The Geography of Knowledge Production: Connecting Islands and Ideas*

Andrew B. Bernard† Andreas Moxnes‡ Yukiko U. Saito§

November 2019

Preliminary

Abstract

This paper examines the importance of economic integration on the flow of ideas. During the late 1980s and -90s, one of Japan's four islands, Shikoku, was for the first time connected to the largest island, Honshu, by three different bridges. This led to a large reduction in travel time compared to existing modes of transport (ferry and air). We examine the impact of bridge connections on team formation and the production of knowledge, as measured by patent data. Using the geolocation of inventors before the opening of the bridges, we find that inventors close to future bridges increased knowledge production more than inventors far from the bridges. Furthermore, the treated inventors matched to more productive inventors, as measured by co-inventors' existing patent portfolio. The results are robust to instrumenting for the location of the bridges using the minimum bridge span distances between Shikoku and Honshu. Our results suggest that economic integration can have sizable effects on idea flows and consequently innovation and growth.

JEL: F14, D22, D85, L10, L14, R12

Keywords: trade, knowledge flows, innovation, patenting, infrastructure.

*This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreements No 715147).

†Tuck School of Business at Dartmouth, CEP, CEPR & NBER; andrew.b.bernard@tuck.dartmouth.edu

‡Princeton University, University of Oslo & CEPR; andreamo@econ.uio.no

§Waseda University & RIETI; saito-yukiko@waseda.jp

1 Introduction

Innovation and invention are key components of economic growth for advanced economies. In spite of the popular archetype of the lone brilliant inventor, most inventions are the output of teams of collaborators.[1] More than 70 percent of patents at the European Patent Office in 2010 were produced by teams of at least two inventors (Akcigit et al., 2018) and more than 75 percent of papers published in the top five economics journals in 2018 had multiple authors. At the same time that innovative activity is increasingly done in teams, the members of those teams are often geographically distant from one another. The structure of the modern economy means that the costs of interacting at a distance are lower than ever before and continue to fall. This paper provides new evidence on the role of reduced travel times and increased integration on the structure and performance of research teams producing patents.

From 1988 to 1999, Japan completed three large bridge projects providing the first physical links between Honshu and Shikoku, the largest and and fourth largest Japanese islands respectively. The bridges were tremendous feats of engineering, taking over a decade to complete and including sections that are still among the longest bridges in the world. These bridges dramatically reduced the travel time between Shikoku and nearby major Japanese cities on Honshu including Osaka, Kobe and Hiroshima. We use the opening of these three bridges as a quasi natural experiment to examine the role of travel time on collaboration and performance of geographically distant inventors.

There is substantial evidence that distance plays a role in inhibiting the spread of knowledge. Jaffe et al. (1993) show citations of patents are geographically localized; Comin et al. (2012) find new technology adoption decreases in distance from the location of the invention; Head et al. (2019) use citation counts in mathematics to argue that the role of geographic distance in inhibiting knowledge flows has declined over time. While these studies focus on the ability of researchers to acquire knowledge at a distance, they do not consider the causal link between geography and the structure and performance of research teams.

A separate group of papers focuses on the research teams. Wuchty et al. (2007) find that research is increasingly done in teams and teams produce the highest quality papers and patents. Akcigit et al. (2018) use patent data on European patents and show that most patents are produced in teams and that interactions with more productive inventors yield higher subsequent productivity. Jones et al. (2008) argue that research increasingly crosses university boundaries and that elite universities are driving this change with high quality research becoming more concentrated rather than less. Agrawal and Goldfarb (2007)

[1] Wuchty et al. (2007) examine scientific publications and patent data and find that most research is done in teams that that the share of team is rising over time.

examine a reduction in communications cost due to the rollout of Bitnet, a precursor to the Internet, and find an increase in multi-institutional collaborations. Most closely related to our research is recent work by Catalini et al. (2018) who find that the rollout of low cost air routes in the US dramatically increased output by existing research pairs. Our work builds on the literature and confirms the rise of teams in knowledge production in Japan as well as the productivity advantage of larger teams. Our main focus differs from these papers in that we examine a reduction in travel time and the formation of new teams by geographically remote inventors.

Rather than focusing on the use of distant ideas in production, we consider the role of geography on the production team itself and the quantity and quality of its output. As with many other studies on geography and knowledge, we use the number of patents by an inventor (or team) as the measure of the output of knowledge production and the citations of each patent as a measure of its quality.

The opening of the bridges had a significant positive effect on patenting activity and the structure of research teams. The number of patents applications filed by inventors on Shikoku rose relatively rapidly after the bridge opening. The increase in patent applications was largest for inventors near one end of the bridge.

Using the inventor information on the Japanese patent applications, we start by documenting several stylized facts. As with many social interactions there is positive assortative matching between collaborators on individual patents: better inventors, as measured by either their number of previous patents or citation-weighted counts on previous patents, match with each other. Innovation quality is increasing in the size of the team and in the quality of the team members. Finally, more productive inventors are part of larger and more geographically dispersed teams.

To examine the effect of the bridges on inventor activity, we estimate a simple difference-in-differences model where the first difference is activity before and after the bridge opening and the second difference is the distance of the inventor to the bridge. In addition we instrument for the location of the bridges using the minimum bridge span distances between Shikoku and Honshu. The effects of the bridge are positive and significant for both inventor output, citation-weighted patents, and the quality of co-inventors and the probability of stopping patenting is lower after the opening of the bridges for inventors close to the bridge.

2 Data, Measurement and Stylized Facts

2.1 Data and Measurement

Data. The main dataset is the universe of patent applications from the Japanese Patent Office (JPO) over the time period 1981 to 2005. The patent data includes the name and work addresses of the applicant(s) and inventor(s) of each patent, the date the patent is filed and granted, and the citations of previous patents. From the raw data, we geocode the location of all domestic inventors. From the geocode, we allocate each inventor to a cell in a 10 km $\times$ 10 km mesh provided by the Ministry of Internal Affairs and Communications (MIC). For foreign inventors, we observe the country of residence.

We use the citation data to calculate the number of future citations of each patent. For the main analysis, we use the 10-year citation count, i.e. the number of citations a patent p receives within the first 10 years after publication, c_p.[2] The unique identity of an inventor is based on the name in the patent application. This leads to measurement error (i) if inventors in the same geographic mesh have identical names leading to multiple inventors being grouped into one inventor id and overcounting patents, and (ii) if a unique inventor appears with different spelling on different patent applications, so that the same inventor receives multiple inventor ids and undercounting patents. Appendix 5 provides more details on the construction of the data set.

Knowledge productivity. We measure the knowledge productivity of an inventor i as the cumulative (citation-weighted) number of patents i has worked on up until year t,

$$z_{it} = \sum_{s=1981}^{t} \sum_{p \in \mathbb{P}_{is}} c_p, \tag{1}$$

where $\mathbb{P}_{is}$ is the set of patents that i is filing in year s.

Co-inventor productivity. We measure match quality as the leave-out mean lagged z of the co-inventors of i in year t,

$$\bar{z}_{it} = \frac{1}{\sum_{s \in \mathbb{P}_{it}} (t_s - 1)} \sum_{p \in \mathbb{P}_{it}} \sum_{j \in \mathbb{I}_p \setminus i} z_{jt-1}, \tag{2}$$

where $\mathbb{I}_p$ is the set of inventors on patent p and t_p is the team size (number of inventors) of patent p. z_{jt-1} is the knowledge productivity of co-inventors j *prior to* joining the team with inventor i.

[2] Patent examiners can also add citations to a patent application. These additional citations are excluded in the baseline data set.

The geography of collaboration. We measure the distance to co-inventors as the leave-out mean log distance to co-inventors of i in year t,

$$\ln \bar{d}_i = \frac{1}{\sum_{s \in \mathbb{P}_{it}} (t_s - 1)} \sum_{p \in \mathbb{P}_{it}} \sum_{j \in \mathbb{I}_p \setminus i} \ln Distance_{ijp}, \tag{3}$$

where $Distance_{ijp}$ is the geodesic distance (in km) between inventors i and j based on their geolocation inferred from patent p.[3]

2.2 Stylized Facts

Fact 1. *There is positive assortative matching among inventors.*

A defining feature of the matching between inventors is that highly productive inventors tend to match with other highly productive inventors. Figure 1 shows a binned scatterplot of inventor productivity z_i (x-axis) against co-inventor productivity $\bar{z}_i$ (y-axis), both in logs. The regression (and plot) includes mesh fixed effects (see Section 2.1), so that we compare productivity differences across inventors in the same 10 km $\times$ 10 km location. The left figure uses the citation-weighted measure of z_i and $\bar{z}_i$, while the right figure uses simple patent counts instead. The relationship is strikingly linear in logs; a ten percent increase in inventor productivity is associated with a 2.2% increase in co-inventors' productivity.

Fact 2. *Productive inventors work in larger teams.*

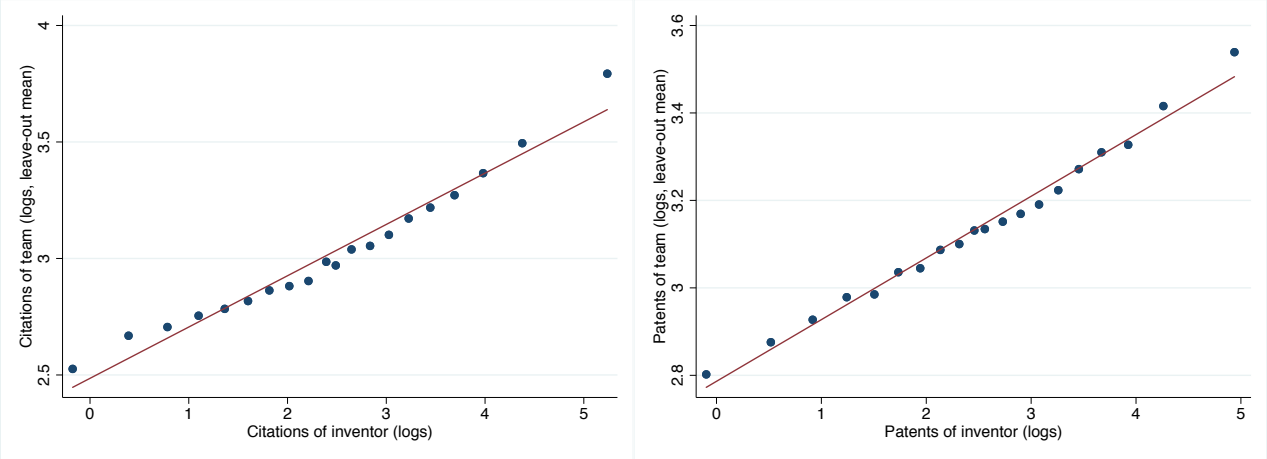
The productivity of inventors is increasing in the size of the research team (i.e., the number of co-inventors). Figure 2 shows a binned scatterplot of inventor productivity z_i (x-axis) against team size t_i (y-axis), both in logs. We find a strong linear relationship between the two; a ten percent increase in inventor productivity is associated with 0.5 percent larger team size.

Fact 3. *Innovation quality rises with the number and productivity of inventors.*

Next, we investigate the association between the inputs into innovation - the number of inventors and the productivity of those inventors - and the output of innovation as proxied by the future citations of a patent. The left binned scatterplot in Figure 3 plots the average inventor productivity z_i across the inventors on patent p on the x-axis against the future citations of patent p, c_p, on the y-axis, both in logs. The right binned scatterplot instead uses the team size of patent p on the x-axis. In both cases, we find that larger and better teams on a patent are associated with better innovations, as proxied by the future citation count.

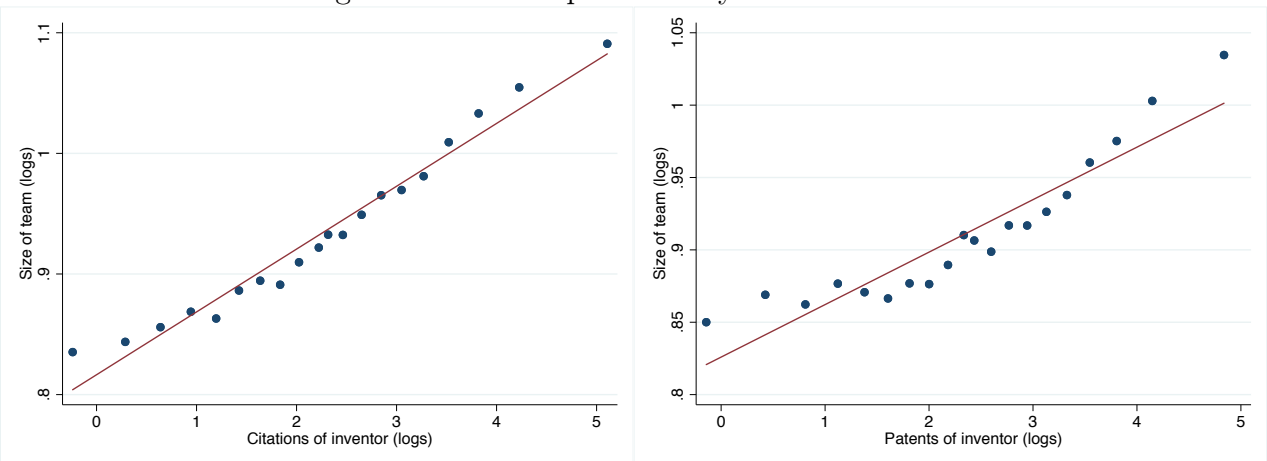
[3] We set $\ln Distance_{ijp} = 0$ in those cases where $Distance_{ijp} = 0$, so that co-located inventors do not drop out of the calculation.

Figure 1: Inventor productivity and team productivity.



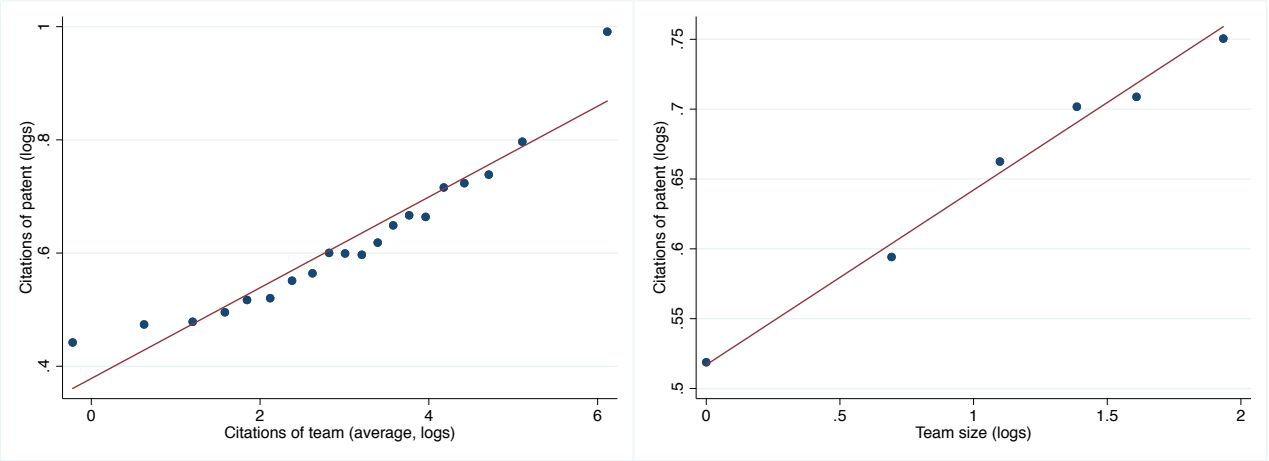
Note: The binned scatterplots group inventors into 20 equal-sized bins according to the variable on the x-axis and compute the mean of the variables on the x- and y-axes in each bin. All variables are demeaned by mesh averages.The OLS slope coefficients (solid lines) are .22 (left plot) and .14 (right plot). The sample includes all inventors filing a patent in 1998.

Figure 2: Inventor productivity and team size.



Note: The binned scatterplots group inventors into 20 equal-sized bins according to the variable on the x-axis and compute the mean of the variables on the x- and y-axes in each bin. All variables are demeaned by mesh averages. The OLS slope coefficients (solid lines) are .05 (left plot) and .04 (right plot). The sample includes all inventors filing a patent in 1998.

Figure 3: Innovation quality and inventor characteristics.



Note: The binned scatterplots group patents into 20 equal-sized bins according to the variable on the x-axis and compute the mean of the variables on the x- and y-axes in each bin. The OLS slope coefficients (solid lines) are .08 (left plot) and .13 (right plot). The sample includes all patents filed in 1998.

Fact 4. *The majority of inventor teams are co-located.*

Inventors working on a patent together are typically located in close geographical proximity. Figure 4 plots the average log distance to co-inventors across all Japanese inventors, i.e. the histogram of $\ln \bar{d}_i$. 59 percent of inventors have zero distance to their co-inventors in 1998. The average of $\ln \bar{d}_i$ is 1.01, which implies that the geometric average of distance is $e^{1.01} \approx 2.7$ km.

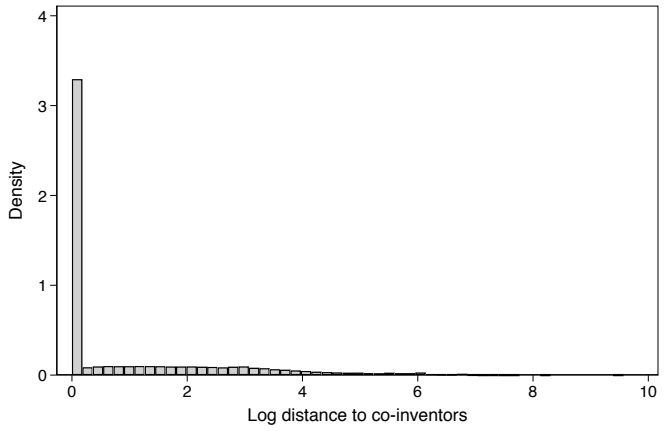
Fact 5. *Productive inventors match to more geographically dispersed teams.*

While many inventor teams are co-located, more productive inventors match to co-inventors farther away. The left plot of Figure 5 shows a binned scatterplot of inventor productivity z_i on the x-axis versus average log distance $\bar{d}_i$ on the y-axis. The right plot uses instead the share of co-inventors within a 10 km of inventor i on the y-axis. We find that higher productivity inventors have greater average distance to their co-inventors and a lower share of co-inventors within a 10 km radius.

3 Background and Methodology

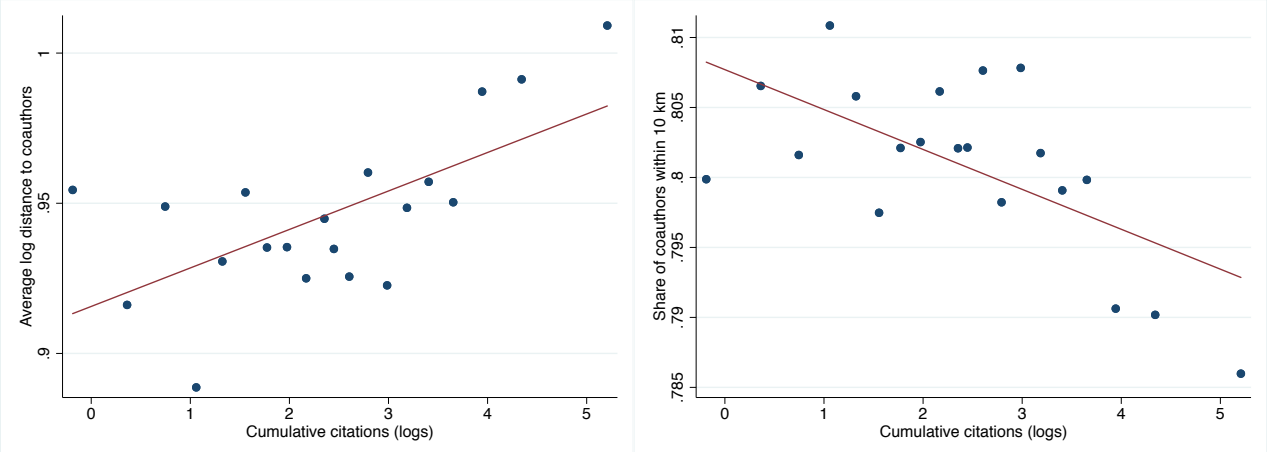
Background. Japan is a collection five main islands, Hokkaido, Honshu, Kyushu, Shikoku and Okinawa, in addition to thousands of smaller islands. Shikoku, the second smallest island with a population of 3.8 million people in 2015, was only accessible by ferry and air until

Figure 4: Histogram of average log distance.



Note: The figure shows the histogram of average log distance to co-inventors across inventors, $\bar{d}_i$. The sample includes all inventors filing a patent with at least one co-inventor in 1998.

Figure 5: Inventor productivity and distance to co-inventors.



Note: The binned scatterplots group inventors into 20 equal-sized bins according to the variable on the x-axis and compute the mean of the variables on the x- and y-axes in each bin. All variables are demeaned by mesh averages. The OLS slope coefficients (solid lines) are .013 (left plot) and -.003 (right plot). The sample includes all inventors filing a patent in 1998.

1988. During the 1980s and 1990s three bridges were completed, which connected Shikoku to Honshu, the largest and most populous island, by expressways and railways.

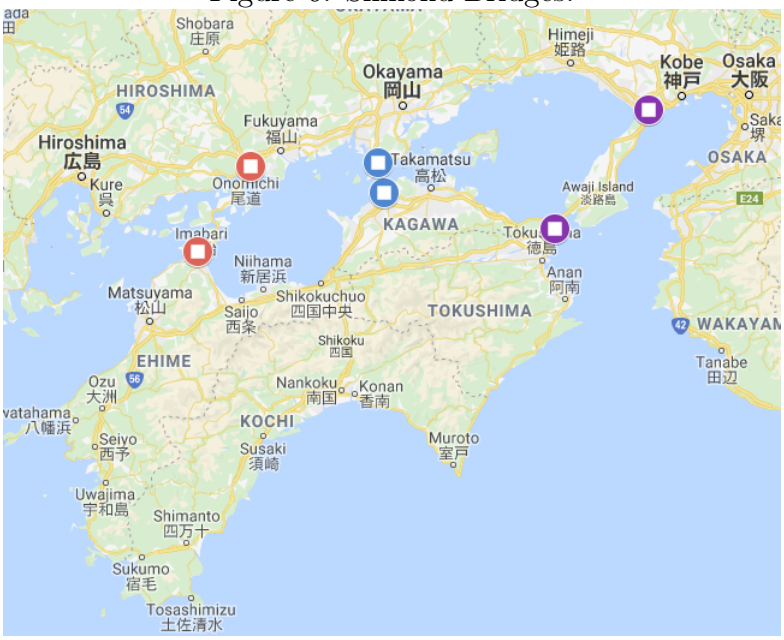
Figure 6 shows a map of the three bridges. The Great Seto bridge was opened in April 1988 (blue points), whereas the Kobe-Awaji-Naruto (KAN) Expressway was opened in April 1998 (purple points) and the Nishiseto Expressway was opened in May 1999 (red points). The bridges connected Shikoku to major economic centers on Honshu, such as Hiroshima in the West and Osaka in the East. All three expressways were major infrastructure projects where planning and construction started many years before their opening. Already in 1889, the member of the Prefectural Parliament Ōkubo Jinnojo, stated in a speech that *"The four provinces of Shikoku are like so many remote islands. If united by roads, they will be much better off, enjoying the benefits of increased transportation and easier communication with each other."*[4] Figure 7 plots daily vehicle volume between Honshu and Shikoku from 1984 to 2007. Daily vehicle volume increased by roughly 160 percent over the period. The share of vehicles passing by ferry decreased from 100 to 12 percent over the period. The largest jumps in vehicle traffic occur in the years just after completion of each of the three bridges.

Innovation on Shikoku increased substantially after the bridges were built. Appendix Figure 10 shows the share of innovation (measured by citation-weighted patents) performed by teams with Shikoku inventors relative to total Japanese innovation. The share almost doubled from the late 1980s until 2004. Appendix Figure 11 shows that the share of innovation (measured by citation-weighted patents) performed by teams with inventors on both Shikoku and not Shikoku also increased over the period, suggesting that the bridges enabled collaboration across islands. For Japan as a whole, however, innovation teams have become less geographically dispersed over time. Appendix Figure 12 shows the share of team members that are within 10/30/60 km of each other (citation-weighted across patents). The share of team members within 10 km has increased by roughly 5 percentage points since the early 1980s. Appendix Figure 13 shows the overall geography of innovation in Japan by plotting citation-weighted patents across all inventors in each 10x10km cell for the period before the 1st bridge (1984-1988) and the period after the final bridge (2000-2004). As expected, innovative activity is concentrated in the Greater Tokyo area. However we also observe innovation in most other regions of Japan, including on Shikoku. One can also eyeball that innovation on Shikoku increased over the time period.

Methodology. We employ a simple differences-in-differences (DID) model to estimate the casual impact of the Shikoku bridges on innovation and matching between inventors. The first difference is the inventors' distance from the bridges and the second difference is the

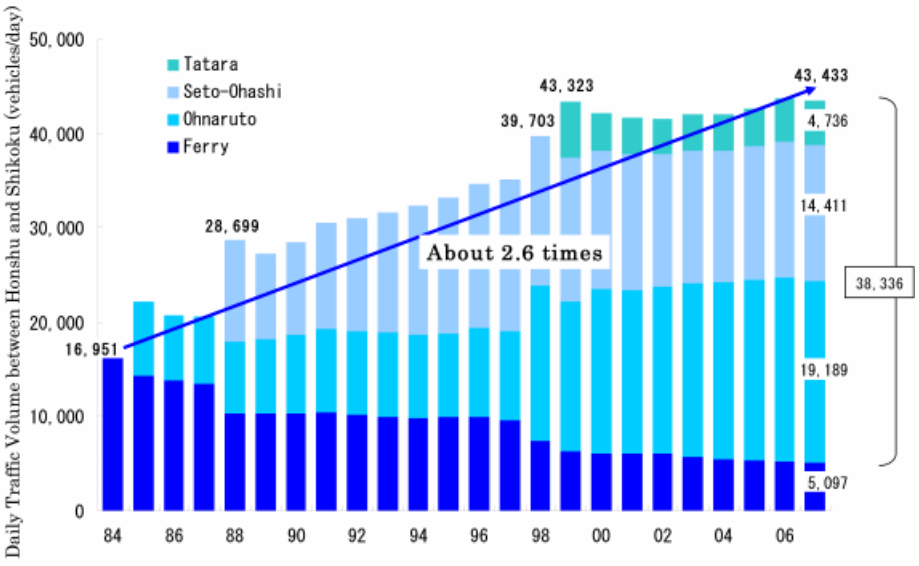
[4] https://en.m.wikipedia.org/wiki/Great_Seto_Bridge. The expressways were built over the periods 1978-1988, 1970-1998 and 1979-1999, respectively.

Figure 6: Shikoku Bridges.



Note: The figure shows the location of the three bridges. The blue/purple/red points indicate the location of the Great Seto Bridge (April 1988), Kobe-Awaji-Naruto Expressway (April 1998) and Nishiseto Expressway (May 1999), respectively.

Figure 7: Honshu-Shikoku Vehicle Traffic Volume, 1984-2007



Note: Source: Business Report of Honshu-Shikoku Bridges.

change in innovation outcomes before and after the opening of the bridges. In the baseline model, we focus on the 1998 and 1999 bridges and the time period 1995-2005. Appendix 5 provides external validity presenting evidence from the 1988 bridge.

Define the continuous treatment variable T_i as inventor i's proximity to the closest bridge, based on her last known location in 1998, i.e.

$$T_i = -\min\left\{\ln Dist_{i,Bridge1998}, \ln Dist_{i,Bridge1999}\right\},$$

where $Dist_{i,BridgeX}$ is the geodesic distance from i to bridge X.[5] Inventors located close to one of the two bridges are then highly treated, whereas inventors far away are mildly treated. In 1998, roughly 30,000 of in total 262,000 inventors were located within 50 kilometers of one of the two bridges. Approximately 3,000 inventors were located on Shikoku itself. We present evidence using indicator variables instead of a continuous treatment variable in the robustness section. The estimating equation is

$$y_{it} = \alpha_i + \delta_t + \beta T_i \times Post_t + \gamma X_{it} + \epsilon_{it}, \tag{4}$$

where y_{it} is an innovation outcome of interest, α_i and δ_t are inventor and year fixed effects and $Post_t = t > 1999$, i.e. it takes the value one the year after the 1999 bridge was completed.[6]

Covariates. The vector X_{it} includes various geographic and inventor characteristics, ensuring that we compare observationally similar inventors that happen to be located close to or far away from the bridges. Specifically, X_{it} includes the following variables interacted with $Post_t$: (i) log distance from inventor i to Tokyo Station, (ii) the quartile of i's pre-bridge productivity (z_{i998}), (iii) the first year i appears as an inventor (i.e., inventor age), as well as (iv) the pre-bridge longitude and latitude of i. In addition, we include a trend term for Shikoku ($t \times Shikoku_i$) that captures Shikoku-specific trends in innovation that are potentially correlated with the completion of the bridges.

Identification. As is standard in DID models, the identifying assumption is that changes in the outcome variable would be similar for treated and non-treated inventors in the absence of the bridges. In the results section below, we test for parallel pre-trends, i.e. whether trends in inventor outcomes are similar among treated and non-treated inventors before the opening of the bridges.

Instrumental variables. As documented above, the planning and construction of the bridges started many years before the opening of the bridges. Therefore, it is relatively

[5] If the inventor did not participate on a patent in 1998, we use geocode of the inventor from the last filed patent before 1998. Distance is measured from the location of i to the nearest side of the bridge (e.g., the Shikoku side of the bridge for Shikoku inventors).

[6] We expect some effect of the 1998 bridge to appear in 1999. Setting $Post_t = t > 1999$ should induce the estimates to be biased down.

unlikely that location-specific trends in innovation in the 1990s triggered the construction of the bridges, i.e. that the bridges are endogenous variables. Nevertheless, we construct a novel instrument that should eliminate any remaining endogeneity concerns. Our instrument is based on the premise that the cost of a bridge is increasing in length of the bridge crossing water. Therefore, the optimal location of the bridge, based on cost considerations only, not taking into account the spatial distribution of economic activity, is the location that minimizes distance over water. This *virtual bridge* is an instrument for the actual bridge because the location of the virtual bridge is by construction invariant to the geography of economic activity, i.e. ϵ_{it}.

We proceed by calculating the optimal location of three virtual bridges, based on their rank in terms of distance over water. Figure 8 shows the locations of the three virtual bridges. Interestingly, both the locations and orders of the virtual bridges are close to the actual ones. The blue circle in the figure is the minimum distance location (virtual bridge no. 1), which is close to the 1988 Great Seto Bridge. The purple circle is the 2nd minimum distance location (virtual bridge no. 2), which is close to the 1998 Kobe-Awaji-Naruto Expressway. The red circle is the 3rd minimum distance location, which is close to the 1999 Nishiseto Expressway (virtual bridge no. 3). The instrument for T_i is then

$$T_i^{IV} = -\min\left\{\ln Dist_{i,VirtualBridge2}, \ln Dist_{i,VirtualBridge3}\right\},$$

where $Dist_{i,VirtualBridgeX}$ is the geodesic distance from the 1998 location of inventor i to virtual bridge X.

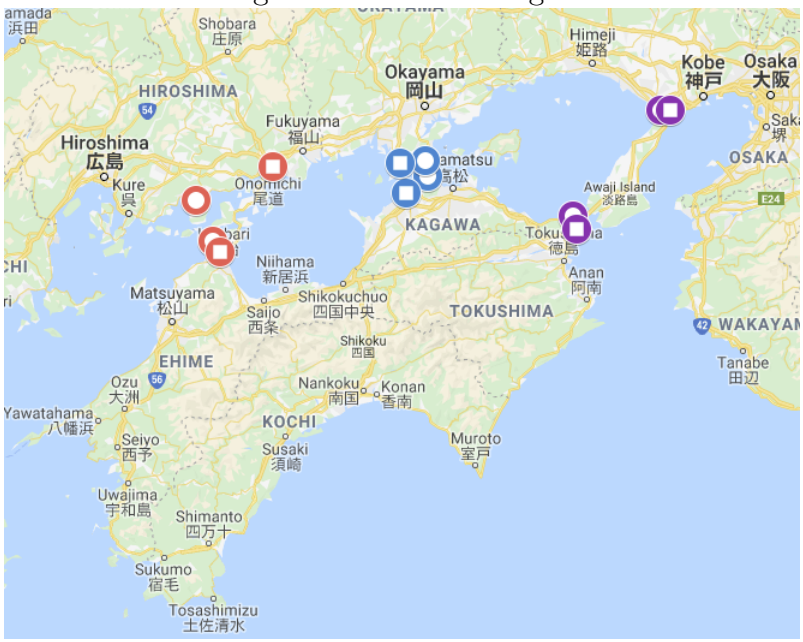
4 Results

4.1 Main Results

We start by presenting evidence on pre-trends inventor productivity z_{it}. Specifically, we estimate a version of equation (4) with separate interaction terms for each year, $T_i \times I\left[Year = t\right]$. Figure 9 plots the coefficient estimates for the interaction terms, with 1998 being the omitted year. We find no evidence on differential pre-trends before the opening of the bridges. After the openings, inventor productivity rises rises markedly for the treated inventors.

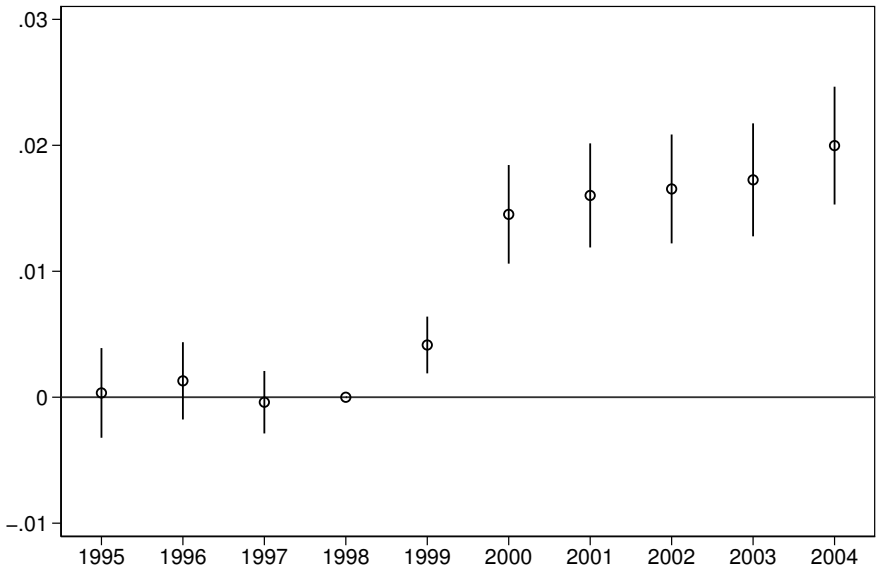
We summarize the results from estimating equation (4) in Table 1. All specifications includes inventor and year fixed effects. The first three columns show results without controls, while the last three columns show results with controls. Column (4) shows the coefficient estimate for log inventor productivity z_{it}. Treated inventors are increasing their productivity relative to non-treated ones. The estimated magnitudes are statistically and economically

Figure 8: Virtual Bridges.



Note: The figure shows the location of the three actual and virtual bridges. The blue/purple/red circles indicate the location of the minimum / 2nd minimum / 3rd minimum distance bridges, respectively. The squares indicate the actual bridges.

Figure 9: Inventor Productivity.



Notes: Coefficient estimates of the interaction terms $T_i \times I[Year = t]$, with 1998 being the omitted year. The outcome variable is $\ln z_{it}$ (citation-weighted). Robust standard errors clustered by inventor are indicated by the vertical lines.

Table 1: OLS Estimates: Inventor Productivity and Matching.

Dep. var.	z_{it}	$\bar{z}_{it}$	t_{it}	z_{it}	$\bar{z}_{it}$	t_{it}
$T_i \times Post_t$	.009^a	.017^a	-.003^a	.015^a	.022^a	-.003^b
	(.001)	(.002)	(.001)	(.002)	(.004)	(.001)
Controls	No	No	No	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Obs	1,779,605	1,479,867	1,993,685	1,779,605	1,479,867	1,993,685

All inventors with a geocode in 1998 are included in the sample. The time period is 1995-2004. The control variables are reported in the main text. Robust standard errors clustered by inventor in parentheses. All dependent variables are in logs. ^{a} p$<$ 0.01, ^{b} p$<$ 0.05, ^{c} p$<$ 0.1.

significant: the bridges increase productivity by 0.11 log points for inventors located next to the bridge compared to inventors with a median distance to the bridge.[7]

The next two columns investigates the impact of the bridges on inventor matching. Column (5) shows the estimate for co-inventor productivity $\bar{z}_{it}$ while column (6) shows the estimate for the average team size of inventor i, t_{it} (both in logs).[8] The bridges lead to the formation of higher quality teams: treated inventors match on average with better co-inventors (based on the citation count of their historical patent portfolio). While innovation teams become better, the size of the teams becomes slightly smaller.

Next, we study the impact of the bridges on the geography of collaboration. Columns (1) and (2) in Table 2 shows the impact on average log distance to co-inventors, $\ln \bar{d}_i$, and the share of co-inventors within a 10 km radius. The bridges lead to teams that are more geographically dispersed, with average distance increasing and the share of nearby inventors decreasing. Using the same calculation as above, the bridge increases average log distance to co-inventors by 0.14 log points for inventors next to the bridge compared to the median inventor. The evidence therefore suggests that the bridges improved search and matching over longer distances. Column (3) uses the share of co-inventors located on Shikoku as the outcome variable. The share declines somewhat, but the negative impact masks two mechanisms. First, inventors on Shikoku might decrease their share of co-inventors located on Shikoku, while inventors on the mainland might increase their share of co-inventors located on Shikoku. We therefore add the interaction variable between being located on Shikoku in 1998 and $Post_t$, $Shikoku_i \times Post_t$, in column (4). The main interaction variable now

[7] Max T_i is 1.38 (.25 km) while the median is -6.04 (420 km), which yields $.015\,(1.38 + 6.03) = 0.11$.

[8] The median of t_i is 2.7 and the share of inventors with no co-inventors is 21 percent in 1998. Inventors with no co-inventors drop out when the dependent variable is $\bar{z}_{it}$.

Table 2: OLS Estimates: Geography.

	$\ln \bar{d}_i$	Share $<$ 10 km	Share co-inventors on Shikoku	
$T_i \times Post_t$	1.927^a	-.121	-.032^c	-.002
	(.502)	(.092)	(.017)	(.017)
$Shikoku_i \times Post_t$				-7.301^a
				(.842)
Controls	Yes	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Obs	1,580,215	1,584,754	1,584,679	1,584,679

All inventors with a geocode in 1998 are included in the sample. The time period is 1995-2004. The control variables are reported in the main text. Robust standard errors clustered by inventor in parentheses. Estimates multiplied by 100. ^{a} p$<$ 0.01, ^{b} p$<$ 0.05, ^{c} p$<$ 0.1.

becomes insignificant, while the new interaction variable is negative and significant, showing that Shikoku-based inventors form teams with fewer co-inventors located on the island. This is therefore direct evidence that the bridges led to collaboration across the islands.

Finally, we turn to the instrumental variable approach. As documented in Figure 8, the location of the virtual bridges are relatively close to the location of the actual bridges, suggesting that their locations were determined by costs rather than economic geography. It is no surprise, then, that the first stage estimates shown in Table 3 are very precisely estimated. Focusing on the same four outcomes of inventor productivity and matching as in Table 1, we find virtually identical point estimates compared to the OLS results.

4.2 Mechanisms and Robustness

Heterogeneous treatment effects. Our estimates indicate that average own- and match productivities are increasing in response to the opening of the bridges, but this may mask heterogeneous effects across inventors. One hypothesis is that less productive inventors have a harder time matching with highly productive inventors because the bridges introduces more competition among inventors. We investigate this by re-estimating the main regression in equation (4), but now include the triple interaction term $T_i \times Post_t \times HighZ_i$ where $HighZ_i$ takes the value one if inventor i has a higher than median value of z_{it} before 1998. Table 4 shows the results. We find that highly productive inventors do not innovate more (column 1), but there is evidence that the bridges enable them to match with better inventors compared to less productive inventors (column 3).

Table 3: 2SLS Estimates: Inventor Productivity and Matching.

	z_{it}	$\bar{z}_{it}$	t_{it}
$T_i \times Post_t$	.016^a	.021^a	-.003^a
	(.001)	(.003)	(.001)
		First stage	
$T_i^{IV} \times Post_t$	1.026^a	1.026^a	1.026^a
	(.002)	(.002)	(.002)
Controls	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs	1,779,605	1,479,867	1,993,685

All inventors with a geocode in 1998 are included in the sample. The time period is 1995-2004. The control variables are reported in the main text. Robust standard errors clustered by inventor in parentheses. All dependent variables are measured in logs. ^{a} p$<$ 0.01, ^{b} p$<$ 0.05, ^{c} p$<$ 0.1.

Table 4: OLS Estimates: Heterogeneous Treatment Effects.

	z_{it}	$\bar{z}_{it}$	t_{it}
$T_i \times Post_t$	.013^a	.015^a	-.005^a
	(.003)	(.005)	(.002)
$T_i \times Post_t \times HighZ_i$	.003	.012^a	.004^b
	(.003)	(.005)	(.002)
Controls	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs	1,779,605	1,479,867	1,993,685

All inventors with a geocode in 1998 are included in the sample. The time period is 1995-2004. The control variables are reported in the main text. All dependent variables are measured in logs. Robust standard errors clustered by inventor in parentheses. ^{a} p$<$ 0.01, ^{b} p$<$ 0.05, ^{c} p$<$ 0.1.

Table 5: OLS Estimates: Inventor exit.

	(1)	(2)
T_i	-.002[a]	-.004[a]
	(.001)	(.001)
Controls	No	Yes
Inventor FE	No	No
Year FE	No	No
Obs	614,697	614,697

All inventors with a geocode in 1998 are included in the sample. The control variables are reported in the main text. Robust standard errors in parentheses. [a] p< 0.01, [b] p< 0.05, [c] p< 0.1.

Treatment indicators. One might expect that the treatment intensity is not linear in log distance from the bridges. We therefore group inventors into three bins: the ones that are within 140 km of the bridges, T_i^{140km}(the 1st quartile), the ones that are between 140 and 440 km, T_i^{440km} (the 3rd quartile) and everyone else. Table 6 shows the results for inventor productivity and matching. As expected, the gains are concentrated among inventors in the 1st quartile: the point estimates are markedly higher for the inventors close to the bridge compared to the ones in the 2nd and higher quartiles.

Inventor exit. In the baseline analysis, inventors that stop inventing will disappear from the sample. We therefore investigate whether the bridges also affect the likelihood of inventor exit. Inventor exit, $Exit_i$, is defined as not observing inventor i on any patent after 1999, i.e. in the $Post_t$ period. Table 5 shows results when regressing $Exit_i$ on the continuous treatment variable T_i. Note that this is a purely cross-sectional regression, so we can no longer include inventor and year fixed effects. Column (1) has no controls, whereas column (2) has the same set of controls as in the baseline regressions. In both cases, we estimate a negative effect on exit, i.e. the bridges lowered the likelihood of inventor exit for individuals close relative to far away from the bridges.

The Great Seto Bridge. Finally, we investigate the impact of the first (1988) bridge. The sample time period is now 1984 to 1993, and $Post_t$ takes the value one after 1988. All inventors with a geolocation in 1988 or earlier are included in the sample. The treatment variable is now (minus) the log distance from inventor i's 1988 location to the 1988 Great Seto Bridge, $T_i = -\ln Dist_{i,Bridge1988}$. We find significant effects on inventor productivity z_{it}, with a point estimate that is greater than the baseline results of Table 1, underscoring the robustness and external validity of the results. The effect on co-inventor productivity $\bar{z}_{it}$

Table 6: OLS Estimates: Inventor Productivity and Matching.

Dep. var.	z_{it}	$\bar{z}_{it}$	t_{it}
$T_i^{140km} \times Post_t$	$.052^a$	$.033^a$	-.002
	(.005)	(.010)	(.004)
$T_i^{440km} \times Post_t$	$.027^a$	$.018^a$	.004
	(.004)	(.007)	(.003)
Controls	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs	1,779,605	1,479,867	1,993,685

All inventors with a geocode in 1998 are included in the sample. The time period is 1995-2004. The control variables are reported in the main text. Robust standard errors clustered by inventor in parentheses. All dependent variables are measured in logs. ^{a} p$<$ 0.01, ^{b} p$<$ 0.05, ^{c} p$<$ 0.1.

is, however, no longer significant.

5 Conclusions

This paper examines the effect of a large reduction in travel time on the production of knowledge. The connection of Shikoku to Honshu by three bridges during the late 1980s and -90s led to a large reduction in travel time compared to existing modes of transport (ferry and air). We find a significant increase in the number and quality of patents produced by inventors near the bridges on Shikoku. In addition, these inventors are more likely to partner with teams on Honshu as well as with more productive teams. Our results suggest that economic integration can have sizable effects on idea flows and consequently innovation and growth.

Table 7: OLS Estimates: The Great Seto Bridge.

	z_{it}	$\bar{z}_{it}$	t_{it}
$T_i \times Post_t$	.022^a	.001	-.009^a
	(.005)	(.007)	(.002)
Controls	Yes	Yes	Yes
Inventor FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Obs	1,046,681	945,436	1,390,313

All inventors with a geocode in 1988 are included in the sample. The time period is 1984-1993. The control variables are reported in the main text. Robust standard errors clustered by inventor in parentheses. ^{a} p$<$ 0.01, ^{b} p$<$ 0.05, ^{c} p$<$ 0.1.

References

Agrawal, Ajay and Avi Goldfarb, "Restructuring Research: Communications Costs and the Democratization of University Innovation," *American Economic Review*, 2007, *322* (4), 1578–1590. 1

Akcigit, Ufuk, Santiago Caicedo, Ernest Miguelez, Stafanie Stantcheva, and Valerio Sterzi, "Dancing with the Stars: Innovation through Interactions," Technical Report, NBER Working Paper 24466 2018. 1

Catalini, Christian, Christian Fons-Rosen, and Patrick Gaule, "How Do Travel Costs Shape Collaboration?," Technical Report, NBER Working Paper No. 24780 2018. 1

Comin, Diego A., Mikhail Dmitriev, and Esteban Rossi-Hansberg, "The Spatial Diffusion of Technology The Spatial Diffusion of Technology The Spatial Diffusion of Technology," Technical Report, NBER Working Paper No. 18534 2012. 1

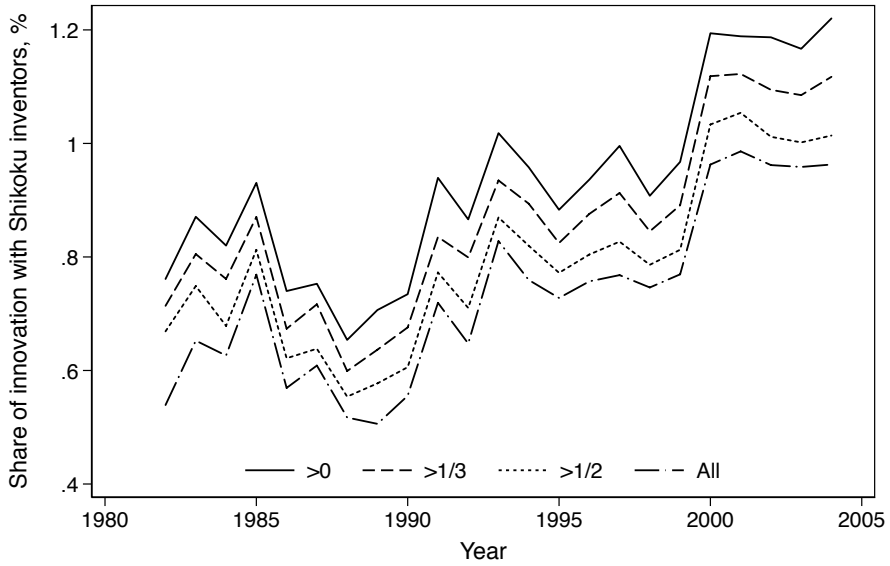
Head, Keith, Yao Amber Li, and Asier Minondo, "Geography, Ties, and Knowledge Flows: Evidence from Citations in Mathematics," *The Review of Economics and Statistics*, 2019, *101* (4). 1

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson, "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations," *The Quarterly Journal of Economics*, 1993, *108* (3). 1

Jones, Benjamin F., Stefan Wuchty, and Brian Uzzi, "Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science," *Science*, 2008, *98*, 1259–1262. 1

Wuchty, Stefan, Benjamin F. Jones, and Brian Uzzi, "The Increasing Dominance of Teams in Production of Knowledge," *Science*, 2007, *316*, 1036–1039. 1

Figure 10: Share of Innovation on Shikoku.



Notes: The figure shows the share (in %) of citation-weighted patents with at least one inventor (>0) / more than 1/3 of inventors / more than half of inventors / all inventors located on Shikoku in the application year. The population is all patents with at least one domestic inventor.

Appendix

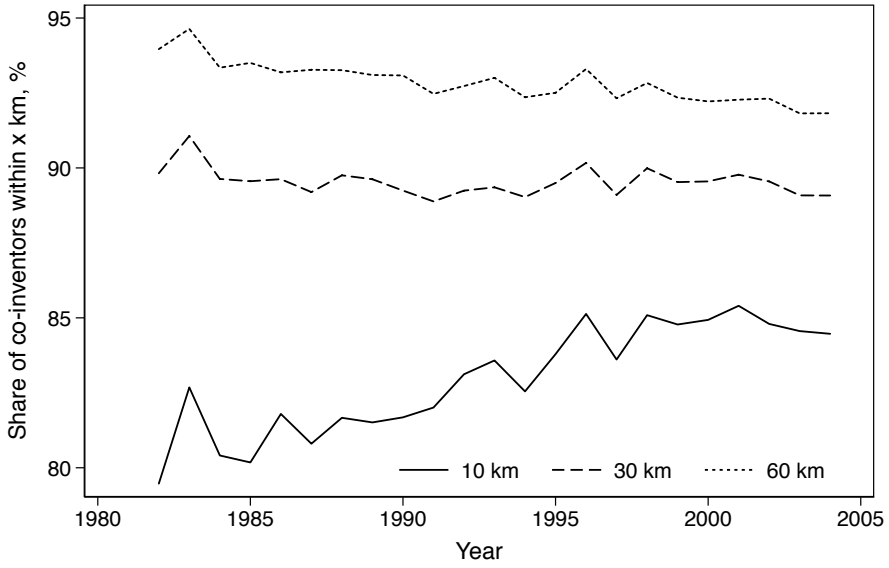
A Additional Tables and Figures

Figure 11: Share of Innovation with both Shikoku and non-Shikoku inventors.



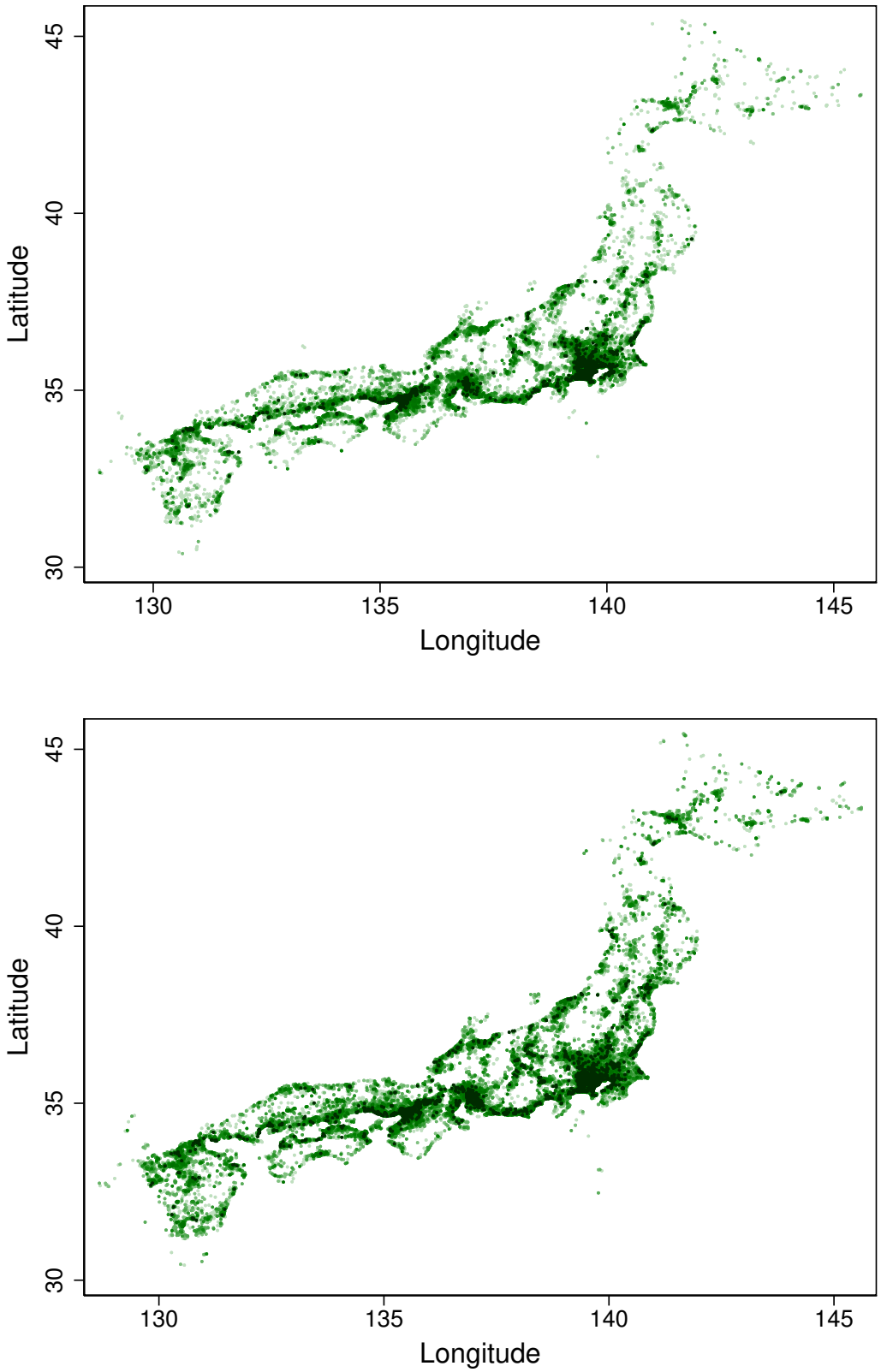
Notes: The figure shows the share (in %) of citation-weighted patents with at least one inventor located on Shikoku and one inventor not located on Shikoku. The population is all patents with at least two domestic inventors.

Figure 12: Share of Co-inventors Within 10/30/60 km.



Notes: The figure shows the weighted average share of co-inventors within 10/30/60 km of each other. Citations are used as weights. The population is all patents with at least two domestic inventors.

Figure 13: The Geography of Innovation.



Notes: The figures show innovation in each 10x10km cell during the period 1984-1988 (upper panel) and 2000-2002 (lower panel). Innovation is measured by the citation-weighted number of patents across all inventors in a given cell. Darker shades represent more innovation.