$: the not-bias-corrected estimate of equation (1)
2. Randomly allocate all citing establishment-technology Fih into two equally sized groups (groups are time-invariant). Call them citing groups a and b.
3. Randomly allocate all cited establishment-technology Gjk into two equally sized groups (groups are time-invariant). Call them cited groups a and b.
4. Create four p subsamples of the original data: (a,a), (a,b), (b,a), (b,b). Subsamples keep the same granularity as the original data $FiGjhkt$.
5. Estimate equation (1) (gravity equation of the main text) in each of the subsamples from the previous step to obtain $\hat{\beta}_{(p,z)}$.[21] Store the four estimated coefficients.
6. Repeat Z times steps 2 to 5.
7. Compute equation 3

To compute bias-corrected bootstrap standard errors we need to bias-correct the point estimate $\tilde{\beta}_m^J$ of each bootstrap iteration m. The *procedure to estimate bias corrected standard errors* is as follows:

[21]Given that we require to identify the fixed effects, the *effective subsample* in all four p estimations does not have the same amount of observations. However, in our estimations the *effective subsample* size across p subsamples does not differ by more than 5%.

1. Sample establishment-technology-pairs $FiGjhk$ with replacement such that we obtain a re-sampled data of the same size as the original data (hence, some $FiGjhk$ will be repeated in the re-sampled data). Sampled $FiGjhk$ are kept for all time periods in order to keep the source of identification of β: across time variation within a establishment pair. Label this new dataset $data_m$.

2. Using $data_m$, estimate equation (1) to obtain $\hat{\beta_m}$ (this is a point estimate of the specific $data_m$)

3. Using $data_m$, repeat Z_M times steps 2 to 5 of the *procedure to estimate bias corrected point estimate*. This step provides $Z_M \times 4$ point estimates $\hat{\beta}_{(p,m,z_M)}$

4. Compute the bias corrected point estimate of bootstrap m $\tilde{\beta}^J_m = 2 \times \hat{\beta}_m - \frac{1}{Z_M} \sum_{z_M} \sum_p \frac{\hat{\beta}_{(p,m,z_M)}}{4}$.

5. Store the bias corrected point estimate of bootstrap m

6. Repeat steps 1 to 5 M times to obtain M bias corrected bootstrap point estimates $\tilde{\beta}^J_m$

7. Compute the variance-covariance matrix of bias corrected bootstrap coefficients $\tilde{\beta}^J_m$ and use it to compute standard errors of $\tilde{\beta}^J_N$

The bias correction of point estimates and bias correction of bootstrap standard errors implies estimating $Z \times 4 + Z_M \times M \times 4$ models. This is a computationally demanding task. To estimate columns (1) and (2) of Table 1 we set $Z = 100$, $Z_M = 5$ and $M = 200$, adding up to $1,100$ models to estimate for each column.

As recommended in **?**, in the Table 9 we repeat Table 1 but reporting 0.025 and 0.975 quantile values of bootstrap estimates (bias corrected for columns (1) and (2)) instead of standard errors:

	PPML		IV PPML	
Dep. variable:		*citations*		
	(1)	(2)	(3)	(4)
log(travel time)	-0.084^{***} $(-0.116; -0.056)$		-0.161^{***} $(-0.232; -0.108)$	
log(travel time) $\times$ 0-300km		-0.015 $(-0.073; 0.031)$		-0.185 $(-0.489; 0.097)$
log(travel time) $\times$ 300-1,000km		-0.085^{***} $(-0.137; -0.050)$		-0.155^{***} $(-0.255; -0.083)$
log(travel time) $\times$ 1,000-2,000km		-0.096^{***} $(-0.148; -0.024)$		-0.132^{**} $(-0.204; -0.033)$
log(travel time) $\times$ +2,000km		-0.166^{***} $(-0.246; -0.115)$		-0.206^{***} $(-0.299; -0.143)$
N obs. effective	5,147,161	5,147,161	5,147,161	5,147,161
R2	0.88	0.88	0.88	0.88

Table 9: 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. 0.025 and 0.975 quantile bootstrap estimates are presented in parentheses. The coefficients and bootstrap estimates in columns (1) and (2) are jackknife bias-corrected. R2 is computed as the squared correlation between observed and fitted values.

F.2. Instrumental variables PPML

To implement the instrumental variables of Poisson estimation we follow the control function approach described in Wooldridge (2014). We explain the procedure using the estimation of the elasticity of citations to travel time. The procedure is similar for the elasticity of (new) patents to knowledge access. We proceed in two steps estimating the following two equations:

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

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

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 in the following way:

1. Sample establishment-technology-pairs $FiGjhk$ with replacement such that we obtain a re-sampled data of the same size as the original data (hence, some $FiGjhk$ will be repeated in the re-sampled data). Sampled $FiGjhk$ are kept for all time periods in order to keep the source of identification of β: across time variation within a establishment pair. Label this new dataset $data_m$

2. Using $data_m$, estimate equations (4) and (5) to obtain the bootstrap estimate $\hat{\beta}_m$. Store $\hat{\beta}_m$.

3. Repeat M times steps 1 and 2.

4. Compute the variance-covariance matrix of $\hat{\beta}_m$ and use it to compute standard errors of $\hat{\beta}$

For columns (3) and (4) of Table 1, and columns (3) to (6) of Table 3 we set $M = 200$.

G. Additional results

G.1. Diffusion of knowledge

G.1.1. Heterogeneous effects

First, 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 11.

Second, 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 11.

Third, we extend the sample of patents to include patents with a patent owner identified as a government organization or university. Column (5) of Table 11 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.

Fourth, 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 11 shows that the elasticity is not statistically different for self citations.

Fifth, 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 10. We do not find a particular pattern related to the initial innovativeness.

Sixth, we check if the elasticity varies with the citing technology, cited technology and citing-cited technology pair. Results are shown in Table 12 and Table 13. We find that the elasticity is negative and significant mainly when the citing and cited technology are the same. In Appendix D we show that most citations happen within a technology, so most identification power would be when citing and cited technologies are the same.

	Concentration citing	Concentration cited	Cited lag forward	Citing lag backward	Citing govt & uni	Cited govt & univ	Self citation
Dep. variable:				*citations*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
log(travel time):0-300km	0.037	0.077	−0.035	0.068	−0.014	0.015	−0.031
	(0.121)	(0.125)	(0.447)	(0.523)	(0.041)	(0.041)	(0.042)
log(travel time):300-1000km	-0.148^{*}	−0.073	−0.346	−0.014	-0.096^{***}	-0.095^{***}	-0.079^{***}
	(0.077)	(0.090)	(0.333)	(0.344)	(0.041)	(0.025)	(0.027)
log(travel time):1000-2000km	−0.146	−0.129	0.039	0.068	-0.093^{**}	-0.093^{**}	-0.092^{**}
	(0.102)	(0.110)	(0.458)	(0.488)	(0.041)	(0.041)	(0.040)
log(travel time):+2000km	-0.300^{***}	-0.278^{***}	0.796^{**}	0.673	-0.176^{***}	-0.178^{***}	-0.144^{***}
	(0.100)	(0.086)	(0.334)	(0.471)	(0.048)	(0.047)	(0.040)
log(travel time):0-300km $\times X$	−0.759	−1.346	0.009	−0.032	0.095	0.565	0.137
	(1.749)	(1.705)	(0.179)	(0.206)	(0.342)	(0.469)	(0.191)
log(travel time):300-1000km $\times X$	0.786	−0.320	0.100	−0.032	−0.105	-0.636^{*}	0.129
	(1.081)	(1.278)	(0.132)	(0.136)	(0.244)	(0.301)	(0.122)
log(travel time):1000-2000km $\times X$	0.757	0.512	−0.053	−0.065	−0.275	−0.304	0.106
	(1.379)	(1.511)	(0.181)	(0.192)	(0.365)	(0.370)	(0.199)
log(travel time):+2000km $\times X$	1.732	1.421	-0.393^{***}	-0.341^{*}	−0.235	0.093	−0.054
	(1.427)	(1.124)	(0.132)	(0.186)	(0.406)	(0.293)	(0.166)
N obs. effective	5,147,161	5,147,161	5,147,161	5,147,161	5,250,386	5,250,386	5,287,792
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 11: Elasticity of citations to travel time: Heterogeneity (part 1)

Result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp[\sum_d \beta_d \mathbb{1}\{distance_{ij} \in d\} \, log(\text{travel time}_{ijt}) + \sum_d \alpha_d \mathbb{1}\{distance_{ij} \in d\} \mathbb{1}\{X_{FiGjhkt}\} \, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}] \times \varepsilon_{FiGjhkt}$, 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*	
	(1)	(2)
log(travel time) × quantile 0.00	-0.152*** (0.055)	-0.119*** (0.036)
log(travel time) × quantile 0.01	-0.091 (0.111)	-0.058 (0.090)
log(travel time) × quantile 0.25	-0.078 (0.098)	-0.166* (0.088)
log(travel time) × quantile 0.50	-0.140 (0.086)	-0.076 (0.078)
log(travel time) × quantile 0.75	-0.187** (0.076)	-0.040 (0.063)
log(travel time) × quantile 0.90	-0.008 (0.061)	-0.097* (0.054)
log(travel time) × quantile 0.95	-0.022 (0.036)	-0.127*** (0.038)
log(travel time) × quantile 0.99	-0.129*** (0.033)	-0.062* (0.036)
log(travel time) × quantile 0.999	-0.078* (0.042)	-0.073* (0.043)
N obs. effective	5,147,161	5,147,161
R2	0.88	0.88

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

Table 10: 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*	
	(1)	(2)
log(travel time) × Chemical	-0.070^{*} (0.041)	-0.094^{**} (0.041)
log(travel time) × Computers & Communications	-0.075 (0.076)	-0.121^{*} (0.074)
log(travel time) × Drugs & Medical	-0.015 (0.156)	0.027 (0.173)
log(travel time) × Electrical & Electronic	-0.071 (0.046)	-0.053 (0.043)
log(travel time) × Mechanical	-0.094^{***} (0.030)	-0.096^{***} (0.030)
log(travel time) × Others	-0.128^{***} (0.043)	-0.107^{**} (0.042)
N obs. effective	5,147,161	5,147,161
R2	0.88	0.88

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

Table 12: 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	-0.095^{**} (0.047)	0.240 (0.252)	0.147 (0.194)	-0.306^{***} (0.090)	−0.040 (0.065)	−0.045 (0.064)
Computers & Communications	−0.115 (0.247)	-0.267^{***} (0.093)	−0.426 (0.966)	0.101 (0.087)	0.089 (0.139)	0.067 (0.161)
Drugs & Medical	0.221 (0.231)	0.330 (1.140)	−0.197 (0.262)	−0.505 (0.489)	−0.264 (0.345)	0.359 (0.302)
Electrical & Electronic	0.176^{**} (0.089)	0.171^{*} (0.092)	−0.146 (0.615)	-0.102^{**} (0.052)	0.094 (0.075)	−0.023 (0.076)
Mechanical	−0.058 (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 13: 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 5,147,161.

G.1.2. IV PPML: first and second stage estimation

	First stage OLS	Second stage PPML
Dep. variable:	log(travel time)	*citations*
	(1)	(2)
log(travel time IV)	0.912***	
	(0.042)	
log(travel time)		-0.161***
		(0.031)
residual 1st stage		0.101
		(0.316)
N obs. effective	$10,907,616$	$5,147,161$
R2	0.99	0.88
Within R2	0.34	

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

Table 14: Elasticity of citations to travel time: first and second stage IV PPML

The table presents the results of 2-step instrumental variables estimation of $citations_{FiGjhkt} = \exp\left[\beta\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, where $log(\text{travel time}_{ijt})$ is instrumented with $log(\text{instrumental travel time}_{ijt})$. Column (1) shows the results of the first stage regression estimated by OLS. Column (2) shows the result of the second stage regression estimated by Poisson Pseudo Maximum Likelihood , including the estimated residuals of the first stage as controls. The number of observations in the second stage estimation is smaller due to not being able to identify fixed effects that are required in PPML estimation.

	OLS First stage 0-300km	OLS First stage 300-1,000km	OLS First stage 1,000-2,000km	OLS First stage +2,000km	Second stage PPML
Dep. variable:			log(travel time)		*citations*
	(1)	(2)	(3)	(4)	(5)
log(travel time IV) × 0-300km	0.330*** (0.120)	0.049 (0.057)	0.013 (0.022)	0.037* (0.019)	
log(travel time IV) × 300-1,000km	-0.162*** (0.041)	1.08*** (0.040)	-0.014 (0.009)	0.011 (0.010)	
log(travel time IV) × 1,000-2,000km	-0.079*** (0.025)	-0.066*** (0.022)	1.06*** (0.043)	0.014 (0.009)	
log(travel time IV) × +2,000km	-0.074*** (0.024)	-0.059*** (0.019)	-0.022** (0.010)	1.10*** (0.017)	
log(travel time) × 0-300km					-0.185 (0.153)
log(travel time) × 300-1,000km					−0.155*** (0.044)
log(travel time) × 1,000-2,000km					−0.132** (0.044)
log(travel time) × +2,000km					−0.206*** (0.042)
residual × 0-300km					0.172 (0.153)
residual 300-1,000km					0.071 (0.052)
residual 1,000-2,000km					0.071 (0.066)
residual +2,000km					0.060 (0.081)
N obs. effective	10,907,616	10,907,616	10,907,616	10,907,616	5,147,161
R2	0.99	0.99	0.99	0.99	0.88
Within R2	0.06	0.45	0.80	0.88	

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

Table 15: Elasticity of citations to travel time: first and second stage IV PPML

The table presents the results of 2-step instrumental variables estimation of Poisson Pseudo Maximum Likelihood of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \times \mathbb{1}\{distance_{ij} \in d\} \times log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, where $\mathbb{1}\{distance_{ij} \in d\} \times log(\text{travel time}_{ijt})$ is instrumented with $\mathbb{1}\{distance_{ij} \in d\} \times log(\text{instrumental travel time}_{ijt})$. Given that there are 4 distance segments d there are 4 first stages. Columns (1) to (4) show the results of the first stage regressions which are estimated by OLS. Coefficients of the 4 interactions of the instrument can be identified due to the presence of the fixed effects, e.g. after demeaning by fixed effects there is residual variation that allows to identify the 4 coefficients in each regression of the first stage. Column (5) shows the result of the second stage regression estimated by PPML, including the estimated residuals of the first stage as controls. The number of observations in the second stage estimation is smaller due to not being able to identify fixed effects that are required in PPML estimation.

G.1.3. Robustness

Sample of establishments

We may be concerned that the changes in the diffusion of knowledge are purely consequence of a potential change in the geographical location of innovation activity. To rule out this possibility, in Table 16 we estimate the baseline regression 1 with different samples. In column (1) we include the baseline results.[22] 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 establishments Fi and cited establishments Gj that filed patents in 1949-1953.[23] We find that the coefficient at more than 2,000km remains comparable to the one in the baseline regression, statistically significant at the 1%.

[22] Coefficients are not bias corrected.

[23] 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.

	All	Citing establishment	Citing & Cited establishment
Dep. variable: *citations*		$cit_{FiGjhkt}$	
	(1)	(2)	(3)
log(travel time) × 0-300km	-0.014 (0.041)	-0.017 (0.045)	-0.008 (0.045)
log(travel time) × 300-1,000km	-0.095^{***} (0.025)	-0.090^{***} (0.027)	-0.087^{***} (0.028)
log(travel time) × 1,000-2,000km	-0.092^{**} (0.041)	-0.093^{**} (0.046)	-0.061 (0.049)
log(travel time) × +2,000km	-0.177^{***} (0.048)	-0.146^{***} (0.051)	-0.163^{***} (0.051)
N obs. effective	5,147,161	3,446,185	2,190,973
R2	0.88	0.88	0.89

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

Table 16: 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. Results in none of the columns is bias-corrected. 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.

Indirectly connected MSAs

If the 1951 flight network was constructed in order to connect city pairs that would see future growth in citations, we can alleviate this endogeneity concern by focusing only on indirectly connected pairs.

Table 17 presents PPML regressions not bias-corrected. Columns (1) and (2) are the baseline regressions (all MSA-pairs), columns (3) and (4) drop MSA-pairs that are ever connected with one leg (a non-stop flight), and columns (5) and (6) drop MSA-pairs that are ever connected with one flight number. The difference between non-stop and one flight number is that one flight number could serve multiple MSAs by making intermediate stops.[24] The estimated coefficients are in the ballpark of the initial estimates, especially for +2,000km, providing evidence that it is reasonable to use the pre-existing network as the baseline to construct the instrument.

[24]For example, in 1951 NYC-LA was connected with one flight number that included one stop in Chicago, that is two legs but only one flight number (passengers did not have to change airplanes.)

	PPML not bias-corrected					
Dep. variable:	*citations*					
	(1)	(2)	(3)	(4)	(5)	(6)
log(travel time)	-0.088^{***} (0.023)		-0.174^{***} (0.049)		-0.231^{***} (0.058)	
log(travel time) × 0-300km		-0.014 (0.041)		-0.253^{**} (0.113)		-0.429^{***} (0.160)
log(travel time) × 300-1,000km		-0.095^{***} (0.025)		-0.102 (0.073)		-0.203^{**} (0.084)
log(travel time) × 1000-2,000km		-0.092^{**} (0.041)		-0.161^{*} (0.087)		-0.223^{**} (0.103)
log(travel time) × +2,000km		-0.177^{***} (0.048)		-0.263^{***} (0.084)		-0.210^{***} (0.091)
N obs. effective	5,147,161	5,147,161	1,735,427	1,735,427	1,396,393	1,396,393
R2	0.88	0.88	0.94	0.94	0.94	0.94
Observation selection:						
All	X	X				
Discard one leg			X	X		
Discard one flight number					X	X

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

Table 17: Elasticity of citations to travel time: dropping directly connected MSA pairs

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) discards all ij that are ever connected with one leg (non-stop flight), while columns (5) and (6) discard all ij that are ever connected with one flight number. The difference between non-stop and one flight number is that one flight number could serve multiple MSAs by making intermediate stops. Standard errors clustered at the non-directional location are presented between parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

Estimation of log-log gravity equation

We modify equation 1 to have a log-log version:

$$log(citations_{FiGjhkt}) = \kappa\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \nu_{FiGjhkt} \quad (6)$$

Results by OLS estimation are presented in Table 18.

	PPML		OLS	
Dep. variable:	*citations*		$log(citations)$	
	(1)	(2)	(3)	(4)
log(travel time)	-0.088^{***} (0.023)		-0.056 (0.037)	
log(travel time) × 0-300 km		0.014 (0.041)		0.046 (0.061)
log(travel time) × 300-1,000 km		-0.095^{***} (0.025)		-0.072^{*} (0.042)
log(travel time) × 1,000-2,000 km		-0.092^{**} (0.041)		-0.104 (0.070)
log(travel time) × +2,000 km		-0.177^{***} (0.048)		-0.161^{*} (0.085)
N obs. effective	5,147,161	5,147,161	2,855,586	2,855,586
R2	0.88	0.88	0.99	0.99

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

Table 18: Elasticity of citations to travel time: PPML and OLS

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) shows the result of OLS estimation of $log(citations_{FiGjhkt}) = \kappa\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt} + \nu_{FiGjhkt}$. The coefficients and standard errors in columns (1) and (2) are jackknife bias-corrected. In columns (3) and (4) standard errors are clustered at the MSA-pair. 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 a sample of prices in the last page of their booklet of flight schedules a sample of prices, which we digitized. We have digitized American Airlines 1951, 1961, 1966; TWA 1951 and United Airlines 1956 and 1961.[25]. The sample includes prices for 11,590 directional airport pair years. 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 (until 1962-1963). A linear regression with an intercept and a slope estimated separately for each year (including 1966), service class (first class or coach service), and aircraft type (propeller or jet) gives a R2 of 0.98 or higher in each regression, with an average R2 of 0.993.

Second, all airlines operating within the same route charged exactly the same price. In 1951, in our digitized price data we have 432 airport pairs in which both American Airlines and TWA were operating and reported the price for first class service. 94% of those airport pairs had exactly the same price in both airlines.

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, the change in prices over time had a similar pattern until 1961: a stronger increase in short distances (probably due to an increase in fixed costs of take-off and landing, although not reflected in the intercept of the linear regressions), and a relatively constant increase for flights between airports more than 1,000 km apart. In the period 1961 to 1966 we observe a drop in prices of around 20% for routes of more than 1,000km distance, breaking the linearity of prices in distance previously observed. We had vi-

[25]The sample of prices digitized was limited due to data availability.

sually inspected price tables and detected that the drop in prices happened in 1962-1963.

Figure 36 shows prices for first class service by year and aircraft type, deflated by the consumer price index to 1951 values. Figure 37 presents the percentage change in deflated prices of first class service. Both figures show the previous facts: prices are generally linear in distance until 1966 in which we observe a break after 1,000 km.

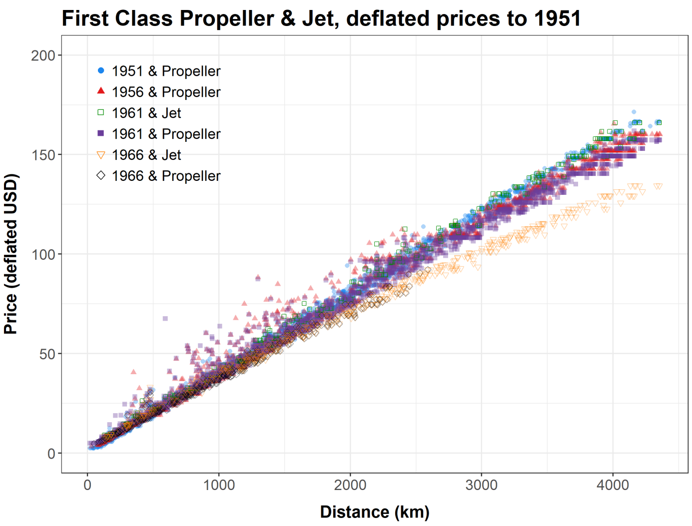


Figure 36: Flight ticket prices, deflated by CPI

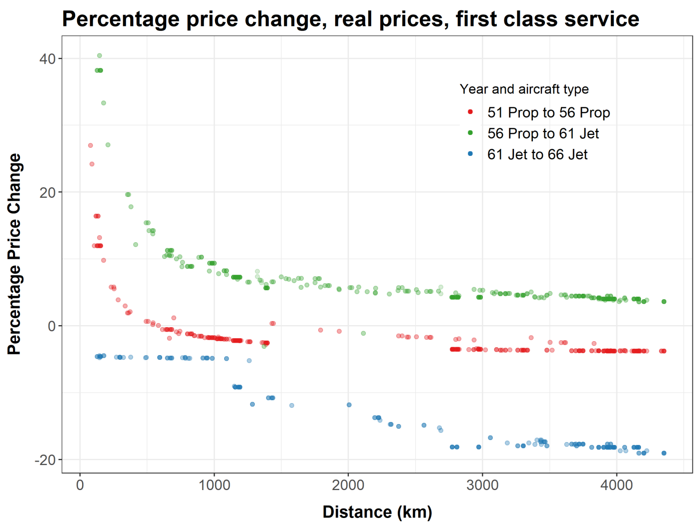


Figure 37: Change flight ticket prices, deflated by CPI

We convert our sample of prices at the airport-pair level to prices of the population of MSA-pairs as follows: first, we obtain a pricing function that can flexibly approximate prices by regressing deflated prices on a cubic polynomial of distance separately for each year. We use prices of first class service for all years, propeller aircraft for 1951 and 1956 and jet aircraft for 1961 and 1966. Second, we predict prices for each MSA-pair and year using the MSA-pair distance and the year's estimated regression.

Highway travel time

Taylor Jaworski and Carl Kitchens have graciously shared with us data on county-to-county highway travel time and nominal travel costs for 1950, 1960 and 1970. Travel time is constructed using maximum speed limit in each highway segment and year. Travel costs uses, for each year, travel time, highway distance, truck driver's wage and petrol costs. See Jaworski and Kitchens (2019) for details. The dataset is constructed using 2010 county boundaries and contains county centroids. We converted it to MSA-to-MSA by matching counties' centroids to 1950 MSAs using the shape file from Manson et al. (2020). We take the minimum travel time and minimum travel costs among all county pairs that belong to the same MSA pair. We convert nominal travel costs to 1950 real travel costs deflating by the consumer price index. We convert 1950, 1960 and 1970 travel times and travel costs to 1951, 1956, 1961 and 1966 by linearly interpolating (e.g. $\text{travel time}_{ij,1951} = \text{travel time}_{ij,1950} \times \frac{1960-1951}{10} + \text{travel time}_{ij,1960} \times \frac{1951-1950}{10}$).

The within MSA-pair correlation of the 1951-1966 change in travel time by highway and airplane is 0.068 for all MSA-pairs, and -0.011 for MSA-pairs more than 2,000 km apart. Figure 38 presents the MSA-pair 1951-1966 change in travel time by highway and airplane, where for exposition we only present MSA-pairs that had a reduction in travel time by both means of transport. Estimating a linear regression of change in air travel time on the change in highway travel time gives a slope of -0.02 not statistically different from zero, with a R2 of 0.00005.[26] Figure 39 repeats the exercise where MSA-pairs are weighted by the amount of establishment-technology pairs used to estimate the elasticity of citations to travel time (equation (1)). In this case the estimated regression has a slope of 0.73 statistically significant at the 1% level and a R2 of 0.09.[27]

In Tables 4 and 19 we present the results of adding highway travel time as control. The low correlation between the change in travel time by highway and airplane implies that the estimated elasticity of citations to air travel time remains almost unchanged,

[26] 8.7% of MSA-pairs had an increase in travel time either by highway or by airplane. The regression with all MSA-pairs has a slope of 0.60 significant at the 1% level. However, the R2 of the regression remains very low: 0.0046.

[27] With all MSA-pairs the slope is 1.01 statistically significant at the 1% level and the R2 is 0.04.

relative to the baseline estimation.[28]

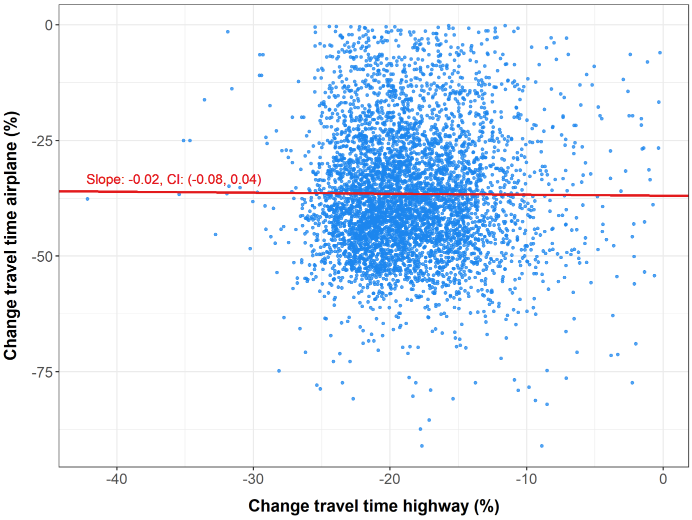


Figure 38: Change travel time by airplane and highway 1951-1966

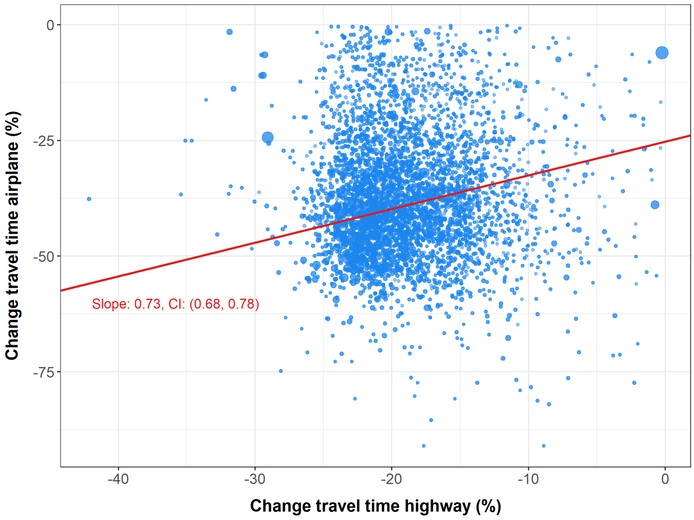


Figure 39: Change travel time by airplane and highway 1951-1966, weighted

[28]In order to perform a test of statistical difference of coefficients we would need to compute the covariance between the two regressions. Assuming the covariance is zero, in columns (1) and (2) 19 the coefficients of air travel time at +2,000km are not significantly different.

	PPML							
Dep. variable:	*citations*							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log(travel time) × 0-300km	-0.014 (0.041)	-0.013 (0.042)	-0.016 (0.041)	-0.004 (0.042)	-0.015 (0.042)	-0.004 (0.042)	-0.007 (0.042)	-0.007 (0.042)
log(travel time) × 300-1,000km	-0.095*** (0.025)	-0.091*** (0.026)	-0.091*** (0.025)	-0.066** (0.029)	-0.089*** (0.026)	-0.070** (0.029)	-0.067** (0.029)	-0.067** (0.029)
log(travel time) × 1000-2,000km	-0.092** (0.041)	-0.086** (0.043)	-0.075* (0.041)	-0.037 (0.050)	-0.073* (0.043)	-0.037 (0.050)	-0.032 (0.050)	-0.032 (0.050)
log(travel time) × +2,000km	-0.177*** (0.048)	-0.170*** (0.050)	-0.167*** (0.049)	-0.112** (0.056)	-0.164*** (0.051)	-0.112** (0.056)	-0.112** (0.057)	-0.112** (0.057)
N obs. effective	5,147,161	5,147,161	5,147,161	5,147,161	5,147,161	5,147,161	5,147,161	5,147,161
R2	0.88	0.88	0.88	0.88	0.88	0.88	0.88	0.88
Controls:								
log(highway time)	-	Yes	-	-	Yes	Yes	-	Yes
log(telephone share) × time	-	-	Yes	-	Yes	-	Yes	Yes
log(distance) × time	-	-	-	Yes	-	Yes	Yes	Yes

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

Table 19: Elasticity of citations to travel time: additional controls

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $citations_{FiGjhkt} = \exp\left[\sum_d \beta_d \mathbb{1}\{distance_{ij} \in d\}\, log(\text{travel time}_{ijt}) + FE_{FiGjhk} + FE_{Fiht} + FE_{Gjkt}\right] \times \varepsilon_{FiGjhkt}$, 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.

Frequency adjusted travel time

The frequency of flights may have changed simultaneously with the introduction of jet airplanes. The change in travel time could then be consequence of higher frequency rather than changes in airplanes' speed. Given that some MSA pairs are connected indirectly (with connecting flights), accounting for frequency is not straight forward: the frequency of each leg of the flight route matters (actually, it is not only frequency of each leg but also the synchronization among all potential legs). In order to take into account potential changes in the frequency of flights we computed the daily average travel time. This travel time is the average across all fastest travel times if the passenger was to depart at each full hour (1am, 2am, ..., 1pm, 2pm, etc.). The computation of this travel time includes the waiting time that is affected by frequency: the time until first departure and layover time of each connecting flight. Hence, the daily average travel time is a frequency-adjusted travel time: changes in the daily average travel time that are larger than in the fastest travel time denote that frequency of flights increased and therefore there is less waiting time. If we observe the reverse that means that frequency did not improve as much as the speed of airplanes.

Figure 40 shows the within MSA-pair decrease in the fastest travel time and the daily average travel time.[29] Both measures of travel time follow a similar pattern: slight decrease in 1956, a stronger decrease in 1961 especially for long distance routes, and a further decline in 1966. However, we observe that the decrease of the fastest travel time is on average larger than the one of the daily average travel time: the frequency of flights, if any, attenuated the potential decrease in travel time from the improvements in airplanes' speed. This observation is also in line with a comparison of the fastest travel time with and without layover time (Figure 28 in the Appendix of the paper): layover time attenuated the change in travel time.

In table 20 we estimated the elasticity of citations to travel time using first the fastest travel time (baseline, columns 1 and 2) and the daily average travel time (columns 3 and 4). The estimated elasticity is similar using both measures, which gives confidence that our results are not driven by changes in the frequency of flights.

[29]The within MSA-pair correlation of the (1951-1966) change in fastest travel time and the change daily average travel time is 0.60.

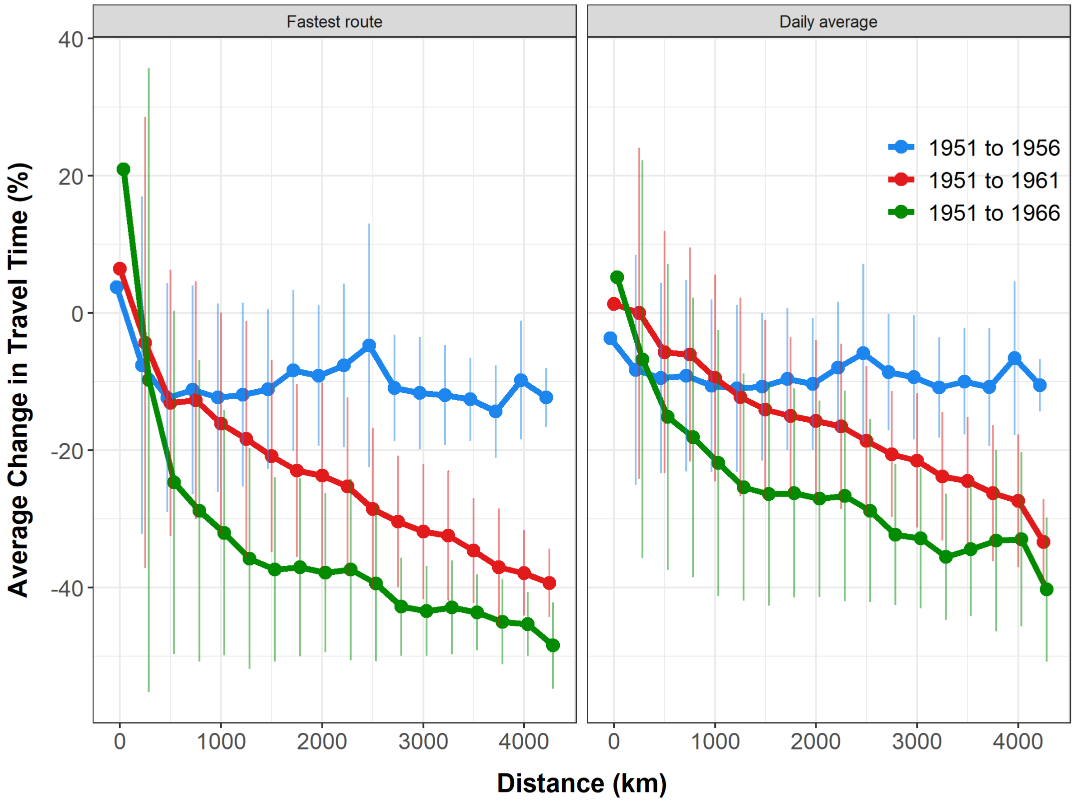


Figure 40: Change in MSAs travel time: fastest travel time and daily average travel time

	PPML not bias-corrected			
Dep. variable:	*citations*			
	(1)	(2)	(3)	(4)
log(travel time)	-0.088^{***} (0.023)			
log(travel time) × 0-300km		-0.014 (0.041)		
log(travel time) × 300-1,000km		-0.095^{***} (0.025)		
log(travel time) × 1000-2,000km		-0.092^{**} (0.041)		
log(travel time) × +2,000km		-0.177^{***} (0.048)		
log(travel time daily avg)			-0.093^{***} (0.034)	
log(travel time daily avg) × 0-300km				-0.001 (0.039)
log(travel time daily avg) × 300-1,000km				-0.097^{**} (0.038)
log(travel time daily avg) × 1000-2,000km				-0.166^{**} (0.069)
log(travel time daily avg) × +2,000km				-0.220^{***} (0.063)
N obs. effective	5,147,161	5,147,161	5,147,161	5,147,161
R2	0.88	0.88	0.88	0.88

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

Table 20: Elasticity of citations to travel time: daily average 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) use the daily average travel time, which is computed as the average of the fastest travel time departing at every full hour (the average across all 24 potential departing times). Standard errors clustered at the non-directional location are presented between parentheses (ij is the same non-directional location pair as ji). R2 is computed as the squared correlation between observed and fitted values.

G.2. Creation of knowledge

G.2.1. Heterogeneous effects

Dependent Variable:	Patents		Patents quality weighted	
	(1)	(2)	(3)	(4)
log(knowledge access)	9.1*** (3.3)		10.2*** (3.6)	
log(knowledge access quality weighted)		6.9** (2.9)		7.8*** (3.1)
R2	0.85	0.85	0.86	0.86
N obs. effective	991,480	991,480	991,284	991,284

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

Table 21: Effect of knowledge access on patents, quality weighted

Column (1) shows the result of Poisson Pseudo Maximum Likelihood (PPML) estimation of $\text{Patents}_{Fiht} = \exp[\rho\ log(\text{KA}_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}] \times \xi_{Fiht}$, 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. Columns (1) and (2) use number of patents as dependent variable while columns (3) and (4) quality-weighted patents. Columns (1) and (3) use $log(\text{KA}_{iht})$ as explanatory variable while columns (2) and (4) use a quality weighted $log(\text{KA}_{iht})$. Quality weights are the 5-year percentile of quality measure after demeaning by year fixed effects computed in Kelly et al. (2021). Weighting by the 10-year percentile of quality gives similar results. Standard errors clustered at the location-technology level *ih* are presented in parentheses. R2 is computed as the squared correlation between observed and fitted values.

G.2.2. IV PPML: centering instrumental knowledge access

The objective of the recentered instrument is to clean any non-random variation that may be mechanically introduced due to geography. Locations that are geographically far from the initial innovation centers are more likely to have a larger increase in knowledge access with the roll out of jet airplanes, in any realization of the flight network. In order to purge out this potentially non-random variation, we compute the expected value of the instrument considering multiple alternative flight networks and subtract it from the *realized* instrument.

We construct the expected instrument $\mathbb{E}[log(\widetilde{KA}_{iht})]$ as follows:

1. Count the amount of airport-pairs connected by a non-stop flight in 1951, label this the *number of 1951 connections*.
2. Set a new seed number for random draws.
3. For each unique origin airport present in 1951, create a counterfactual connection by drawing a random destination airport (different to the origin) present in 1951.

Repeat as many times until the amount of unique counterfactual connections is equal to the *number of 1951 connections*. As *number of 1951 connections* is larger than the number of origins and destinations, some origins and destinations will be repeated.

4. Check if all 1951 origin and destination airports are present in the randomized connections. Some destinations may not be present due to the random draws. If some destination (origin) is missing, drop a counterfactual connection of a destination (origin) that has at least two origins (destinations). Draw a new random connection for the missing destination (origin). Repeat this step until all origins and destinations are present in the counterfactual network.

5. Check if the counterfactual network is a connected set (i.e. it would be possible to route from any airport to any other airport through intermediary connections). If it is not a connected set, drop this iteration of the counterfactual network and go back to step 2.

6. Predict flight duration of each counterfactual connection in each year using airport-to-airport distance and the estimated intercept and slope of each year

7. Using the predicted travel time in 1951, compute the fastest path between each airport pair, including directly and indirectly connected airport pairs

8. Match airports to MSAs. For each MSA-pair, get the airport-pair and the respective path that has the minimum travel in 1951

9. Using the predicted flight duration for each counterfactual connection in each year, compute the counterfactual travel time of the 1951-optimal path for each MSA-pair and year

10. Repeat steps 2 to 9 for 2,000 times

11. With each counterfactual network, compute the counterfactual knowledge access of each MSA-technology-year.

12. Obtain the expected instrument $\mathbb{E}[log(\widetilde{KA}_{iht})]$: within each MSA-technology-year, compute the across-counterfactual network average of the log counterfactual knowledge access

We then recenter the instrument as follows:

$$log(\widetilde{KA}_{iht})_{centered} = log(\widetilde{KA}_{iht}) - \mathbb{E}[log(\widetilde{KA}_{iht})] \quad (7)$$

G.2.3. IV PPML: first and second stage estimation, non-centered instrument

	First stage OLS	Second stage PPML
Dep. variable:	log(knowledge access)	*Patents*
	(1)	(2)
log(knowledge access instrument)	1.01^{***} (0.03)	
log(knowledge access)		10.20^{*} (5.60)
residual		-2.29 (6.15)
N obs. effective	991,480	991,480
R2	0.99	0.85
Within R2	0.53	

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

Table 22: Elasticity of patents to knowledge access: first and second stage IV PPML, non-centered instrument

The table presents the results of 2-step instrumental variables estimation of $\text{Patents}_{Fiht} = \exp[\rho\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}] \times \xi_{Fiht}$, where $log(KA_{iht})$ is instrumented with $log(\widetilde{KA}_{iht})$. Column (1) shows the results of the first stage regression estimated by OLS. Column (2) shows the result of the second stage regression estimated by Poisson Pseudo Maximum Likelihood , including the estimated residuals of the first stage as controls.

	OLS First stage reference quartile	OLS First stage 3rd quartile	OLS First stage 2nd quartile	OLS First stage 1st quartile	Second stage PPML
Dep. variable:		log(knowledge access)			*Patents*
	(1)	(2)	(3)	(4)	(5)
log(knowledge access instrument)	1.00^{***} (0.03)	0.01 (0.06)	0.03 (0.03)	0.00 (0.01)	
log(knowledge access instrument) × 3rd quartile	0.01 (0.01)	1.12^{***} (0.03)	−0.00 (0.01)	−0.00 (0.01)	
log(knowledge access instrument) × 2nd quartile	0.00 (0.01)	−0.02 (0.04)	1.13^{***} (0.03)	−0.01 (0.01)	
log(knowledge access instrument) × 1st quartile	0.01 (0.01)	0.01 (0.04)	−0.04 (0.04)	1.16^{***} (0.04)	
log(knowledge access)					9.39^{*} (5.62)
log(knowledge access) × 3rd quartile					2.10^{***} (0.65)
log(knowledge access) × 2nd quartile					3.79^{***} (1.01)
log(knowledge access) × 1st quartile					5.20^{***} (1.31)
residual					−2.37 (6.18)
residual × 3rd quartile					-2.30^{*} (1.29)
residual × 2nd quartile					-3.78^{**} (1.86)
residual × 1st quartile					-7.19^{**} (3.09)
N obs. effective	991,480	991,480	991,480	991,480	991,480
R2	1.00	1.00	1.00	1.00	0.85
Within R2	0.53	0.89	0.90	0.90	

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

Table 23: Elasticity of patents to knowledge access: first and second stage IV PPML

The table presents the results of 2-step instrumental variables estimation of $\text{Patents}_{Fiht} = \exp\left[\sum_q \rho_q \times \mathbb{1}\{quartile_{ih} = q\} \times log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, where $log(KA_{iht})$ is instrumented with $log(\widetilde{KA}_{iht})$. Column (1) to (4) show the results of the first stage regression estimated by OLS. Coefficients of the 4 interactions of the instrument can be identified due to the presence of the fixed effects, e.g. after demeaning by fixed effects there is residual variation that allows to identify the 4 coefficients in each regression of the first stage. Column (5) shows the result of the second stage regression estimated by Poisson Pseudo Maximum Likelihood, including the estimated residuals of the first stage as controls.

G.2.4. IV PPML: first and second stage estimation, centered instrument

	First stage OLS	Second stage PPML
Dep. variable:	log(knowledge access)	*Patents*
	(1)	(2)
centered log(knowledge access instrument)	1.24^{***} (0.06)	
log(knowledge access)		10.22^{*} (5.63)
residual		2.15 (5.89)
N obs. effective	991,480	991,480
R2	0.99	0.85
Within R2	0.48	

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

Table 24: Elasticity of patents to knowledge access: first and second stage IV PPML, centered instrument

The table presents the results of 2-step instrumental variables estimation of $\text{Patents}_{Fiht} = \exp\left[\boldsymbol{\rho}\, log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, where $log(KA_{iht})$ is instrumented with centered $log(\widetilde{KA}_{iht})$. Column (1) shows the results of the first stage regression estimated by OLS. Column (2) shows the result of the second stage regression estimated by Poisson Pseudo Maximum Likelihood , including the estimated residuals of the first stage as controls.

	OLS First stage reference quartile	OLS First stage 3rd quartile	OLS First stage 2nd quartile	OLS First stage 1st quartile	Second stage PPML
Dep. variable:		log(knowledge access)			*Patents*
	(1)	(2)	(3)	(4)	(5)
centered log(knowledge access instrument)	1.24*** (0.06)	0.36 (0.31)	0.31** (0.15)	0.01 (0.03)	
centered log(knowledge access instrument) × 3rd quartile	-0.01 (0.01)	-1.61*** (0.21)	0.07** (0.03)	0.00 (0.00)	
centered log(knowledge access instrument) × 2nd quartile	0.00 (0.01)	0.04 (0.25)	-1.49*** (0.19)	0.03 (0.02)	
centered log(knowledge access instrument) × 1st quartile	-0.01 (0.02)	0.03 (0.23)	0.47** (0.21)	-1.91*** (0.18)	
log(knowledge access)					7.61 (5.60)
log(knowledge access) × 3rd quartile					3.54*** (1.06)
log(knowledge access) × 2nd quartile					6.89*** (2.06)
log(knowledge access) × 1st quartile					8.05*** (2.21)
residual					0.27 (5.86)
residual × 3rd quartile					−2.54** (1.12)
residual × 2nd quartile					−5.04** (2.10)
residual × 1st quartile					−5.17** (2.28)
N obs. effective	991,480	991,480	991,480	991,480	991,480
R2	1.00	1.00	1.00	1.00	0.85
Within R2	0.48	0.26	0.30	0.43	

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

Table 25: Elasticity of patents to knowledge access: first and second stage IV PPML, centered instrument

The table presents the results of 2-step instrumental variables estimation of $\text{Patents}_{Fiht} = \exp\left[\sum_q \rho_q \times \mathbb{1}\{quartile_{ih} = q\} \times log(KA_{iht}) + FE_{Fih} + FE_{it} + FE_{ht}\right] \times \xi_{Fiht}$, where $log(KA_{iht})$ is instrumented with centered $log(\widetilde{KA}_{iht})$. Column (1) to (4) show the results of the first stage regression estimated by OLS. Coefficients of the 4 interactions of the instrument can be identified due to the presence of the fixed effects, e.g. after demeaning by fixed effects there is residual variation that allows to identify the 4 coefficients in each regression of the first stage. Column (5) shows the result of the second stage regression estimated by Poisson Pseudo Maximum Likelihood, including the estimated residuals of the first stage as controls.

G.2.5. Robustness

	Baseline	Quartile absolute	Quartile per capita
Dependent Variable:		*Patents*	
	(1)	(2)	(3)
log(knowledge access)	9.11*** (3.29)	8.41** (3.31)	6.98** (3.33)
log(knowledge access) × quartile 0.50		1.86*** (0.53)	0.68** (0.30)
log(knowledge access) × quartile 0.25		3.42*** (0.81)	1.42*** (0.46)
log(knowledge access) × quartile 0.00		4.50*** (1.17)	3.63*** (0.69)
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 26: 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.

	PPML		β by distance		+300km		+1,000km		+2,000km	
Dependent Variable:	*Patents*									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log(knowledge access)	9.11*** (3.29)	8.41** (3.31)	13.98*** (4.10)	12.71*** (4.19)	8.97** (4.19)	7.73* (4.19)	16.98*** (5.25)	17.21*** (5.18)	11.42 (7.38)	9.22 (7.15)
log(knowledge access) × 3rd quartile		1.86*** (0.53)		2.11*** (0.63)		1.92*** (0.52)		1.87*** (0.48)		1.75*** (0.44)
log(knowledge access) × 2nd quartile		3.42*** (0.81)		4.06*** (1.03)		3.78*** (0.80)		3.58*** (0.73)		3.29*** (0.66)
log(knowledge access) × 1st quartile		4.50*** (1.17)		5.67*** (1.49)		4.95*** (1.13)		4.77*** (1.11)		4.22*** (0.97)
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 27: 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*		$log(Patents)$	
	(1)	(2)	(3)	(4)
log(knowledge access)	9.11*** (3.29)	8.41** (3.31)	6.16** (2.87)	5.69** (2.88)
log(knowledge access) × 3rd quartile		1.86*** (0.53)		0.83* (0.46)
log(knowledge access) × 2nd quartile		3.42*** (0.81)		2.38** (0.93)
log(knowledge access) × 1st quartile		4.50*** (1.17)		3.45** (1.62)
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 28: 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.

Access to capital

We construct four measures of access to capital using 1949-1953 market capitalization of firms listed in the stock market. The four measures are similar in their essence but differ in the computation of a firm's technology and the firm's location. The measure is computed as follows:

$$\textit{capital access}_{iht} = \sum_k \psi_{hk} \sum_{j,j \neq i} \text{Capital stock}_{jk,t=1951} \times \text{travel time}_{ijt}^{\xi} \tag{8}$$

where Capital $\text{stock}_{jk,t=1951}$ is a proxy for the capital which is specific to technology k located in j at the initial time period 1951. ψ_{hk} is an input-output weight of capital flows and ξ is the elasticity of capital flows between to travel time. As a proxy for capital we use market capitalization of firms.

We construct four measures of $\textit{capital access}_{iht}$ which differ on: (i) the way we define the allocation of the firm's capital to each location (either using all inventors' locations or only the assigned headquarters), and (ii) the way we allocate a firm's capital across technologies (using the share of a technology within the firm, or relative to the national share of that technology). We use COMPUSTAT as our source of data for market capitalization.

We proceed as follows:

1. Use share's market price at closure calendar year multiplied by the number shares outstanding. We use the variables *prcc_c* and *csho* to maximize coverage of firms given that other variables have missing value for many firms.

2. Take the yearly average market capitalization to maximize coverage (many firms have missing in a certain year). This step potentially introduces measurement error due to changes in total stock market capitalization but allows us to increase the amount of firms included in the sample.

3. Determine a firm's MSA using patent inventor location. Two ways to determine

the location, 1. only HQ location, 2. all locations where the firm had inventors applying for patents in 1949-1953

4. Determine the share of each technology within the firm using patent data. Two ways to determine the share of technology: 1. The share of each technology within firm, 2. The share of each technology within firm relative to national share

5. In the absence of data on a capital input-output weight, assume it is the same as the technology input-output weight, i.e. $\psi_{hk} = \omega_{hk}$

6. In the absence of data on the elasticity of capital flows to travel time assume $\xi = -1$

The four measures of access to capital are as follows:

1. Attribute all capital to headquarters and use the absolute share of each technology in the firm

2. Attribute all capital to headquarters and use the share of each technology in the firm relative to the national share

3. Attribute capital to establishments using their pat share and use the absolute share of each technology in the firm

4. Attribute capital to establishments using their pat share and use the share of each technology in the firm relative to the national share

Table 29 shows the results of estimating the elasticity of new patents to knowledge access while at the same time controlling for capital access.

Dependent Variable:	*Patents*								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log(knowledge access)	9.11*** (3.29)					8.97** (4.05)	10.17*** (3.88)	9.61** (4.23)	11.64*** (3.99)
log(finance access hq)		0.54** (0.26)				0.02 (0.30)			
log(finance access hq rel)			0.40 (0.25)				-0.14 (0.28)		
log(finance access est)				0.56* (0.31)				-0.07 (0.39)	
log(finance access est rel)					0.31 (0.30)				-0.40 (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 29: 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.

Sensitivity to β

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
-0.206	9.11	-1.88	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
-1.0	1.72	-1.72	3.51	0.79	1.28	0.24
-2.0	0.58	-1.16	2.8	0.63	1.55	0.3
-5.0	0.04	-0.19	1.19	0.27	3.65	0.7
-8.0	0.09	-0.76	8.22	1.84	6.96	1.33
-10.0	0.11	-1.08	15.16	3.4	8.19	1.56
-20.0	0.13	-2.63	69.8	15.65	21.66	4.14
-50.0	0.16	-8.22	531.34	119.16	219.49	41.94
-100.0	0.12	-12.33	5428.85	1217.49	2971.74	567.91

Table 30: Effect of knowledge access on new patents: varying the value of elasticity of knowledge diffusion

The table shows for different values of β (column 1) used to compute knowledge access, the estimated value of ρ (column 2). Columns 3 shows that for β values in the range estimated, the multiplication of β and ρ remains stable. Column 4 shows the predicted yearly growth rate of patenting, averaged across MSAs, and column 5 divides the predicted value of column 4 by the observed value. Column 6 shows, using quartile-specific coefficients ρ, the predicted differential yearly growth rate between MSAs in the lowest and highest quartile of initial innovativeness. Column 7 shows the ratio between the predicted value of column 6 and the observed value.