Ticket to Impact: Estimating the impact of transport connectivity on collaborative research projects[1]

Carolin Ioramashvili[1,2] & Nicholas Sweeney[3]

[1] Science Policy Research Unit, University of Sussex, c.ioramashvili@sussex.ac.uk

[2] Department of Geography & Environment, London School of Economics

[3] University of Manchester

1. Introduction

Modern science and research is increasingly collaborative, involving larger and more diverse and interdisciplinary teams. Despite advances in communication technology, face-to-face interaction still matters, to convey complex information, establish trust, facilitate informal knowledge sharing, or simply to use shared tools and equipment. Indeed, research grants often include substantial funding to support collaborator meetings or conference attendance. However, suitable transport infrastructure is also required to allow collaborators to get together. In the UK context, debates on new transport infrastructure, particularly over rail, have been contentious, including the HS2 high-speed link between London and Birmingham, with links to the North now being cancelled, regional links such as 'Crossrail for the North', linking large cities in the North West and Yorkshire, and city-level projects such as a tram for Leeds. In these policy debates, both the taxpayer value of such projects as well as the distribution of costs and

[1] This project has been supported by an Innovation and Research Council GtR+ grant, under grant number FFGTR001. We are grateful to the Department for Science, Innovation and Technology to making the Gateway to Research + data available for this work.

benefits have been hotly contested. Conversely, the newly opened Elizabeth Line in London has already surpassed passenger expectations in the first years of its operation.

Connectivity affects collaborative research projects in two distinct ways. First, connectivity may affect which projects take place. Researchers may decide that a collaboration is too cumbersome or its viability is reduced if transportation costs become prohibitive. This may exclude some researchers and organisations from collaborative projects, exacerbating regional inequalities. For example, there is evidence that public research funding in the UK is disproportionately concentrated in London and the South East (Jones & Forth, 2020). It may also mean that the collaborations that do happen are suboptimal, in that more suitable collaborations partners drop out of projects, substituted with less suitable but more accessible partners. For projects that do take place, poor connectivity may mean that the projects are less productive than they otherwise would be if research teams have to operate in isolation. Evidence suggests that face-to-face interaction and close collaboration are particularly important for developing high-impact, groundbreaking ideas.

It is well established that the germination, development and sharing of new ideas benefits from face-to-face interaction. New, particularly complex and contextual knowledge is described as tacit, or non-codifiable (Polanyi, 1966). This idea underpins theories of why dense cities are more innovative (Storper & Venables, 2004). Empirically, it has been shown that knowledge is surprisingly geographically sticky (Storper & Sonn, 2008; Van der Wouden & Youn, 2023) and that breakthrough ideas are more likely to emerge from urban areas (Berkes & Gaetani, 2021). Furthermore, face-to-face interaction including through conferences (Chai & Freeman, 2019), non-stop flights (Bahar et al., 2023) and high-speed rail (Kang et al., 2023) increases the likelihood of collaboration. In this, vein, the research question for this project asks: What is the impact of transport connectivity on research productivity in collaborative research projects?

The present report adds some specific evidence to these debates by asking how rail connectivity impacts collaborative research projects. Poor connectivity may hamper the performance of collaborative research projects, if researchers are not able to meet in person, or not as frequently as planned. Poor connectivity may also hamper researchers to participate in such projects in the first place.

2. Data and methods

The analysis combines data on collaborative research projects funded by UKRI, travel times between project partners, and information on rail performance from the Office for Road and Rail (ORR). Combining information on projects, travel times and short-term variation in connectivity from variation in performance allows to model both selection into collaboration as well as the impact of connectivity on project performance.

UKRI is the UK's main funding agency for research and innovation. It was formed in 2018 by bringing six discipline specific research councils, Innovate UK, Research England, and the Science and Technology Facilities Council under one roof. Through these various arms, UKRI funds training for researchers and scientists, basic and applied research carried out in universities and independent research organisations, as well as research, development and innovation activities of businesses, charities and social enterprises. Consistent information is available from projects started in 2006.

UKRI funding awardees are required to report on project outputs and impacts. Together with details on the project objectives and team, these are reported on the Gateway to Research portal (GtR), accessible to the public. This project relies on the GtR+ dataset, which is collated by DSIT and provides some improvements on the public version, including deduplication of organisations, identification of organisations' sectors, and tagging of disciplines and technologies. For the purposes of the analysis presented here, this project focuses on research collaborations with two organisations located on the British mainland.[2] The focus is on organisations rather than individuals as organisations can be used to determine the location of where the project work is happening. Collaborative projects may have more than two researchers working in them, as long as they are associated with only two organisations, while projects with multiple researchers all in the same organization are excluded.

The number of publications, including journal publications, books and working papers, is used as a proxy for the performance of a project. While publications are widely used as a measure of performance of researchers, including in the Research Excellence Frameworks and other UKRI assessments, it is recognized that this is a flawed measure. The sheer number of

[2] Projects with collaborators in Northern Ireland or overseas had to be disregarded as the main focus is on rail connections, which are only available on the UK mainland.

publications does not take into account quality and scholarly publications would not be expected to be an intended outcome of many R&D and innovation projects, particularly those funded by Innovate UK. To limit the sample to projects where publications are at least applicable, we limit the sample to projects with at least one publication[3].

Using the sample of projects involving two organisations, we use the TravelTime API to calculate the travel time between the organisations postcodes. We collect travel times using public transport, a car or a combination of the two, i.e. driving to the station and then taking the train. The sample is capped at a travel time of four hours From the public transport connection, we also collect routing information. Crucially, this provides information on the operator of the trains someone travelling between the two locations would have to take. For example, for a collaboration between University College London and the University of Birmingham, the fasted connection would be to first take an Avanti Westcoast train and then a West Midlands Railway train.

Information on the train operators running services between two organisations then allows linking projects to information on rail performance. Specifically, the Office for Road and Rail publishes quarterly performance indicators[4]. The focus here is on the Public Performance Measure (PPM) indicating the share of trains arriving on time (within 5 minutes of schedule for commuter trains and 10 minutes for long-distance trains). The sample is restricted to projects with organisations connected by at least one mainline rail line. Therefore, projects within the same city connected by underground, tram or bus services are excluded[5]. We assign to the PPM of the worst performing operator connecting two project partners to a project. For connections requiring a change between operators, the rationale is that the overall delay will be determined by the worst operator, as a delay on an early operator may lead to missed connections later on.

While the PPM has been reported on a quarterly basis since 1997, we are only able to collect information on contemporaneous connections. While there were no substantial changes to the network over this period, service patterns may have changed that mean that the overall travel

---

[3] This approach also eliminates ‘false zeros’ due to a failure to report from the project team.

[4] We are grateful to the Information & Analysis Team at the Office of Rail and Road for making available historical performance metrics.

[5] Projects connected by a combination of mainline and local public transport services are included, e.g. requiring to take a London Underground train to get to a mainline terminal.

time as well as the fastest connection have change, but overall impacts would be expected to be minor.

An overview of the sample is presented in Table 1. The sample consists of 2,599 projects that received just over £420,000 in funding on average each, for a duration of just over 2.5 years (10.8 quarters). On average, 4 publications are attributable to each project, but there is a wide variation in outcomes, ranging between 1 and 26. The vast majority (90%) fall into the category of research projects. This is unsurprising, given that most fellowships are awarded to individuals. Similarly, around 90% of lead organisations are universities. On average, collaborative research projects bridge substantial distances: about 2 hours and 15 minutes driving (or driving and public transport combined), or just under 2.5 hours by public transport. Again, there is substantial variation, between partners just being minutes away from each other, up to the upper threshold of four hours. Likewise, rail performance is variable. Measured as the average PPM between the two partners over the duration of the project, the share of trains on time ranged from 64% to 97%, with an average of 84.8%.

Table 1: Summary statistics

	Mean	Std. dev.	Min	Max
Project funding (£)	420,297.8	614,419.3	1,480	9,600,000
Project duration (quarters)	10.8	5.5	0	44
Total publications	4.1	2.6	1	26
Grant category				
Research	0.9	0.3	0	1
Fellowship	0.09	0.3	0	1
Training	0.004	0.06	0	1
Other	0.02	0.1	0	1
Sector of lead organisation				
Private	0.006	0.08	0	1
Public	0.05	0.2	0	1
Charity/Non-profit	0.003	0.06	0	1

University	0.9	0.2	0	1
Travel time between project partners in minutes				
Driving	136.6	54.4	7	240
Public transport	143.1	51.0	9	240
Driving and public transport	136.1	52.0	7	240
Average PPM during project	85.8	4.9	64	97
Straight line distance (km)	138.5	84.8	2	403
Number of observations	2,599			

3. Findings

First, we take a broad look at connectivity between the UKRI funded projects in the sample. The modal project is in the 1-2 hour travel time bracket by either driving, public transport or a combination of the two. The number of projects declines with increasing travel time, particularly at more than three hours, where a return journey is realistically not feasible in a day. On the opposite side, there is also a spike in local collaborations with a travel time of under 30 minutes.

Next, we turn to the relation between travel time and project size. Figure 2 shows that project size increase with travel distance between partners. Several factors may be driving this. On the one hand, a bigger budget may allow for more travel. On the other, a more ambitious project may warrant collaboration with more specialized partners, or may make up for the inconvenience of distance. This is not necessarily an expected finding. Prior literature in the economics of innovation also shows that riskier projects may require more face to face interaction, which may be hindered by distance. However, it suggests that there is strong selection at play in terms of how researchers form collaborations.

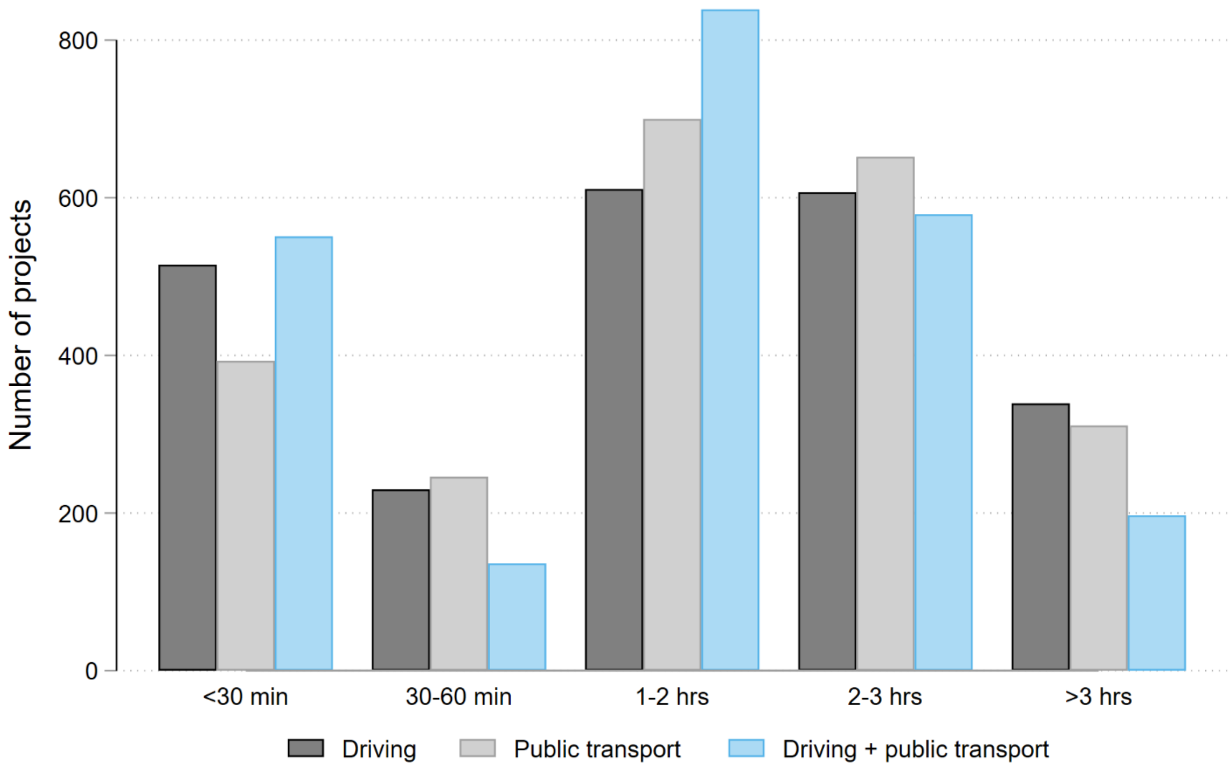

Figure 1: Number of projects by travel time

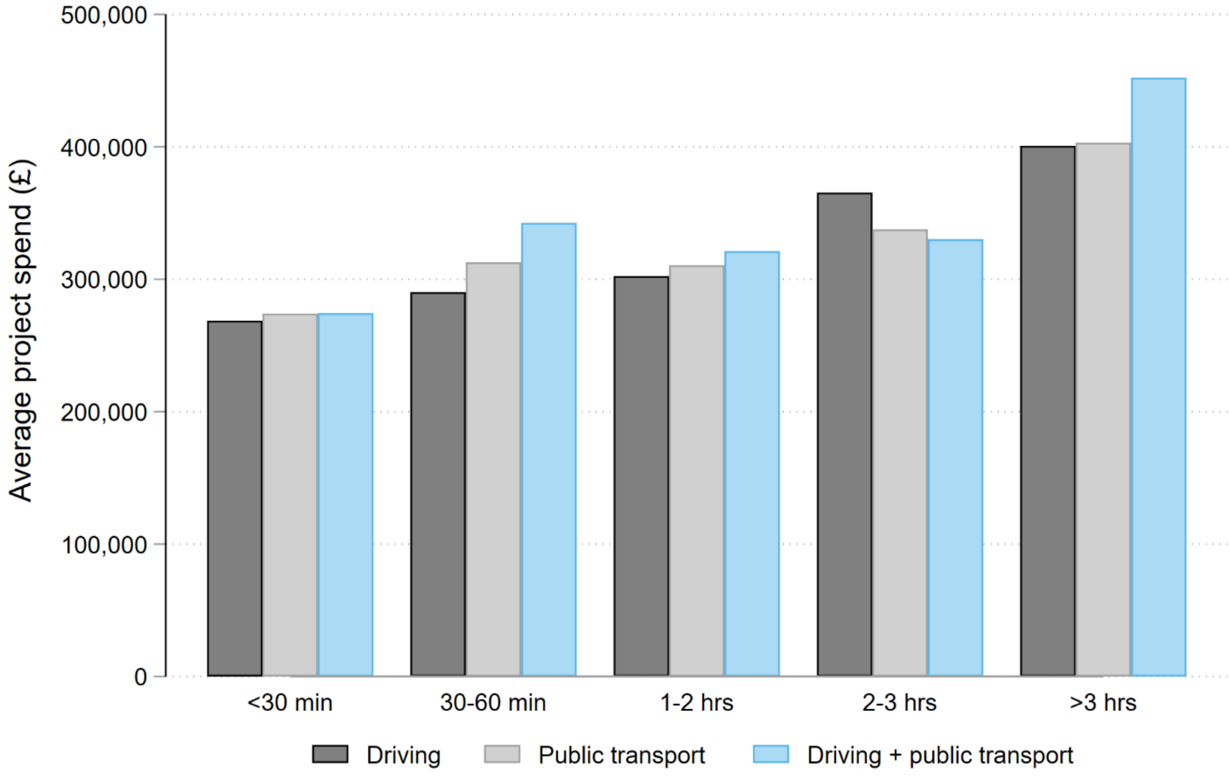

Figure 2: Project spend by travel time

From a regional angle, there is significant variation in terms of how long project partners have to travel, on average. Figure 3 shows average travel distance by region, from the perspective of either partner. That means that projects with partners in two different regions are considered twice, for the measure of both regions. Despite its relative remoteness in a UK context, projects with a Scottish partner have on average the shortest travel time. This might indicate a bias in that Scottish organisations are more likely to collaborate with other local organisations, reflecting devolved policy responsibility, particularly for higher education. Conversely, for projects with partners in Wales, which also has a devolved administration, albeit with less power than the Scottish one, travel times are among the highest across all regions. In England, travel times are lowest within London, followed by the East of England and the South East, reflecting generally better transport infrastructure as well as a higher density of universities and research organisations than in the rest of the country.

Comparing differences between modes of travel within regions is also instructive. In most regions, a combination of driving and public transport is the fastest option, and driving the whole distance takes the longest, but in Scotland, the South East and the West Midlands, driving is faster than taking only public transport, suggesting a lack of local transport option, e.g. to get to a train station. In most regions, the difference between the different modes of travel is small, with the widest gaps recorded for projects with partners in London and the North East. While the remainder of the analysis will focus on public transport connectivity only, the differences in travel time across modes is important to keep in mind as researchers may switch to driving if train travel becomes too unreliable.

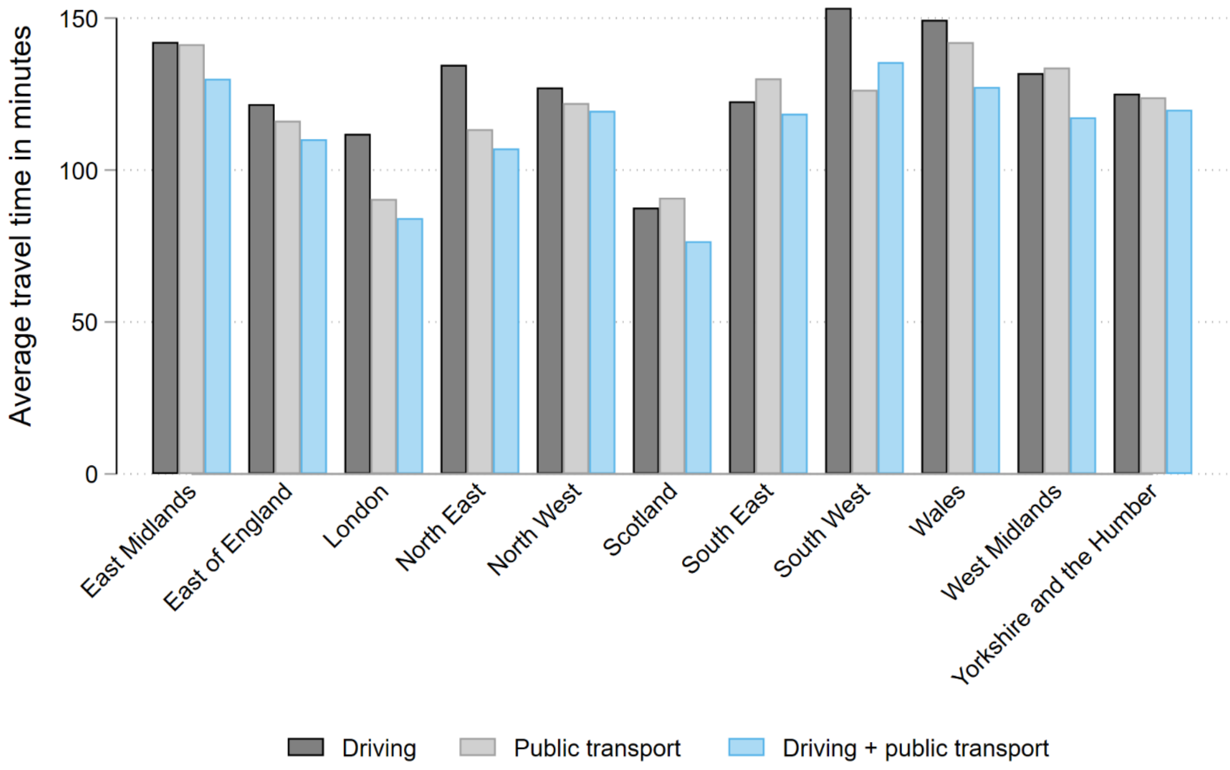

Figure 3: Travel time by region

Next, we turn to the reliability of public transport as measured by the Public Performance Measure (PPM) of mainline rail operators. Figure 4 shows the PPM by operator on a quarterly basis between 1997 and 2024. Each line represents an operator. As there are 22 operators, these are not labelled. However, the figure highlights substantial variation in performance both across operators as well as over time. Performance improved between 2000 and 2010 and then stayed at a high level until 2015, when it started slowly declining. In 2020 to 2022, when many services ran at a sharply reduced frequency due to Covid-19 restriction, performance was generally high, but then fell sharply. At times of good performance, all operators tend to perform well, whereas there is more variability during times of on average lower performance. For further analysis, this provides some reassurance that performance is not driven only by national trends but that there is local heterogeneity that can be exploited.

However, there is relatively little variation by project region. As before, we take the average PPM by region of both project partners over the duration of the project. As figure 5 shows, there is minimal variation across regions. This is unsurprising, as operators do not neatly map onto regions, with most regions being served by multiple operators both for short- and long-distance travel.

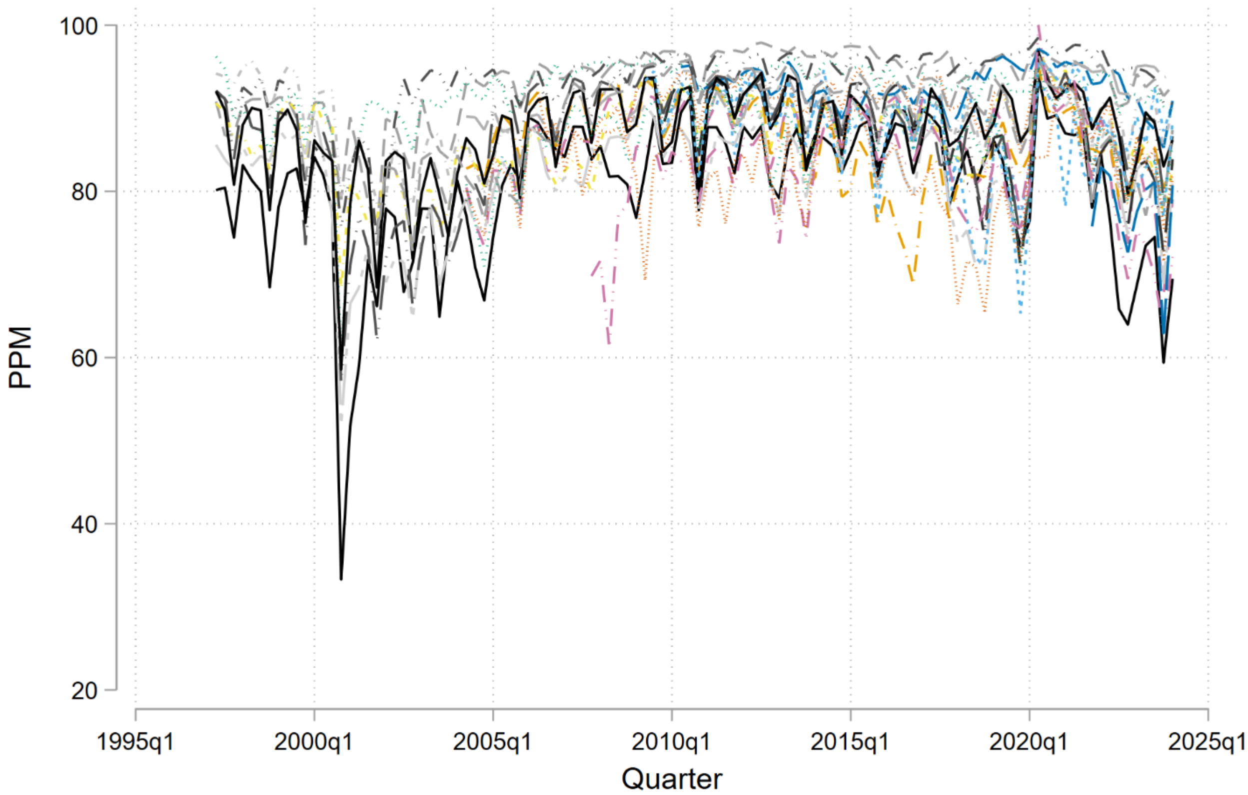

Figure 4: Public Performance Measure by operator over time

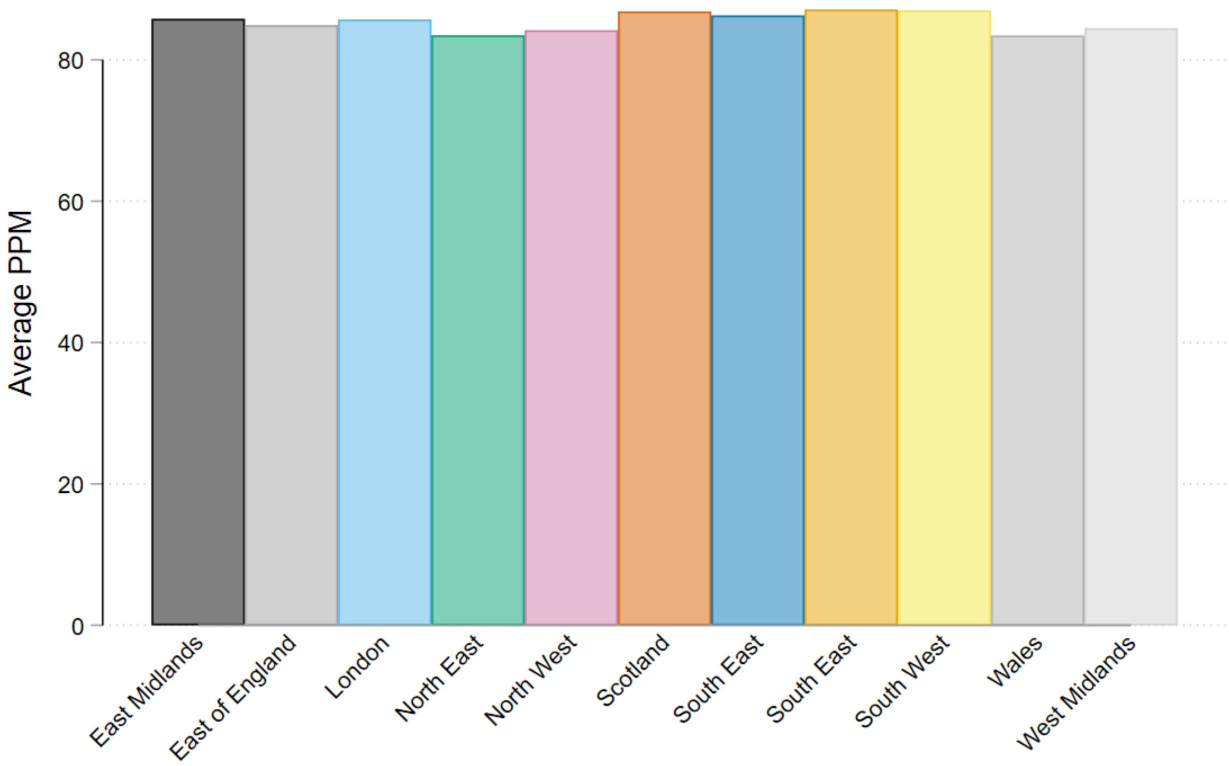

Figure 5: Public Performance Measure by project region

Connectivity and project performance

To summarise, there is some heterogeneity in general connectivity across projects, as well as in the performance of rail operators and hence the experienced connectivity of researchers, irrespective of distance and expected travel time. Next, we turn to the impact connectivity has on project performance. As shown above, travel times are longer for larger projects, which would be expected to deliver more publications. It is therefore difficult to disentangle the link between connectivity and performance as selection and project size are confounding factors. While it is possible to control for size, other factors contributing to selection are unobserved. While general travel times by public transport are readily observable and relatively stable over time, we will argue that public transport delays as measured by the PPM provide a source of variation in connectivity that is not expected at the time of project conception.

Therefore, to arrive at the impact of connectivity on project performance, we regress publications on the PPM during the course of the project. The model is estimated both using

Ordinary Least Squares (OLS), with the dependent variable as log publications, as well as using Poisson regression, which is more suitable for count data. Furthermore, we estimate a simple model as well as models controlling for the PPM prior to project start to account for anticipation effects, log total project funding, the grant category, the sector of the lead organization, project length, start year or the project and straight line distance in logs. Furthermore, we split the sample between social and natural sciences[6] to account for differences in the nature of research and collaboration across differences.

Figure 6 shows the regression coefficient of interest of the PPM over the duration of the project[7]. The point estimates indicate the expected change in publications if the PPM increases by one point. The bars around the central coefficients indicate the 95% confidence intervals. The first coefficient is from the simple OLS model. It suggests that for a one point increase in the PPM, the number of publications associated with a project increases by around 2.8%. The coefficient becomes smaller once other variables are controlled for, but is still statistically significant at 0.015. This suggests that for at 10 point increase in the PPM, the number of publications increases by 15%, or just over half an extra publication for the average project. The next two models, estimated using the Poisson estimator, suggest slightly smaller but qualitatively unchanged coefficients.

In the bottom half of figure 6, the sample is split into social sciences and natural sciences. All of these models control for the full set of control variables. The sample of social science projects is significantly smaller, with only 496 observations. This results in wide confidence intervals and an imprecise estimation of impacts that are overall statistically insignificant from zero. For the natural sciences, the estimated impacts are similar to, and slightly larger than those for the full sample.

[6] There is no definitive way to draw this distinction. First, we assign projects by funder, with projects funded by the Economic and Social Research Council and the Arts and Humanities Research Council labelled as social science, and projects funded by other research councils labelled as natural sciences. For projects not directly funded by research councils, or by UKRI-wide grants, we utilise the Centre for Science and Technology Studies (CWTS) taxonomy as assigned by DSIT. There is a straightforward Social Sciences label, and we assign the Natural, Physical Health and Life Sciences categories to the Natural Sciences. For remaining projects that are not labelled, which are predominantly funded by Government departments, we assign projects funded by DCMS, DFID and DWP to the social sciences and all others to the natural sciences.
[7] Full regression results can be found in Appendix Table 1.

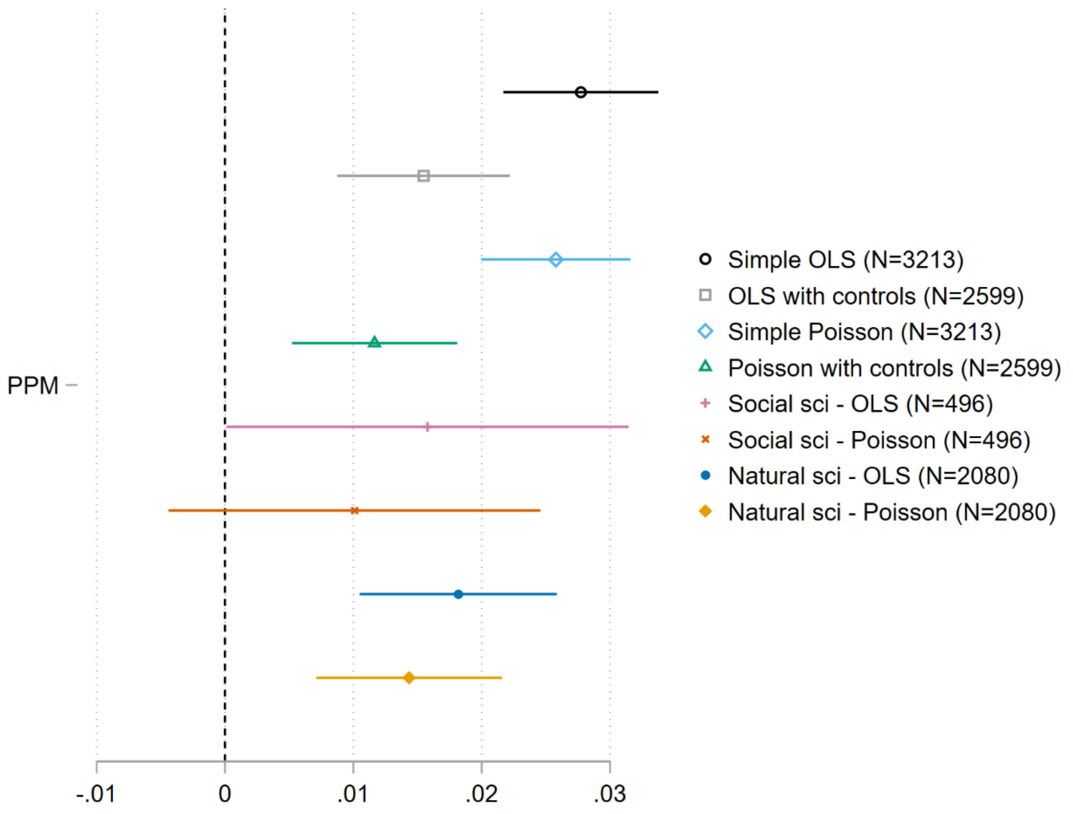

Figure 6: Impact of PPM on project performance

Overall, these findings suggest that transport connectivity, and in particular reliability, as a significant impact on collaborative research projects and better connectivity results in a higher number of project outputs. To test this mechanism further, the following estimations split the sample by train performance prior to the start of the project. Prior performance is measured over the three quarters before the quarter when the project started. Performance is considered to be 'good' if the PPM is on average above 90 and above 85 in each of the three quarters. Conversely, performance is considered 'bad' if the PPM is on average below 80 and below 85 in each individual quarter. Projects associated with an average PPM between 80 and 90 or with strong variation across three quarters are considered 'mixed'.

Similar to the results presented above, figure 7 presents a coefficient plot of point estimates with 95% confidence intervals for the three samples of projects with prior good, bad and mixed rail performance[8]. We estimate OLS and Poisson models, each including the full set of control variables. For projects with prior good rail performance, the estimated point estimates are smallest, and statistically insignificant from zero. This suggests that for collaborators that enjoyed a good connection around the time the project was planned, subsequent performance does not have an impact. However, there is significant effect for projects with a previously bad connection, and this is estimated to be larger – between 0.02 and 0.03 – than for the whole sample. The impact is estimated to be similar to the whole sample for projects that fall in the middle with average and/or mixed performance, and the estimated impacts remain statistically significant.

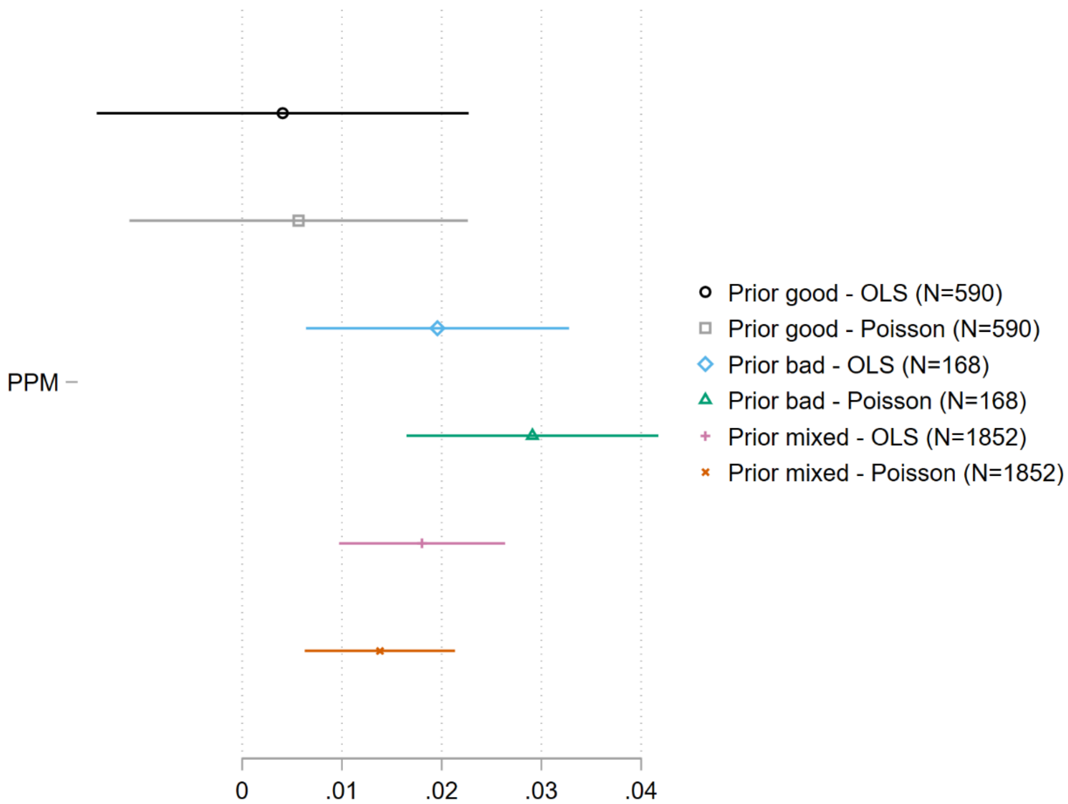

Figure 7: Impact on project performance by rail performance prior to project start

[8] Full regression results can be found in Appendix Table 2.

To interpret these results, it helps to consider rail performance separately for the three groups. Table 2 shows the PPM over the project duration by performance prior to start. As can be seen, for projects with prior good connectivity, performance drops slightly on average after the start of a project, while that for projects with prior bad performance remains low, below 80. However, for projects with bad performance in the run-up to project start, there is a wider range of performance, between 64 – less than two-thirds of trains are on time – and 94 – almost all trains are on time. Further testing reveals that the positive coefficients for the 'Prior bad' sample is driven by projects that continue to suffer from poor performance after the start of the projects. Hence, poor performance in this sample is associated with fewer publications.

Table 2: Average PPM by performance prior to project start

Performance before project start	Mean	SD	Min	Max
Good	89.0	3.3	75	97
Bad	78.6	6.9	64	94
Mixed	85.4	3.7	69	94

In the appendix, we provide two robustness checks for these results. Figure A1 shows regression coefficients with the PPM prior to project start as the explanatory variable. If the reasoning here is correct, that connectivity affects the ability of project partner to collaborate, rail performance prior to project start may affect the propensity to collaborate, but should not affect performance itself. In general, the results presented in figure A1 are reassuring, except for the sample of natural science projects, where there is a small positive effect, hinting at potential selection effects, which will be probed further in the next section.

In figure A2, travel time instead of rail performance is entered as the explanatory variable. As discussed above, travel time is confounded with multiple project characteristics, including (likely positive) size effects and (likely negative) selection effects. Indeed, the regression results indicate no significant correlation between travel time between partners and the number of project publications.

Connectivity and project participation

To summarise, poor rail performance is associated with fewer research outputs from collaborative projects that are affected by delays. Given the importance placed on publications both for individuals' career progression as well as organizational assessments such as the REF, unequal connectivity may have significant impacts on individuals and organisations. To probe these impacts further, we also investigate the impact of connectivity on the probability to collaborate. As discussed above, travel time increases project size, suggesting that there are 'missed' collaborations in smaller projects between more distant partners. There is also selection bias in that we only observe the tip of the iceberg of successful project applications, and not the vast majority of projects applications that are not funded. Naturally, strong selection effects are at play in terms of which partners collaborate, as researchers are highly specialized, choosing partners carefully, and a collaborative project is a significant commitment from researchers and their institutions. To control for these unobserved selection effects, we estimate the propensity for two organisations to collaborate again after an initial collaboration. An initial collaboration signals general compatibility, while the decision to repeat a collaboration may be driven by the experience during the first collaboration.

To test for this mechanism, we focus on first collaborations of two specific project partners, and then define a variable equal to one if they ever collaborate again, or zero otherwise. There are several shortcomings with the approach. Large research organisations collaborate across many different fields. Repeat collaborations do not necessarily involve the same individuals, so the link between collaborations over time may be weak. Additionally, more recent projects would have less time to form another successful collaboration.

With these caveats in mind, the somewhat inconclusive results presented in figure 8 are perhaps unsurprising. For the simple probit model, there is a positive correlation between rail performance during the initial project and the propensity to collaborate again. However, this significant effect disappears once controls are introduced, and not statistical differences are observable between the social and natural sciences.

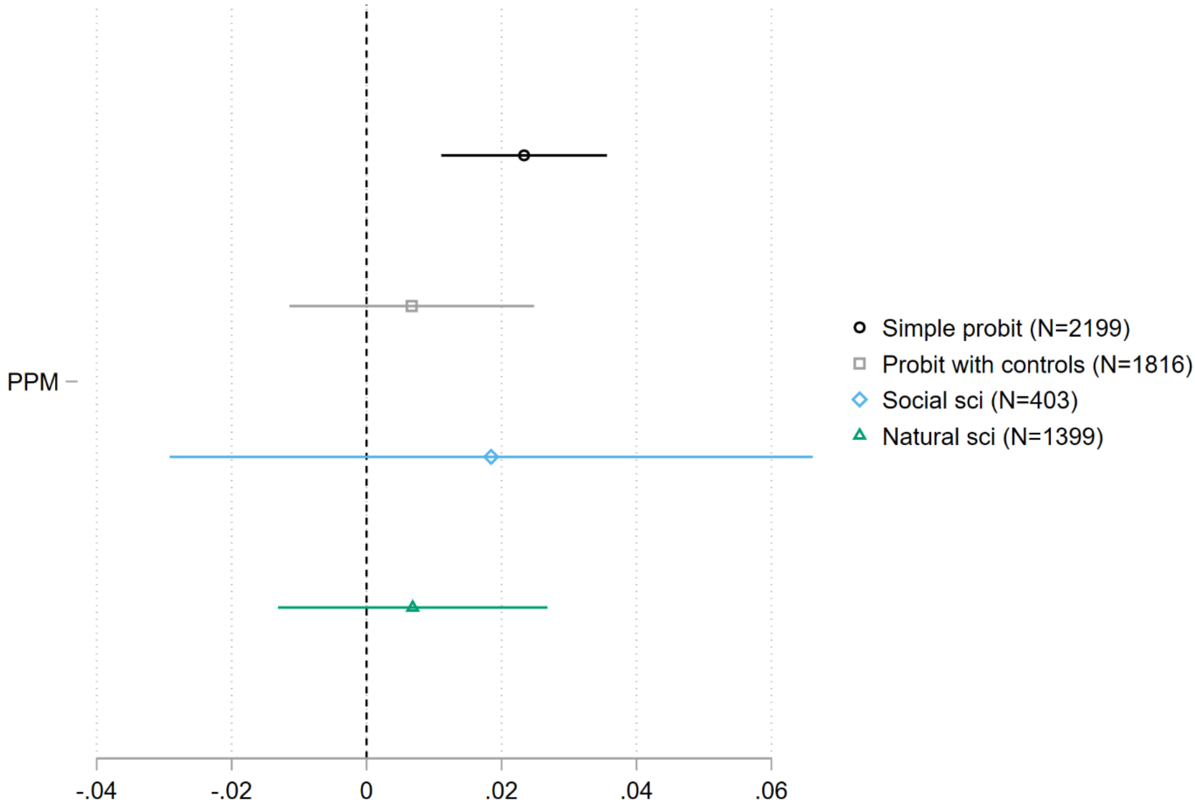

Figure 8: Impact of rail performance on propensity to collaborate again

In figure 9, we test the correlation between travel time between partners and the propensity to collaborate again. While we cannot account for changes in travel time by public transport over time (and assume these are negligible), trave times would already be known to partners prior to the first collaboration. However, partners may learn about the feasibility of collaboration during the course of a collaboration, e.g. whether significant travel time has an impact on their productivity or whether it is a burden on the collaboration. Figure 9 shows no significant association between travel time and the propensity to collaborate again, apart from the sample of social science projects, where there is a negative effect. Thus, a longer travel time is associated with a lower probability of two partners to collaborate again.

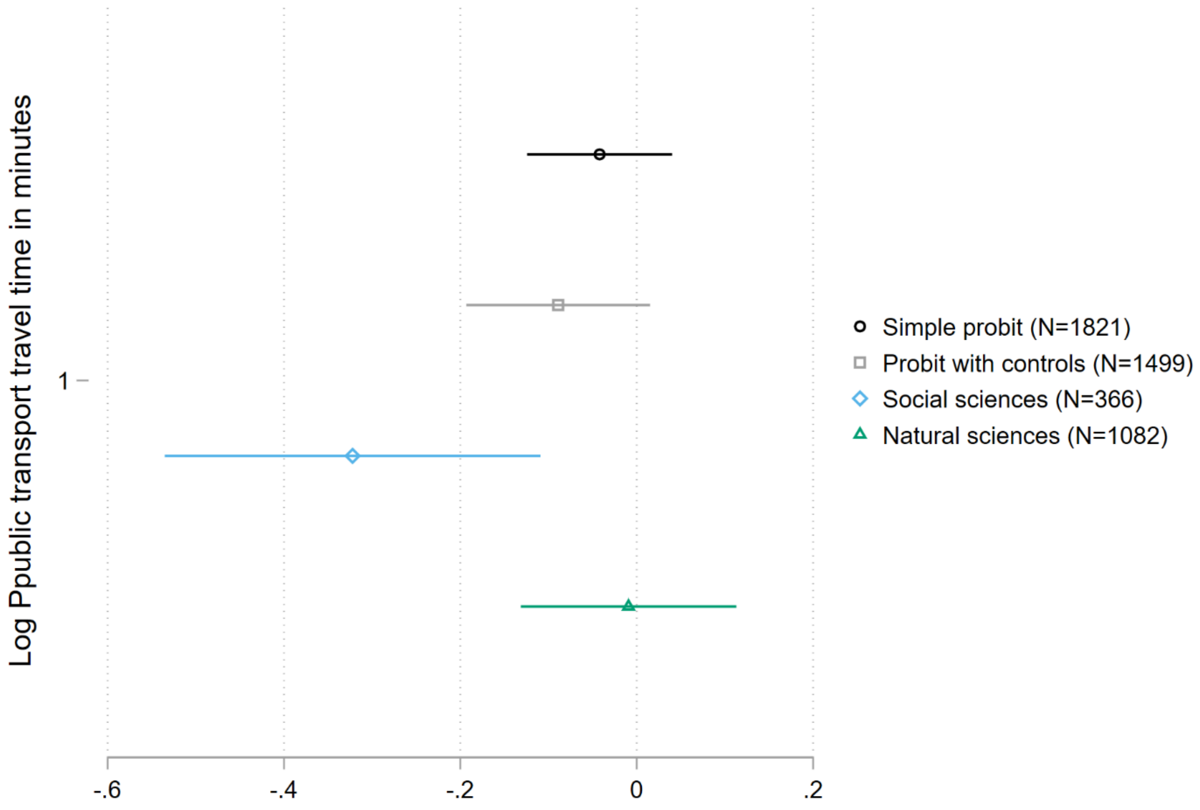

Figure 9: Impact of travel time on propensity to collaborate again

4. Conclusion

This report set out to investigate the relevance of transport connectivity to collaborative research projects. The findings suggest that poor connectivity, particularly unexpected disruptions, has a significant negative impacts on research productivity, as measured by published outputs. These effects are driven by projects involving the worst performing connections. Rail performance and the performance of collaborative research projects are therefore inextricably linked to wider regional inequalities. Public research funding is disproportionately concentrated in the South East, where transport connectivity is better and more investment has been undertaken recently.

Further research can further unpack the selection behaviour of researchers into collaborations. The evidence presented here shows that project size increases with travel distance. Researchers may discount the feasibility of smaller projects over longer distances, suggesting that there may be some fixed costs associated with distance. This would suggest that establishing initial trust or a common understanding of project goals for a successful collaboration is important, which can then carry a project. Alternatively, the inconvenience of further travel may be worth it for more promising projects, suggesting increasing returns to scale. These mechanisms could be tested by considering prior collaboration between teams, for example by accounting for joint prior publications, or by looking at the quality of outputs, not just the quantity.

A narrow focus on publications as the output measure also precludes most innovation projects, such as those funded by Innovate UK from the analysis, as they do not tend to result in publications. A more nuanced measure of outcomes and impacts would strengthen the analysis, for example using text analysis on impact narratives. Innovation projects, particularly those involving collaborations between business and universities struggle with other 'distances', such as cultural and organisational that need to be overcome, making the ability to interact in person even more important (D'Este, Iammarino & Guy, 2013).

References

Bahar, D., Choudhury, P., Kim, D.Y. and Koo, W.W., 2023. Innovation on wings: Nonstop flights and firm innovation in the global context. *Management Science*, *69*(10), pp.6202-6223.

Berkes, E. and Gaetani, R., 2021. The geography of unconventional innovation. *The Economic Journal*, *131*(636), pp.1466-1514.

Boschma, R., 2005. Proximity and innovation: a critical assessment. *Regional Studies*, *39*(1), pp.61-74.

Chai, S. and Freeman, R.B., 2019. Temporary colocation and collaborative discovery: Who confers at conferences. *Strategic Management Journal*, *40*(13), pp.2138-2164.

D'Este, P., Guy, F. and Iammarino, S., 2013. Shaping the formation of university–industry research collaborations: what type of proximity does really matter?. *Journal of economic geography*, *13*(4), pp.537-558.

Forth, T., and Jones, R. (2020). *The missing £4 billion: Making R&D work for the whole UK*. London: Nesta.

Ioramashvili, C. (2024). Technological invention and local labour markets: Evidence from France, Germany and the UK. *Research Policy*, *53*(7), 105021.

Polanyi, M. (1966). *The Tacit Dimension*. London: Routledge.

Sonn, J. W., and Storper, M. (2008). The increasing importance of geographical proximity in knowledge production: an analysis of US patent citations, 1975–1997. *Environment and Planning A*, *40*(5), pp.1020-1039.

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

Van der Wouden, F., and Youn, H. (2023). The impact of geographical distance on learning through collaboration. *Research Policy*, *52*(2), 104698.

Appendix

Robustness checks

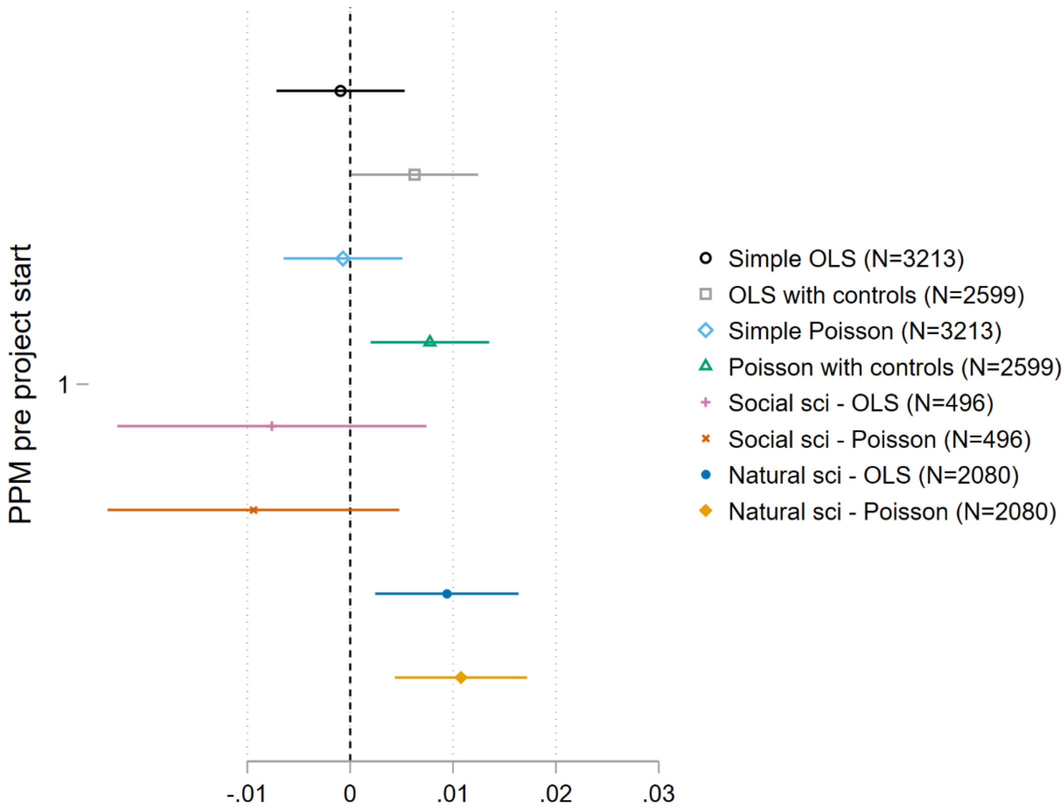

Figure A1: Regression of total publications on PPM prior to project start

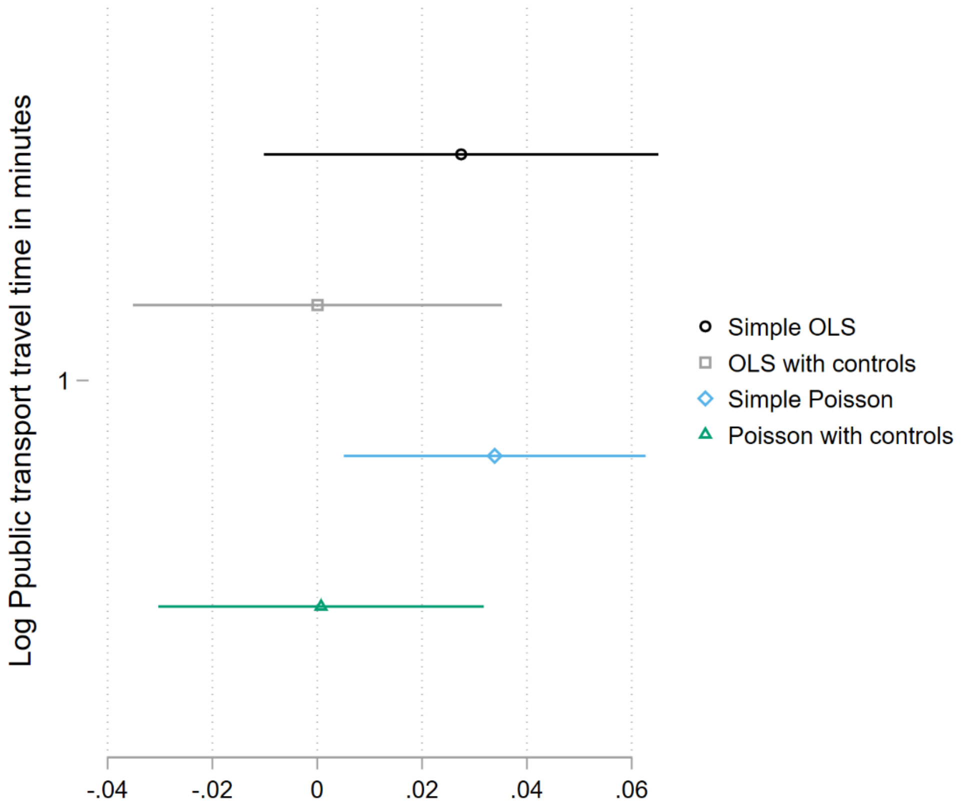

Figure A2: Regression of total publications on travel time between project partners