High-Speed Railways and the Geography of Innovation: Evidence from France

Gabrielle Gambuli Fernando Stipanicic*

February 28, 2025

Preliminary and incomplete

Abstract

This paper provides new causal evidence of the impact of train travel time on patent collaboration between inventors. We construct a novel dataset of train travel times in France between 1980 and 2010 and exploit the roll out of High Speed Railways (HSR) as a quasi-natural experiment. The median decrease in travel time of 25% led to a 5% increase in patent collaborations across departments. The effect on the increase in collaboration was stronger for department pairs in which both departments were more developed, likely increasing the collaboration gap with less developed departments.
JEL Classification: O31, O33, R12, R41, N72, F60

*Gambuli: Université Gustave Eiffel & INSEE (gabrielle.gambuli@gmail.com)
Stipanicic: University of Oslo (fernando.stipanicic@gmail.com)

1 Introduction

Innovation and technological progress are key drivers of economic growth, yet their distribution across space remains highly uneven. Since at least the publication of Vannevar Bush's report, "Science, the Endless Frontier," in 1945, countries have promoted policies to increase innovation. Recent policies and recommendations, such as the CHIPS Act in the United States (U.S. Congress (2022)) and the Draghi Report in the European Union (Draghi (2024)), make a strong emphasis on fostering regional innovation with the goal of reducing regional disparities. A growing body of research highlights the importance of face-to-face interactions in driving innovation thorugh knowledge spillovers (Catalini (2018), Atkin, Chen and Popov (2022)), and that this knowledge spillovers are strongly affected by distance (Jaffe, Trajtenberg and Henderson (1993), Belenzon and Schankerman (2013)). Recent studies show improvements in highway infrastructure (Agrawal, Galasso and Oettl (2017)) and reductions in air travel time (Pauly and Stipanicic (2022)) have facilitated knowledge spillovers, potentially affecting regional innovation. However, little is known about the role of high speed railways (HSR) in shaping the geography of innovation, particularly in Europe.

This paper studies the impact of reductions in travel time due to the expansion of high-speed railways (HSR) in France on patent collaborations, called copatents. The introduction of HSR led to a substantial decrease in travel time, facilitating long-distance face-to-face interactions among inventors. We construct a new dataset of train travel times from 1980 to 2020 and match it geo-referenced patent data from the European Patent Office (1980–2010) aggregated to the department level (NUTS3 region). We estimate a gravity model and find that the median decrease in travel time of 25% led to 5% increase in patent collaborations across departments. However, we find that the increase is stronger for department pairs in which both departments are more developed, likely increasing the innovation gap with departments less developed. Additionally, we find that the increased patent collaborations is driven by new inventor-pair collaborations and by collaborations across firms, rather than within.

Our identification strategy exploits the staggered rollout of HSR, which introduced variation in travel times across department pairs. The identification assumption is that the timing of adoption of HSR is exogenous, which is reasonable given the uncertainty around completion times. We include department-pair, origin-time and destination-time fixed effects in the regression, which absorb time invariant characteristics at the department-pair level, and time varying shocks at the origin and destination departments. One potential endogeneity threat would be the existence of time-varying shocks at the department-pair level that affect collaboration patents and are systematically correlated with the timing of HSR rollout. We addres this concern by re-estimating the baseline model in a subsample of department pairs that do not have HSR station, and hence benefit from the roll out of HSR only indirectly through connecting stations. We find results that go in the same direction as in the full sample.

The roll out of HSR in France started with the inaguration of the first line in 1981, connecting Paris and Lyon, which are located approximately 400 kilometers apart. This development reduced train travel time from 3 hours and 40 minutes to 1 hour and 40 minutes, making it viable for individuals to complete a return trip within a single day. Over the years multiple other cities became connected to Paris: Le Mans

(1989), Tours (1990), Lille (1993), Marseille (2001), Montpellier (2001), Strasbourg (2007-2016), Bordeaux (2017), and Rennes (2017). By the end of 2017, the high-speed rail network covered more than 1,500 kilometers.

We construct a new dataset of train travel time in France. To do so we develop a new method to estimate the unobserved counterfactual travel time before the arrival of HSR. Our method is general and can be applied to reconstruct unobserved travel time in other set ups when the econometrician counts with partial information on the transportation technology. We document in a sample of city-pairs that our method replicates between 58% and 87% of the observed changes in travel time. With our estimated travel times, we document that between 1980 and 2010 the average reduction in train travel time was 14%, while the median and the 75th quantile were 25% and 35%, respectively.

We assemble a dataset of collaborative patents in France granted by the European Patent Office. We take patents with two or more inventors with high-quality geolocation based on Morrison, Riccaboni and Pammolli (2017) and aggregate it to the department-pair year level. Between 1980 and 2010, the share of collaborative patents increased from 36% to 62%, with the share of inter-regional collaborations growing at twice the rate of intra-regional collaborations. By 2010, the share of inter-regional patents accounted for around half of all copatents in France.

We estimate a regression of copatents on travel time. The empirical strategy exploits only changes in travel time and copatents that are differential across department pairs, absorbing any aggregate changes in travel time or copatents. Following Silva and Tenreyro (2006) we estimate the regression by Poission Pseudo Maximum Likelihood (PPML) which allows for zeros in the left hand size variable, and gives an unbiased estimate in the case of heteroskedasticity of the underlying multiplicative model. We estimate that the elasticity of copatents to travel time is -0.2, with the effect mostly coming for department-pairs at 100km-400km distance. This result is plausible, as inventors may choose to travel by car for short distances, and by airplane for longer distances.[1] We split departments by whether they are mostly urban or rural according to Eurostat, which we label Core and Periphery, and find that the elasticity is only negative and significant for Core-Core collaborations. We find a similar result based on under/over median population density. As consequence, HSR leads to increased collaborations between department that are likely to be already the more developed ones.

Additionally, while we find similar point estimates for the elasticity for copatents of inventor-pairs that previously collaborated and new inventor-pairs, it is only statistically different from zero for new inventor-pairs. We also find that the elasticity is larger in absolute value and statistically significant for across-firm collaborations, while it is not statistically different from zero for within-firm collaborations. We also find that the effect remains when weighting copatents by amount of claims and technology classes, suggesting that the increased copatenting does not have reduced scope nor breadth. We find that the elasticity becomes positive when weighting copatents by their citations received in a 5-year window, though we are cautious interpreting this result as most patents receive zero citations and hence the effective sample may be substantially different. Moreover, we find that the increase in collaborative

[1] In 2010, the average train travel time for distances over 400km in the effective sample was 5 hours and 10 minutes.

patents happens between inventors that have both similar and different knowledge bases, as captured by the technology classes of their previous patents.

Literature. This paper contributes to the literature on transportation infrastructure and innovation by providing new evidence on the effect of HSR on patent collaborations. While previous research has shown that highways (Agrawal, Galasso and Oettl (2017)) and air travel (Pauly and Stipanicic (2022); Bahar et al. (2023)) facilitate knowledge spillovers, less is known about the role of HSR.[2] A distinct characteristic of HSR is that it is a means of transportation for people rather than goods, making face to face interactions the relevant mechanism at play. Our results show that reductions in train travel time increase inter-regional patenting, particularly at intermediate distances where neither car nor air travel is the dominant mode of mobility.

Second, we contribute to the literature on spatial inequality and regional innovation (Moretti (2012); Duranton et al. (2009); Carlino and Kerr (2015); Gross and Sampat (2023)). While recent policies emphasize fostering innovation in less developed areas (U.S. Congress (2022); Draghi (2024)), our findings suggest that HSR primarily strengthens collaborations between already developed regions. By showing that increased collaboration is driven by new inventor-pair collaborations and across-firm interactions, we also shed light on how reduced travel time shapes the organization of innovation.

Finally, we make a methodological contribution by developing a new approach to estimate historical travel times in the absence of comprehensive records. Our method replicates a large share of the variance in observed changes in travel time and can be applied to other settings where transportation infrastructure has evolved. This provides a useful tool for studying the long-term effects of transportation improvements on economic outcomes.

2 Train travel time data

We construct a new train travel time dataset between departments of continental France at the yearly frequency between 1980 and 2020. This period includes the roll out of high speed railways, which are around twice as fast as the previous train technology. We document that during this period the median and 3rd quartile reduction in travel time are 20% and 32%, respectively. This reduction in travel time affected only passenger transport, as high speed railways are not used to transport goods. While the plans for construction of high speed railways were publicly advertised, the date of opening of new routes was uncertain, introducing randomness in the timing of travel time reductions. Appendix B provides more details on each of the following subsections.

The method used for constructing this dataset is general and can be applied for constructing other data sets in which previous values of a variable are not observed. This method is especially useful when the analyst knows when the new technology was implemented and either has knowledge of the efficiency of the previous technology (e.g. what was the speed of trains in those routes before, combined

[2] Dong, Zheng and Kahn (2020) studies the impact of HSR in China on scientific collaboration and Tsiachtsiras et al. (2022) do so on local innovation and technological specialization. To the best of our knowledge, no study has examined the impact of HSR on innovation in a European setting

with route distance), or the efficiency can be estimated from currently observed values (e.g. the older train types are still used in other routes). Hence, our method is a more general contribution on how to construct datasets where past values are not observed, and the current application operates as a proof of concept for such method.

2.1 Roll out of High Speed Railways in France

High speed railways in France have a maximum speed around 320 km/h, compared to around 160 km/h for the alternative Intercités train. Starting in 1981 with the connection between Lyon and Saint Florentin, high speed railways have been rolled out in France, counting more than 1,500km of rail by the end of 2017.

Figure 1 shows the roll out of high speed railways by decade. The new high speed railways connected Paris with largest cities in a gradual manner. The first segment to Lyon opened in 1981, then Le Mans (1989), Tours (1990), Lille (1993), Marseille (2001), Montpellier (2001), Strasbourg (2007-2016), Bordeaux (2017), and Rennes (2017). Given the staggered roll out and the network nature of train travel, one city pair could have reductions in travel time multiple times. For example, Paris-Marseille travel time went from 6 hours in 1980, to 4 hours 36 minutes in 1983, to 3 hours in 2001.

The date of opening of high-speed railways was uncertain due to financial constraints, political negotiations, judicial rulings, and environmental concerns.[3] For example, the Paris-Lille line was originally tied to the construction of the Channel Tunnel, but the British government withdrew support in 1975 for financial reasons, leading France to shift its focus to the Paris-Lyon line instead. Political agreements later revived the Paris-Lille project, and despite a relatively quick construction period, public opposition in Lille led to security forces being deployed to control protests. Similarly, the Bordeaux-Toulouse line did not begin construction until three decades after its initial announcement due to legal battles, environmental opposition, and financial disputes, with a court ruling initially halting the project before a later decision allowed it to proceed. Even projects included in national railway plans, such as the Paris-Le Havre line, were ultimately abandoned due to concerns over financial viability, further highlighting the unpredictability of HSR expansions. This uncertainty in completion dates, driven by factors unrelated to regional innovation, introduces arguably exogenous variation in the timing of travel time reductions, which we exploit for identification.

2.2 Data construction: train travel times

We construct a data set on train travel times between all continental French departments (NUTS3 region) at the yearly frequency between 1980 and 2020. We use as input the December 2021 travel times from the French railway company SCNF and manually gather the opening dates of each high speed railway segment.[4] Based on the 2021 train network, we estimate the travel speed of each train-railway type.

[3] Details are provided in Appendix A.

[4] High speed trains (TGV – Trains à Grande Vitesse) operate in both high speed railways and normal railways. Other train types do not operate on high speed railways.

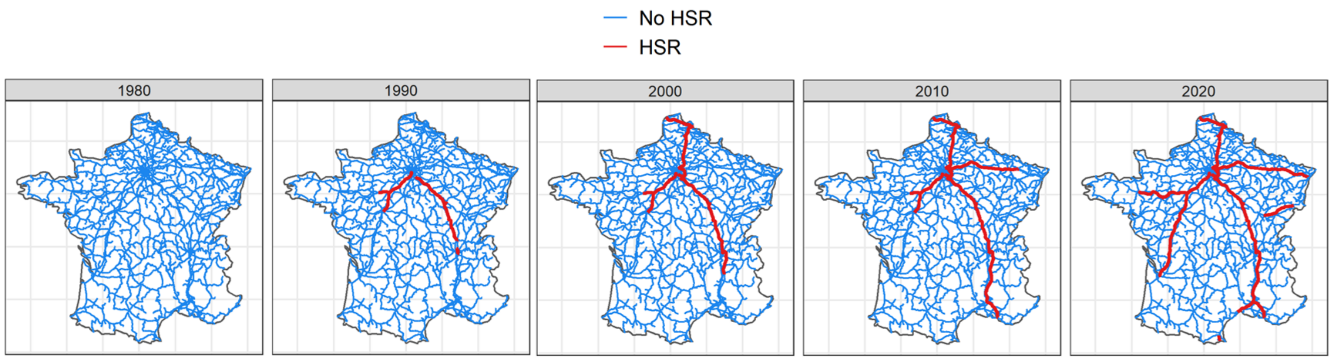

Figure 1: Roll of high speed railways in France

With these estimates we simulate the travel time that would have taken place in each of the segments before the opening of the high speed railway. We validate the travel time against observed travel times for a sample of city-pair years provided by SNCF and document that the constructed dataset accounts for between 58% and 87% of the observed changes in travel time.[5]

We start with all trips by train and railway type during a week of December 2021, adding up to 47,298 trips in 416 routes (origin-destination station pairs). We split each trip, which may include multiple stops, into segments of non-stop station pairs. For each segment and train-railway type, we take the minimum travel time across all trips. We then estimate a linear regression of the minimum travel time on distance for each train-railway type. Appendix Table 3 presents the results for the estimations. For the relevant train types, high speed railways and Intercités, these simple linear models explain around 95% of the variance showing the predicted values would be a good approximation for observed values. The estimated travel speed for high speed trains operating in high speed railways is 229 km/h, while it is 112 km/h for the second fastest train type Intercités.

With the estimated speeds, we impute the counterfactual Intercités travel time for each non-stop station-pair operated with high speed railway in December 2021. We use the counteractual travel time for each non-stop station-pair in all years previous to the year in which the high speed railway opened in that segment, switching then to the observed travel time with high speed railway. For all other station-pairs that are not connected with high speed railways we use the observed travel times from the December 2021 dataset. Hence, across-time variation in travel time comes only from switching from the counterfactual Intercités travel time to the observed travel time with high speed railways, everything else remains constant.

Next, using the travel time between non-stop station pairs for each year, we run the Dijkstra algorithm (Dijkstra (1959)) to obtain the fastest route and travel time between all station pairs in each year.[6]

[5]Appendix Table 4 regresses observed travel time on our predicted travel time for the sample of city-pair years. The regression gives a R2 of 87% when including origin-destination fixed effect. The R2 is 58% when including origin-destination, origin-time and destination-time fixed effects. Hence, our predicted travel times capture a large share of the observed across-destination within-origin variation of changes in travel time, which is the variation that we use for identification in the baselines specification.

[6]We set a zero-minute penalty for switching trains within a station. We allow for changes of station within a city using the

Then, we keep only origin-destination stations that belong to the most populated municipality in each department.[7] Finally, we take the minimum travel time across station-pairs within each department-pair for each year.

2.3 Descriptive statistics: Train travel times

For each department pair, we compute the change in travel time relative to 1980. Figure 2 shows, for each decade, the change in travel time by department-pair relative to 1980. Figure 2 shows the change in travel time within department-pair relative to 1980, averaged within 100km distance bins. The change in travel time is non-uniform, with larger reductions in travel time for department-pairs that are farther apart. The average reduction in travel time is 14% in 2010 and 19% in 2020. The median, 75th percentile and 90th percentile reduction in travel time are, respectively, 11%, 25% and 35% in 2010, and 20%, 32% and 39% in 2020.

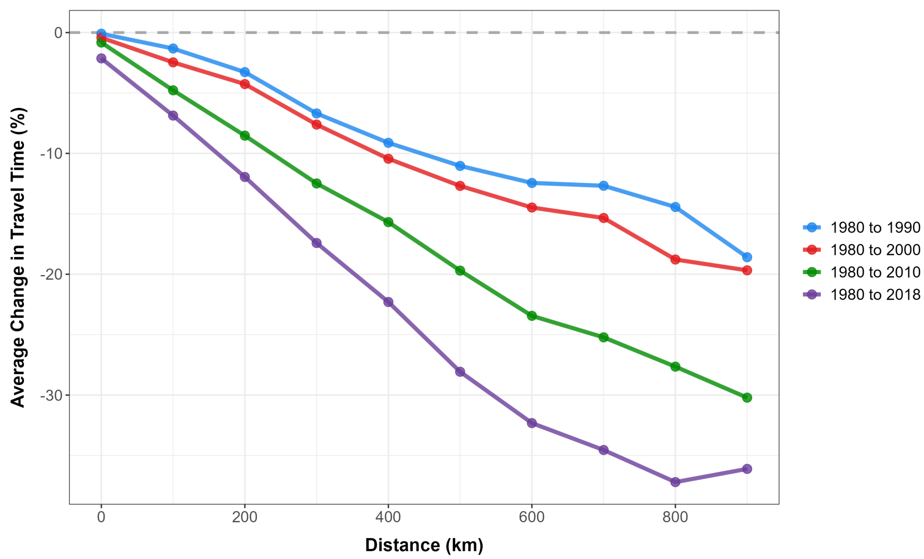

Figure 2: Change in travel time relative to 1980

3 Patent data

We assemble a dataset of patents granted by the European Patent Office (EPO) to inventors residing in France with application year 1980-2010. We obtain geo-referenced patent information from Morrison, Riccaboni and Pammolli (2017), with time invariant identifiers both for inventors and assignees. This dataset covers 1978-2014. We restrict the sample to patents applied between 1980-2010 because, first, we observe a spike in patent applications in 1978, year of creation of the EPO. Second, there is a lag between patent application and patent granting, which introduces measurement error towards the latest years of

distance across stations and assuming the change happens at 45 km/h.

[7] We use average municipality population in 1980-2020.

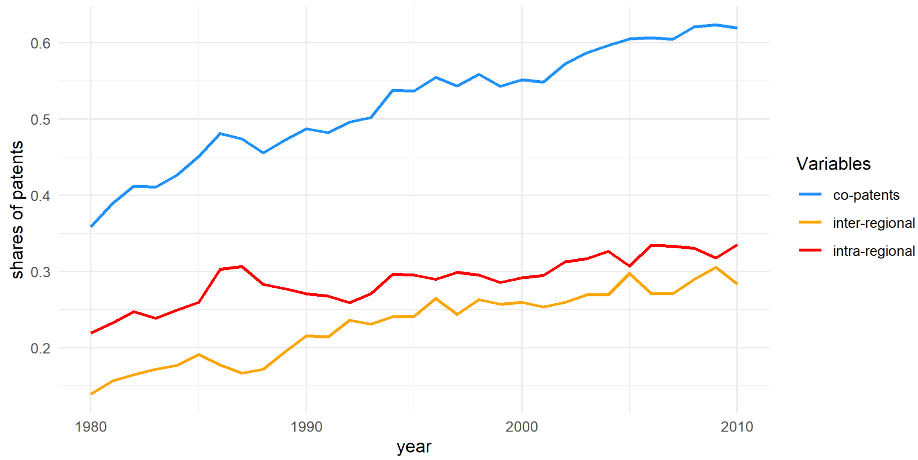

Figure 3: Share of copatents in all copatents

the sample. We keep only multi-inventor patents that have at least 2 inventors with high-quality georeferencing in Morrison, Riccaboni and Pammolli (2017). Our selected sample includes 63,041 patents done by 77,345 inventors located in 94 departments.[8] Then, for each department-pair year, we count the amount of patents that have inventors in both departments.

3.1 Descriptive statistics: Patent data

Figure 3 shows the share of copatents in the total number of patents, and splits the share by copatents that have all inventors in the same department (intra-regional) and copatents that have inventors in multiple departments (inter-regional). The share of copatents increased from 36% in 1980 to 62% in 2010, with increase in both the share of intra-regional and inter-regional copatents. Nonetheless, the growth rate of inter-regional copatents was twice the one of intra-regional copatents. The increase in the share of copatents is accompanied by an increase in the distance across inventors within a team. The average team distance increased by 50%, which was in part due to the shift towards more inter-regional copatents. Comparing team distance of inter-regional copatents, the average team distance increased by 27%.

4 Analysis

4.1 Time varying effect of geography

We begin the analysis by providing evidence that the effect of geography on copatents has been decreasing over time. We do so by estimating in gravity equation with, first, the time varying effect of distance, and second, the time varying effect of within-department collaboration.

[8]Given commuting patterns, we aggregate Paris and its surroundings departments Hauts-de-Seine,Seine-Saint-Denis and Val-de-Marne into one department. Hence, the final dataset has 91 departments, with one of them being Paris and its surroundings.

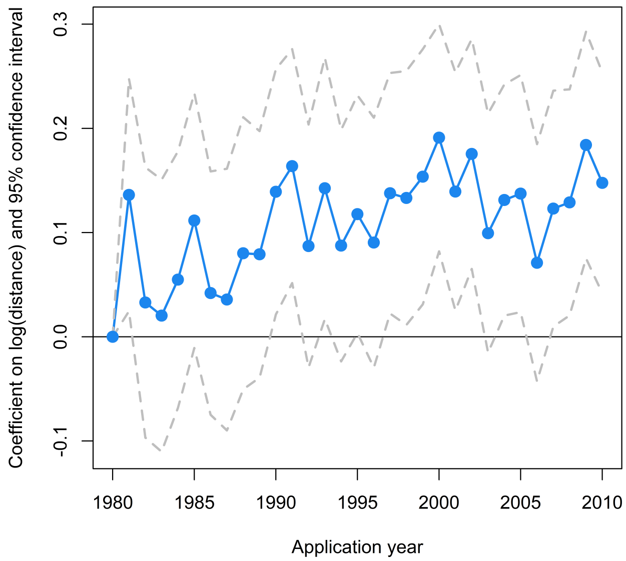

Figure 4: Time varying effect of distance on copatents
The plot shows the point estimate and 95% confidence interval of α_t obtained estimating $copatents_{ijt} = exp(\sum_t \alpha_t \log(\text{distance}_{ij}) + \mu_{ij} + \mu_{it} + \mu_{jt}) \times \varepsilon_{ijt}$, normalized to 1980. Confidence intervals are obtained with ij clustered standard errors.

Figure 4 shows that, relative to 1980, the negative effect of distance on copatents has "softened" – became less negative– over time.

Figure 3 shows that around two thirds of copatents had all inventors located within the same department in 1980, reducing to around one half by 2010. Figure 5 shows that, after controlling for pair, origin-time and destination-time fixed effects, the effect of within-department collaboration on copatents has decreased relative to 1980.

4.2 Effect of travel time

To study the effect of travel time on copatents we estimate the following gravity equation:

$$\text{copatents}_{ijt} = \exp\left[\beta \log(\text{travel time}_{ijt}) + \mu_{ij} + \mu_{it} + \mu_{jt}\right] \times \varepsilon_{ijt} \tag{1}$$

for origin department i, destination department j and year t. copatents_{ijt} is the number of copatents and $\log(\text{travel time}_{ijt})$ is the log of train travel time.[9,10] β is the elasticity of copatents to travel time.

[9] While copatents are in principle non-directional, we treat them as directional as this allows to estimate the gravity equation in a similar manner as in international trade models. Hence, both copatents_{ijt} and copatents_{jit} appear in the data and $\text{copatents}_{ijt} = \text{copatents}_{jit}$.

[10] We assume passengers take a round trip, hence we make travel time symmetric, i.e. $\text{travel time}_{ijt} = (\widetilde{\text{travel time}}_{ijt} + \widetilde{\text{travel time}}_{jit})/2$, where $\widetilde{\text{travel time}}_{ijt}$ is our constructed travel time.

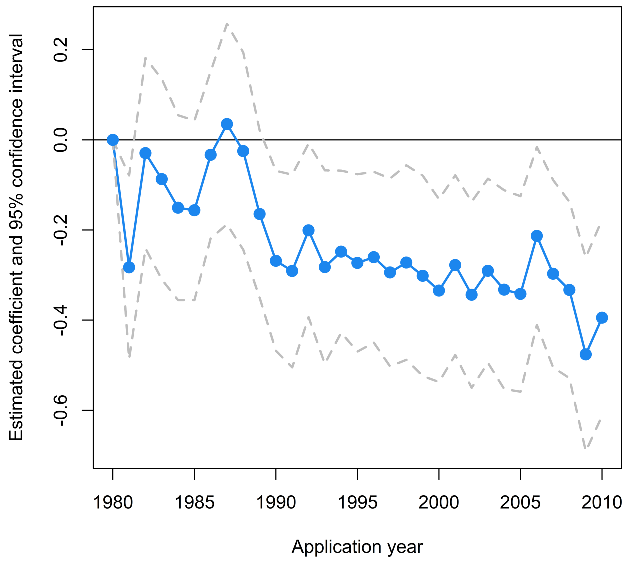

Figure 5: Time varying effect of within-department on copatents

The plot shows the point estimate and 95% confidence interval of α_t obtained estimating $copatents_{ijt} = exp(\sum_t \alpha_t \mathbb{1}\{i = j\} + \mu_{ij} + \mu_{it} + \mu_{jt}) \times \varepsilon_{ijt}$, normalized to 1980. Confidence intervals are obtained with ij clustered standard errors.

Following Silva and Tenreyro (2006) we estimate equation 1 by Poisson Pseudo Maximum Likelihood (PPML) as this method allows to accomodate zeros in left hand side variable, and is unbiased in the case of heteroskedasticity of the underlying multiplicative gravity model. We cluster standard errors at the department-pair level.

The fixed effect μ_{ij} absorbs time invariant factors at the department-pair level as distance, and cultural proximity. In particular, if there are unobserved time-invariant determinants for which department-pairs receive a high speed railway that connects them as for example business ties, μ_{ij} would absorb such determinants. The fixed effect μ_{it} absorbs time varying shocks at the origin department, as changes the determinants of supply or demand of innovation like changes in population and local policies. Similarly, the fixed effect μ_{jt} absorbs time varying shocks at the destination department.

In equation 1 the identification of β comes from across-time changes in copatents and travel time within a department-pair, relative to other department-pairs with the same origin department, conditional on time varying shocks to the destination department. Hence, the identification is not driven by whether a certain department becomes more central in either the copatent or train the network, or whether it has increases or decreases of population or economic activity. Rather, the identification comes from how the train connectivity of a department-pair evolves over time relative to other department-pairs starting from the same origin department.

The identification assumption is that the timing of roll out of high speed railways, and hence reductions in travel time, is exogenous to the copatent activity at the department pair level. This assumption is plausible because there was uncertainty about the opening dates of high speed railways. In robustness analysis we re-estimate equation 1 including only department pairs that do not have a high speed railway station, meaning that they benefited of high speed railways only indirectly, and find results that go in the same direction.

Figure 6 presents an event study version of equation 1, where department pairs can be treated multiple times at different intensity.[11] As we can see, there is little anticipation effect before the reduction in travel time.

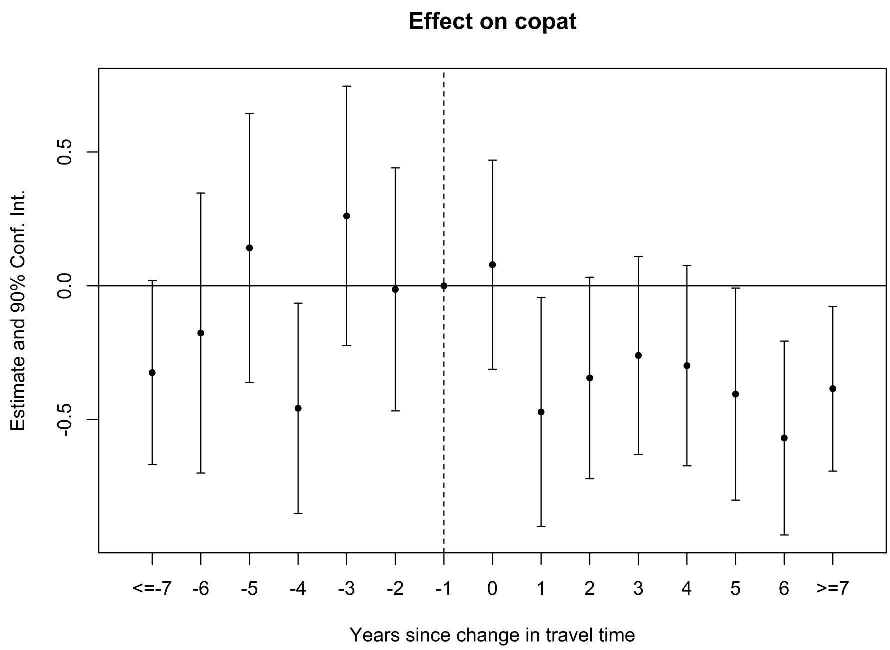

Figure 6: Event study scaled by change log(travel time)
The plot shows the point estimate and 90% confidence interval of $\text{copatents}_{ijt} = exp(\sum \beta_k \, \Delta_{t,t+k} log(\text{travel time}_{ij}) + \mu_{ij} + \mu_{it} + \mu_{jt}) \times \varepsilon_{ijt}$, normalized to the year previous to the change in log(travel time), t=-1. Confidence intervals are obtained with ij clustered standard errors.

Table 1 presents the results of estimating equation 1. Column 1 shows that the elasticity of copatents to travel time is -0.20, significant at the 10% level. In column 2 we open up the elasticity by distance between origin and destination departments. It is likely that travel time by train is not relevant at short distances, where the relevant measure may be car travel time, or at long distances, where the relevant measure may be airplane travel time. We find that the elasticity is largest in absolute value at distance between 100km and 400km, and it is imprecisely estimated distances under 100km or over 400km. While Figure 2 shows that the decrease in travel time was larger for longer distances, in 2010 the average travel time for distances over 400km in the effective sample was 5h 10min. Hence, it is likely that total travel

[11]We adapt the linear version of event study analysis with multiple treatment of different intensities of Schmidheiny and Siegloch (2023) into a non-linear PPML estimation.

time by airplane –after accounting for travel time to/from the airport and security checks– may still be lower than by train at distances over 400km.

	Co-patents	
	(1)	(2)
log(Travel time)	-0.199*	
	(0.113)	
log(Travel time) × distance < 100km		0.032
		(0.237)
log(Travel time) × 100km ≤ distance < 400km		-0.254*
		(0.133)
log(Travel time) × 400km ≤ distance		-0.136
		(0.187)
Fixed effects		
μ_{ij}	✓	✓
μ_{it}	✓	✓
μ_{jt}	✓	✓
Observations	122,714	122,714
Pseudo R^2	0.85	0.85

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 1: Effect of travel time on copatents

The table presents the result of estimating by PPML $\text{copatents}_{ijt} = \exp\left[\beta \log(\text{travel time}_{ijt}) + \mu_{ij} + \mu_{it} + \mu_{jt}\right] \times \varepsilon_{ijt}$, for departments i and j, and application year t. The sample includes application year from 1980 to 2010. Standard errors clustered at the non-directional department pair are presented in parentheses.

4.3 Effect of travel time: robustness

The identifying assumption for an unbiased estimate of β is that the timing of reductions of travel time at the department-pair is exogenous. This assumption is plausible given the uncertainty about the date at which high speed railways would be operative for each department pair. The concern then would be that there may be unobserved time-varying shocks at the department pair level that may be systematically correlated with the timing of opening of high speed railways. An example of such concern would be that an economic group, which is increasing or plans to increase its across-department integration, lobbies in order to improve the train connectivity.

To alleviate this concern we re-estimate equation 1 including only department pairs that do not have a high speed railway station. These department pairs see reductions in travel time only due to reductions in connecting segments, and hence it is unlikely that time-varying shocks to such department pairs are correlated with the reduction in travel time. The results in Table 2 show that the elasticity of copatents to travel time is -0.48 for department pairs without a high speed railway station.

	Co-patents	
	(1)	(2)
log(Travel time)	-0.199*	-0.478*
	(0.114)	(0.273)
Fixed effects		
μ_{ij}	✓	✓
μ_{it}	✓	✓
μ_{jt}	✓	✓
Sample selection	All	No HSR station
Observations	122,714	74,100
Pseudo R^2	0.85	0.78

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 2: Robustness: department-pairs without HSR station

4.4 Effect of travel time: heterogeneity

In this section we open up copatents and department pairs by different characteristics. Importantly, we find that the reduction in travel time leads to more collaboration between more developed department pairs, as proxied by their population density, rather than between those and less developed ones, or among less developed ones. Additionally, we uncover that the effect is larger for across-firm copatents, and while the point estimate of reductions of travel time is similar to both new and pre-existing collaborations across inventors, it not precisely estimated for pre-existing collaborations. We also find that the effect persists once we weight patents by amount of claims or technology classes, proxies for patent scope and patent breadth, implying that the reduction in travel time does not increase copatents by reducing the scope or breadth of each patent.

Core-Periphery analysis

We classify departments as Core or Periphery based on the Eurostat classification of NUTS3 regions (departments in the case of France). Eurostat classifies regions as urban (80% of population lives in urban clusters), intermediate (between 50% and 80% of population lives in urban clusters) and rural (at least 50% of population lives in rural areas).[12] We consider urban and intermediate as Core, and rural as Periphery. Appendix Figure 2 shows a map of departments colored by Core/Periphery status.

Table 3 shows the results of estimating the elasticity of copatents to travel time by whether both departments in the pair are Core, Periphery, or one Core and the other Periphery. We find that the elasticity is only statistically significant for Core-Core pairs. This result shows that reductions in travel time reductions in travel time led to an increase in patent collaboration only between department pairs

[12] We use the 2024 classification retrieved from the R package eurostat (Lahti et al. (2017)). See `https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Territorial_typologies_manual_-_urban-rural_typology` for more details on this classification.

in which both departments are more developed. This result shed light into a new potential consequence of reductions in travel time, while it may lead to an increase in patent collaborations, it may at the same time led to increased inequality in terms of collaborative innovation, with more developed department pairs strengthening innovation ties with other developed departments. In Appendix Table 5 we show the results using 1975 population density to classify departments under/over median density, finding results that go in the same direction.

	Co-patents
log(Travel time) × Core-Core	-0.268**
	(0.116)
log(Travel time) × Core-Periphery	0.396
	(0.305)
log(Travel time) × Periphery-Periphery	-0.158
	(1.43)
Fixed effects	
μ_{ij}	✓
μ_{it}	✓
μ_{jt}	✓
Observations	122,726
Pseudo R^2	0.85

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 3: Copatents by department-pair type

The table presents the result of estimating by PPML $\text{N. co-patents}_{ijt} = \exp[\beta \log(\text{Travel time}_{ijt}) + \rho_{ij} + \gamma_{it} + \delta_{it}] \times \eta_{ijt}$, for departments i and j, and application year t. Sample includes application year from 1980 to 2010. The regression is estimated as directional (ij different from ji). Standard errors clustered at the non-directional department pair are presented in parentheses.

Type of collaboration

Table 4 shows the results of estimating equation 1 where copatents are classified into whether the copatent belows to an inventor-pair that did not collaborate in the past (column 2) or it did (column 3), and whether the patent involves only one firm (column 4) or multiple firms (column 5). We find that the point estimate elasticity is comparable in new and old collaborations, though it is only statistically significant for new collaborations. This result suggests that reductions in travel time lead to new collaborations, but once the collaboration exist, further reductions in travel time may not be as relevant. We note, however, that as patenting is a rare event at the inventor level there is likely less statistical power for estimating the effect on continuing-collaborations. We also find that the elasticity is larger in absolute value and statistically significant for across-firm collaborations, and not statistically significant for within firm collaborations. This result suggests that face to face interactions are relevant for across firms collaborations, but that within firm collaborations may be less reliant on repeated interactions that are affected by changes in travel time. This result is related to Giroud (2013) who finds that non-stop flight connections between a subsidiary and its headquarters leads to increased investment in the subsidiary.

	Baseline (1)	New collab (2)	Old collab (3)	Within firm (4)	Across firms (5)
log(Travel time)	-0.199*	-0.246**	-0.247	-0.134	-0.403*
	(0.113)	(0.108)	(0.294)	(0.132)	(0.230)
Fixed effects					
μ_{ij}	✓	✓	✓	✓	✓
μ_{it}	✓	✓	✓	✓	✓
μ_{jt}	✓	✓	✓	✓	✓
Observations	122,714	121,742	32,691	108,950	47,813
Pseudo R^2	0.85	0.83	0.81	0.85	0.59

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 4: Copatents by type of collaboration

Patent characteristics

We investigate wether the decrease in travel time leads to more patents by leading to patents of different quality. We estimate a variation of equation 1 weigthing patents by the amount of citations received in a 5-year window, amount of claims in the patent, and amount of technology classes included in the patent. Table 5 shows the results.

When weighting patents by the amount of citations received, we find that the reduction of travel time, if any, decreased the amount of citation-weighted copatents. However, given that most patents have zero citations, the identification of fixed effects drops around two thirds of the department-pair-year observations of the baseline estimation in Table 1.[13] As the underlying effective sample is so different, one should be cautious when comparing coefficients across regressions.[14]

Column (2) of Table 5 presents the results with copatents weighted by amount of claims. The amount of claims in a patent is a way to measure the patent scope – the amount of "things" over which the patent is claiming intellectual property rights (Lanjouw and Schankerman (2004)). Comparing the coefficient with the baseline estimation, we find a larger coefficient, suggesting that the reduction in travel time led to copatents with larger scope. Column (3) presents the results weighting each copatent by the amount of different technology classes that it has.[15] The amount of technology classes in a patent is a way to measure the patent breadth – the amount of technological domains that it covers (Lerner (1994)). The

[13]In the effective sample of the baseline regression shown in Table 1, the average observation at the department-pair-year has 1.17 copatents while it has 0.30 citation weighted copatents. In the effective sample, 79% of observations have zero copatents, while the share increases to 94% when considering citation-weighted copatents.

[14]Estimating equation 1 in the effective sample of department-pair-year of model (1) in Table 5 gives an elasticity of copatents to travel time -0.155 with clustered standard error 0.118. Using the same effective sample, the results for claims weighted and technology class weighted copatents are quantitatively similar to the ones in Table 5.

[15]We count the amount of different IPC35 technology classes, having a maximum of 35 technology classes.

reduction in travel time led to a comparable increase in technology-weighted copatents as unweighted copatents, suggesting that the reduction in travel time did not affect the technological breadh of copatents.

	Citation weighted	Claims weighted	Tech. classes weighted
	(1)	(2)	(3)
log(Travel time)	0.452*	-0.402***	-0.241**
	(0.241)	(0.139)	(0.113)
Fixed effects			
μ_{ij}	✓	✓	✓
μ_{it}	✓	✓	✓
μ_{jt}	✓	✓	✓
Observations	39,477	122,726	122,726
Pseudo R^2	0.78	0.90	0.87

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 5: Copatents weighted by patents' characteristics

Inventor characteristics

We investigate whether the reduction in travel time leads to more copatents among technologically similar or different inventors. We compute the technological similarity of inventors based on their previous patents and then classify copatents by whether the inventor-pair is over or under the median similarity across inventors.[16] Table 6 shows that the decrease in travel time led to increased copatents among inventors under and over the median similarity.

	Similarity > median	Similarity < median
	(1)	(2)
log(Travel time)	-0.314**	-0.260**
	(0.132)	(0.128)
Fixed effects		
μ_{ij}	✓	✓
μ_{it}	✓	✓
μ_{jt}	✓	✓
Observations	74,010	59,427
Pseudo R^2	0.83	0.83

Signif. Codes: ***: 0.01, **: 0.05, *: 0.1

Table 6: Copatents by inventors' technological similarity

[16]We compute the cosine similarity of technologies using the IPC35 technologies of previous patents.

5 Conclusion

This paper exploits the roll out of High Speed Railways (HSR) as a quasi-natural experiment to provide new causal evidence of the impact of train travel time on patent collaboration between inventors. To do so, we constructed a new dataset of train travel time in France between 1980 and 2010, documenting a median decrease in travel time between department-pairs of 25%. We find that the decrease in travel time led to an increase in patent collaboration across departments, driven by collaborations between inventors that have not collaborated before, and by collaborations across firms. At the same time, we find that the reduction in travel time only leads to an increase in collaborations between department pairs in which both departments with higher population density, which are likely the more developed ones. As consequence, the decrease in travel time could increase the collaboration gap between more developed and less developed departments.

This paper provides evidence on the impact of transportation infrastructure on innovation. Recent policies and recommendations make a strong emphasis on fostering regional innovation. Our results suggest that transportation infrastructure can affect regional innovation by facilitating long-distance face-to-face interactions among inventors. However, the gains from improvements in connectivity may be unevenly distributed, favoring more developed regions.

References

Agrawal, Ajay, Alberto Galasso, and Alexander Oettl. 2017. "Roads and innovation." *Review of Economics and Statistics*, 99(3): 417–434.

Atkin, David, M Keith Chen, and Anton Popov. 2022. "The returns to face-to-face interactions: Knowledge spillovers in Silicon Valley." National Bureau of Economic Research.

Bahar, Dany, Prithwiraj Choudhury, Do Yoon Kim, and Wesley W Koo. 2023. "Innovation on wings: Nonstop flights and firm innovation in the global context." *Management Science*, 69(10): 6202–6223.

Belenzon, Sharon, and Mark Schankerman. 2013. "Spreading the word: Geography, policy, and knowledge spillovers." *Review of Economics and Statistics*, 95(3): 884–903.

Carlino, Gerald, and William R Kerr. 2015. "Agglomeration and innovation." *Handbook of regional and urban economics*, 5: 349–404.

Catalini, Christian. 2018. "Microgeography and the Direction of Inventive Activity." *Management Science*, 64(9): 4348–4364.

Dijkstra, Edsger W. 1959. "A note on two problems in connexion with graphs." *Numerische mathematik*, 1(1): 269–271.

Dong, Xiaofang, Siqi Zheng, and Matthew E Kahn. 2020. "The role of transportation speed in facilitating high skilled teamwork across cities." *Journal of Urban Economics*, 115: 103212.

Draghi, Mario. 2024. "The future of European competitiveness. Part A: A competitiveness strategy for Europe." Accessed: 2025-02-28.

Duranton, Gilles, Philippe Martin, Thierry Mayer, and Florian Mayneris. 2009. "The economics of clusters: evidence from France."

Giroud, Xavier. 2013. "Proximity and investment: Evidence from plant-level data." *The Quarterly Journal of Economics*, 128(2): 861–915.

Gross, Daniel P, and Bhaven N Sampat. 2023. "America, jump-started: World War II R&D and the takeoff of the US innovation system." *American Economic Review*, 113(12): 3323–3356.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic localization of knowledge spillovers as evidenced by patent citations." *the Quarterly journal of Economics*, 108(3): 577–598.

Lahti, Leo, Janne Huovari, Markus Kainu, and Przemyslaw Biecek. 2017. "Retrieval and Analysis of Eurostat Open Data with the eurostat Package." *The R Journal*, 9(1): 385–392.

Lanjouw, Jean O, and Mark Schankerman. 2004. "Patent quality and research productivity: Measuring innovation with multiple indicators." *The economic journal*, 114(495): 441–465.

Lerner, Joshua. 1994. "The importance of patent scope: an empirical analysis." *The RAND Journal of Economics*, 319–333.

Moretti, Enrico. 2012. *The new geography of jobs.* Houghton Mifflin Harcourt.

Morrison, Greg, Massimo Riccaboni, and Fabio Pammolli. 2017. "Disambiguation of patent inventors and assignees using high-resolution geolocation data." *Scientific data*, 4(1): 1–21.

Pauly, Stefan, and Fernando Stipanicic. 2022. "The creation and diffusion of knowledge: Evidence from the Jet Age."

Schmidheiny, Kurt, and Sebastian Siegloch. 2023. "On event studies and distributed-lags in two-way fixed effects models: Identification, equivalence, and generalization." *Journal of Applied Econometrics*, 38(5): 695–713.

Silva, JMC Santos, and Silvana Tenreyro. 2006. "The log of gravity." *The Review of Economics and statistics*, 88(4): 641–658.

Tsiachtsiras, Georgios, Deyun Yin, Ernest Miguelez, and Rosina Moreno. 2022. "Trains of Thought: High-Speed Rail and Innovation in China." *Available at SSRN 4280769*.

U.S. Congress. 2022. "CHIPS and Science Act of 2022." Public Law No: 117-167, Accessed: 2025-02-28.