Intercity innovation collaboration and the role of high-speed rail connections: evidence from Chinese co-patent data

Li Yao\textsuperscript{a}  and Jun Li\textsuperscript{b}

\textbf{ABSTRACT}
This study explores the extent to which changes in transport infrastructure counterbalance pre-existing geographical friction and foster innovation collaboration, using the Chinese high-speed rail (HSR) construction as a quasi-natural experiment. Using a comprehensive dataset of city-pair co-patents from 2005 to 2018, we show that HSR connections significantly increase intercity co-patents, patent quality and collaborative partnerships, and such effects are strongest for city-pairs within 250 km and decrease for longer distances. Moreover, the HSR effect is stronger for cities in similar institutional settings, indicating a negative moderating effect of institutional distance. Various robustness methods are used to confirm the validity of our findings.

\textbf{KEYWORDS}
innovation collaboration; co-patent; high-speed rail; geographical proximity; institutional proximity

\textbf{INTRODUCTION}
Intercity innovation collaborations need to overcome geographical friction. Early research (e.g., Fischer et al., 2006; Maurseth & Verspagen, 2002) that followed Jaffe et al.’s (1993) seminal work found a substantial influence of geographical proximity on innovation collaboration. However, later research that embraced Boschma’s (2005) fivefold classification of proximity as an analytical framework has produced the more balanced findings. First, geographical proximity is found to be less influential than previously assumed once non-geographical forms of proximity are considered in innovation collaboration (Balland et al., 2015; Boschma, 2005; Torre & Rallet, 2005). Second, geographical and non-geographical forms of proximity are often found to be positively correlated (Balland et al., 2015). Nonetheless, geographical proximity is still found to positively influence the formation of innovation collaboration when other forms of proximity are included in studies (Cao et al., 2019; Hong & Su, 2013; Marek et al., 2017).

If geographical distance remains a noticeable barrier to intercity innovation collaboration, to what extent do changes in infrastructure counterbalance pre-existing geographical friction and foster innovation collaboration?

While the limited studies have shown that reductions in communication costs and travel costs as a result of technological advancement mitigate geographical friction (Agrawal & Goldfarb, 2008; Catalini et al., 2020), these studies also point to the differential impact on innovators across urban systems. For example, Agrawal and Goldfarb (2008) examined the adoption of the Internet on university research collaboration in engineering and found reductions in communication costs increased research collaboration between top- and middle-tier institutions from the same region. Catalini et al. (2020) looked at the impact of the introduction of new routes by a low-cost airline on scientist collaboration. They found that reductions in travel costs mitigated geographical friction to collaboration and increased the number of collaborations between 0.3 and 1.1 times. Still, we do not know much about the collaboration-enhancing effect of changes in infrastructure and the cost-induced complementary effect of geographical and institutional proximities.

More recently, the construction of high-speed rail (HSR) in many countries has again raised the question of how HSR can help overcome geographical distance and facilitate intercity innovation collaboration. Currently, the world has 52,418 km of high-speed network in commercial operation and 11,693 km of high-speed lines
under construction, of which China made up no less than 50% (Guigon, 2020). HSR is one of the most advanced modes of ground transportation that could operate at speeds of over 200 km/h. As a key national development strategy in China, the construction of a nationwide HSR network aims to improve connectivity between regions and promote a more balanced and equitable regional development (Chen & Haynes, 2017). The sheer scale of HSR connectivity in China, as can be seen in Figure 1, and the advantages of HSR over other alternatives—such as high speed, convenience, a comfortable experience, proximity to city centres, punctuality and safety—have tremendously transformed the way people travel and become one of the most popular forms of intercity passenger transportation.

So far, the limited research on HSR and innovation has produced some interesting findings. For example, Inoue et al. (2017) used the case of the opening of the Shinkansen in Japan to estimate the impact of HSR on innovative activities along the line. They found that HSR significantly increased patent submissions and patent citations by establishments along the line. Gao and Zheng (2020) used innovation surveys to assess the impact of HSR on innovation in manufacturing firms in the Yangtze River Delta and Pearl River Delta, China’s two most developed regions. The results show that HSR connection promotes firm innovation in peripheral areas. Dong et al. (2019) specifically investigated the impact of HSR on intercity university research collaboration using a dataset of research paper publication and citations. They found that when cities are connected by HSR, co-author productivity from existing collaborations rises, new co-author pairs emerge and more highly productive scientists migrate to the HSR cities. Building on studies with a focus on infrastructure and innovation, we aim to unpack the differential impact of HSR by addressing our first research question: To what extent does HSR affect inter-organizational innovation collaboration across HSR-connected cities?

First, we use dyadic city-pair co-patient panel data to capture more accurately the differential effect of HSR on intercity co-patient collaboration through two mechanisms, that is, HSR-induced intensity of face-to-face interactions, and HSR-induced partner matching in larger labour markets. Second, we disentangle the HSR effect across urban systems and on different types of innovators in the manifestation of inter-firm collaboration (II) and university and research institute collaboration (URI).

Extant research on innovation and proximity has explored the impact of institutional proximity on innovation collaboration. As Balland et al. (2015) noted, however, researchers have used two different definitions of institutional proximity, that is, the degree to which organizations share similar institutional settings at a macro-level (Boschma, 2005), and values and norms in the same sub-system within academia, industry or government, following the triple helix model (Etzkowitz & Leydesdorff, 2000; Ponds et al., 2007). As compared with studies that examined the impact of organization-level institutional proximity using the second definition (e.g., Cao et al., 2019; Hansen, 2015; Ponds et al., 2007), the impact of region-level institutional proximity adopting the first definition is under-researched. So far, only two studies can be found. Hong and Su (2013) analysed the effect of institutional proximity on non-local university–industry collaborations in China. The results show that institutional proximity caused by subordination to the same administrative unit significantly enhances the probability of collaboration, and those effects are more significant when the distance increases, suggesting the substitution effect. Marek et al. (2017) explored the impact of proximity measures on knowledge exchange measured by granted research and development (R&D) collaboration projects in German NUTS-3 regions. They found a ‘U’-shaped impact of institutional distance on interregional collaboration with a negative impact, which shrinks after passing a distinct threshold level. We complement these works to assess the synergetic effect of HSR and institutional distance from a decentralization perspective. China has an economic system characterized by both high centralization and strong decentralization (Bai et al., 2004). Decentralization comes with local economic agendas and local protection (Bai et al., 2004). This decentralized system has a huge impact on how organizations collaborate. Yet, our understanding of how HSR affects intercity innovation collaboration in a decentralized institutional system is limited. Our research intends to fill this void by exploring our second research question: To what extent does institutional proximity moderate the HSR-mitigated relationship between geographical proximity and intercity innovation collaboration?

Our study makes three contributions to the proximity and innovation literature. First, we enrich the literature on geographical proximity and innovation collaboration by unravelling the HSR-mitigated differential effect of geographical friction on the quantity and quality of intercity innovation collaborations. Second, we draw on a regional decentralization perspective to explore how institutional proximity moderates the HSR effect. Thus, we enrich the literature by providing a more nuanced understanding of how different dimensions of proximity interact to affect intercity innovation collaboration. Third, we again follow the decentralization perspective to explore whether the HSR effect is more constrained in geographical distance for II collaboration and more constrained in institutional distance for URI collaboration. Our research, therefore, sheds new light on how different types of innovation actors with varying inherent incentives and behaviours may respond differently to the improved transportation infrastructure and manifest themselves differently in intercity innovation collaboration.

THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

For innovation collaboration, face-to-face communication is of importance throughout all stages of the process. At the initial stage, it is essential for firms to match up with the right types of partners—such as vertical (suppliers),
horizontal (competitors) or institutional (universities and research institutes) partners – that complement themselves with the necessary knowledge and resources to achieve the goals pursued. Once a partnership or alliance is formed, it is necessary for team members to become familiar with each other, to develop an enhanced understanding of the problem-solving procedure, to cultivate personal trust, and eventually to build effective research routines so as to improve efficiency and prospect for project success (Bercovitz & Feldman, 2011). Therefore, close face-to-face contact is essential throughout the whole process of innovation collaboration.

HSR, as an advanced transportation infrastructure, provides a faster, safer, more comfortable and arguably the most punctual transportation service than other alternatives, such as regular train, automobile and flights (Sun et al., 2017). It fills a blank of travelling at a distance too far for cars and too close for flights. Hence, HSR facilitates more cost-effective intercity travel for face-to-face contact and collaborative innovation. Overall, HSR affects intercity innovation collaboration through two mechanisms. First, HSR helps increase the intensity of interaction between collaborative partners between connected cities. HSR connections significantly shrink the geographical distance between cities because of high travel speed at relatively lower costs. This cost-effective transport mode allows team members in inter-organizational collaborative projects to interact face to face more in order to build rapport, share tacit knowledge and resolve differences. The increasing intensity of interaction has two implications. First, it enhances the efficiency of innovation collaboration that leads to a better performance of co-patenting in quantitative terms. Such an effect should be stronger for city-pairs that within HSR travel time. Studies have shown that within 600 km, the HSR travel experience dominates the alternatives, including that of air travel (Lawrence et al., 2019). Therefore, the HSR effect on intercity innovation collaboration should be more salient for city-pairs that are geographically close and within the HSR range. Second, more face-to-face interactions between innovation partners across HSR-connected cities also improve collaborators’ cognitive proximity due to greater knowledge-sharing, hence elevating innovation outcomes, such as patent quality. Dong et al.’s (2019) research on academic co-publication found the positive effect of HSR on both quantity and quality of co-publication. Therefore, we propose the following.

Hypothesis 1a: HSR increases the quantity of collaborative innovation between connected city-pairs.

Hypothesis 1b: HSR connection improves quality of collaborative innovation between connected city-pairs.
Second, HSR increases opportunities for partner matching in larger markets. First, because HSR increases travel speed across cities, it thus creates a larger market for organizations to match partners for innovation projects. The increased cross-city travel speed at lower costs facilitates the better matching of researchers with complementary skills in a larger scientist labour market, leading to the formation of more innovation collaborative partnerships between the connected city-pairs. Second, by establishing connections between core markets (megacities) and outside markets (smaller cities), transport improvements expand the geographical reach of knowledge spillover, thereby enabling organizations in outside markets to do new things or to accomplish old tasks in new ways and energize innovation in other sectors. Accordingly, HSR connections again lead to the formation of more innovative collaborative partnerships between the connected city-pairs. Hence, we propose the following:

**Hypothesis 1c:** HSR increases the number of innovation collaborative partnerships between connected city-pairs.

In addition to geographical distance, intercity innovation collaboration can be constrained by the intercity discrepancy in formal and informal institutions (North, 1990), or institutional distance. A distinct feature of the economic system in China is administrative decentralization (Perkins, 1988). The devolution of decision-making power from the centre to local governments allows locally available information to be used more effectively and local preferences to have greater influence over local spending decisions (Chen, 1998). Inevitably, the system leads to the diversity of regulative institutions in terms of rules, laws and sanctions through the coercive mechanism (Scott, 2013). The unintended consequences of such a decentralized system are local governments’ zeal for gross domestic product (GDP) growth and local protectionism (Bai et al., 2004). Local governments prefer to support local business development and local inter-organizational collaboration as government officers are more likely to get promotions if local economic growth can benefit from significant innovations within their territories (Hong & Su, 2013). Hence, institutional diversity between regions gives rise to unpredictable and unreliable conditions under which effective inter-organizational innovation collaborations are more difficult to take place. Regional institutional discrepancies and protectionism impede cross-region innovation collaboration (Ding & Li, 2015).

Despite lower travel costs incurred by HSR connections, institutional differences persist among provinces. Therefore, the HSR effect is moderated by institutional distance, that is, HSR connections increase more innovation collaboration between city-pairs that are within the same province than those across provinces. For the aforementioned reasons, we propose the following hypothesis:

**Hypothesis 2:** Institutional distance moderates the positive effect of HSR connection on innovation collaboration between connected city-pairs.

Patents as an embodiment of commercializable technologies are created and applied by various innovative actors. Firms, universities and research institutions are three key innovative actors in the national innovation system (NIS). It is important to distinguish two types of inter-organizational collaborations, that is, collaborations involving universities and research institutes (URI)\(^1\) and intra-industrial collaborations (II) because different knowledge types are affected in different dimensions of proximity (Davids & Frenken, 2018). Since most universities and research institutes are public-funded institutions in China, intercity URI collaborations are driven by the government’s social and economic policies and aim to produce public goods. In contrast, intercity II collaborations are motivated by business interest and aim to deliver private goods. Comparatively, URI collaborations led by universities and research institutes are influenced by government agendas in the Chinese context, and therefore are less sensitive to economic costs but are more likely to succumb to the carrot-and-stick approach used by local governments. For such reasoning, we propose the following hypotheses:

**Hypothesis 3a:** The HSR effect is stronger within close distance for II collaborations than URI collaborations.

**Hypothesis 3b:** The HSR effect is stronger within the same province for URI collaborations than II collaborations.

## DATA

### Data and statistics

Co-patent data from the China National Intellectual Property Administration (CNIPA) are used in this study. We collected data of all patents with two or three applicants over the period 2005–18 from the database incopat.com.

All co-patent data include two or three applicants who could be individuals, firms, universities or research institutes. We construct intercity co-patents by using the information of applicant addresses. For a patent with applicants from city A and city B, we construct the city-pair as A–B, while for three-applicant from cities A, B and C, we then create three city-pairs, A–B, A–C and B–C, but each with one-third of the weight. There are 293 prefecture cities and four municipalities in China in 2018. We only consider the cities that had intercity co-patents and exclude those without, which leaves a total of 285 prefecture-level and above cities. The original dataset contains 1 million co-patents, and after data cleaning it reduced to 708,022, of which 386,735 are intercity co-patents. We exclude intracity collaborations in our research sample.

We use three dependent variables. The first is CoPat\(_{ijt}\), the count of co-patents between city i and city j during year t as the measure of the quantity of collaborative innovation. The second is the value-weighted patent CoPatW\(_{ijt}\) as a measure of the quality of collaborative innovation. The third dependent variable is CoPatP\(_{ijt}\) (collaborative partners) as a measure of the quantity of innovation

\(^{1}\) URI: University–Research Institute.
collaborative partnership. Once two organizations cooperate for at least one co-patent in one year, they are defined as innovation collaborative partners.

A city is connected to the HSR network once at least one station is opened. The opening dates and routes of HSR are collected from the official website 12306.cn, maintained by the National Railway Administration of China. The dummy variable $\text{Connect}_{ijt}$ is coded 1 if both cities $i$ and $j$ are HSR connected in year $t$, and 0 otherwise.

We construct a city-pair panel dataset including all cities with at least one co-patent over 14 years from 2005 to 2018. We exclude the city-pairs that never had any collaborations, and only keep the city-pairs that had at least one collaboration over 14 years. There are a total of 79,058 observations over 14 years on 5647 unique city-pairs. The main independent variables of interest in addition to $\text{Connect}_{ijt}$ are $\text{Distance}_{ij}$ and $\text{SameProv}_{ij}$, which indicate the distance between cities $i$ and $j$, and whether the two cities belong to the same province. Geographical distances are measured in straight-line (or Euclidean) distance between cities using geographical information. The literature has used both straight-line distance and travel time to measure geographical distance, but they generate very similar results (Marek et al., 2017). Most of the gravity-based models use straight-line distance for its lower cost. Following Hong and Su (2013), we use the provincial-border definition of institutional distance. It implies that cities belonging to the same province are considered institutionally approximate and incur no institutional friction because they are subject to the same regulative institution.

For all the specifications, we control city and year fixed effects, which tease out time-invariant city-level characteristics and time-trending effects. Besides, we control other time-variant confounding variables that possibly affect the intercity co-patents, including city-level GDP, total single-applicant patents, science and technology government expenditure, etc. Table 1 summarizes the definition and sources of the dependent, independent and control variables.

### Table 1. Definitions of variables.

| Variable                | Definition                                                                 |
|-------------------------|---------------------------------------------------------------------------|
| **Outcome variables**   |                                                                           |
| $\text{CoPat}_{ijt}$   | Total co-patent count between city $i$ and city $j$                      |
| $\text{CoPatW}_{ijt}$   | Co-patents weighted by patent values. Patent value is calculated by incopat.com using big data techniques and including information such as patent citation, assignment, licensing, legal status and other determinant variables |
| $\text{CoPatP}_{ijt}$   | Co-patent partnerships between city $i$ and city $j$ in year $t$. Calculated using unique collaborative pairs from the variable $\text{CoPat}_{ij}$ |
| $\text{CoPat(URI)}_{ijt}$ | Co-patents involving universities and research institutes between city $i$ and city $j$ in year $t$. $\text{CoPat(URI)}_{ijt}$ is URI partnerships |
| $\text{CoPat(I)}_{ijt}$ | Intra-firm co-patents between different firms between city $i$ and city $j$ in year $t$. $\text{CoPat(I)}_{ijt}$ is II partnerships |
| **Independent variables** |                                                                           |
| $\text{Connect}_{ijt}$ | Time-variant dummy variable indicating whether city $i$ and city $j$ are connected to high-speed rail (HSR) in year $t$ with a one year lag. The data are collected from www.12306.cn, which is maintained by the National Railway Administration of China |
| $\text{Distance}_{ij}$ | Time-invariant variable measures the straight-line (Euclidean) distance between city $i$ and city $j$ (km) |
| $\text{Dist.S}_{(0-250km)}$ | Time-invariant dummy variable coded 1 when the city-pair is within 250 km. $\text{Dist.M}_{(250-600km)}$ and $\text{Dist.L}_{(600-1000km)}$ are for median and long-range dummy variables, respectively |
| $\text{SameProv}_{ij}$ | Time-invariant dummy variable indicating whether city $i$ and city $j$ are located in the same province |
| **Control variables**   |                                                                           |
| $\text{SinglePat}_{it}$ | Total single-applicant patents for city $i$ in year $t$                   |
| $\text{GDP}_{it}$       | Total gross production (RMB millions)                                   |
| $\text{SciExp}_{it}$    | Government expenditure in science and technology (RMB millions)          |
| $\text{HwayRidership}_{it}$ | Total highway ridership (thousands)                                     |
| $\text{MobileUsers}_{it}$ | Total mobile phone users (thousands)                                    |
| $\text{PatentPerFirm}_{it}$ | Total patents per industrial firm                                       |
| $\text{STPerFirm}_{it}$ | Expenditure in science and technology per industrial firm (RMB millions) |
| $\text{RDPPerFirm}_{it}$ | Research and development (R&D) employment per industrial firm (thousands) |
| $\text{FIEs}_{it}$      | Total foreign invested firms                                             |
| $\text{SecondaryShare}_{it}$ | Secondary industry output share                                         |
| $\text{TertiaryShare}_{it}$ | Tertiary industry output share                                          |
Table 2 reports the descriptive statistics of the main dependent and independent variables. As can be seen, the distribution of CoPat is quite skewed with a mean of 4.13 and a standard deviation of 37.02, and a large number of those observations are zeros. In light of considerable over-dispersion and a larger number of zero observations, negative binomial regressions are used throughout the analyses. Table 2 also reports other outcome variables including weighted co-patents, partnerships, and also URI and II co-patents and partnerships.

The correlation analysis in Table 2 suggests the key independent variables such as Connect, SameProv and Distance all have relatively low and moderate correlation coefficients with other variables. Moreover, to test multicollinearity, we inspected the variance inflation factors (VIFs) of the variables using linear regression (the convention is to use linear regression when the main regression model is non-linear). The VIFs for Connect, SameProv and Distance are 5.1, 1.66 and 1.68, respectively, which are below the acceptable level of 10 (Neter et al., 1996). Hence, multicollinearity is not a concern in our case.

**EMPIRICAL STRATEGY AND RESULTS**

We apply a difference-in-differences (DID) approach to test the effects of HSR on intercity innovation collaboration. Because intercity innovation collaborations are dyadic in nature, we apply a variant of gravity model—an empirical method originally used in international trade literature, and then increasingly adopted in invention collaboration and knowledge flow literature (Cappelli & Montobbio, 2016; Picci, 2010). Our regional gravity model considers innovation collaboration in three measures as a function of the distance between cities and time-variant economic variables that potentially affect intercity innovation collaboration. We formulate the following baseline regression model:

\[
Y_{ijt} = \alpha_0 + \alpha_1 \text{Connect}_{ij} + \alpha_2 \text{Dist}_{ij} + \alpha_3 \text{SameProv}_{ij} + X \beta + \delta_i + \delta_j + \tau_t + \epsilon_{ijt} \tag{1}
\]

The main variable of interest is Connect$_{ij}$, the one-year lagged dummy variable indicating whether city $i$ and city $j$ are both connected to the HSR network in the year $t-1$. Dist$_{ij}$ is a continuous variable of the straight-line distance (km) between city $i$ and city $j$. SameProv$_{ij}$, a dummy variable, indicates whether two cities are located in the same province, which is interpreted as institutional proximity. $X$ stands for the time-variant control variables for city $i$ and city $j$. $\delta_i$ and $\delta_j$ indicates the fixed-effects of city $i$ and city $j$, respectively. $\tau_t$ is the year dummy. $\epsilon_{ijt}$ is the error term.

The data and method we use have two advantages. First, we use a full set of Chinese co-patent data across all industries, so our data are representative of all ranges of technologies. Second, dyadic city-pair observations combined with the gravity model could better identify the effects of geographical and institutional distance. In
all estimations, we apply the negative binomial gravity model framework.

Baseline result
Table 3 presents the baseline results from three model specifications. Model 1 excludes city-level variables and fixed effects. From pooled regression in model 1, HSR connections show a positive and significant effect, meaning connected city-pairs tend to have a larger number of co-patents than those unconnected.

In model 2, we additionally control for the interaction term of distance ranges with HSR connection: Connect × (Dist.S + Dist.M + Dist.L). Dist.S is coded as 1 if the city-pair is within 250 km, about 1.5 h travel time by HSR. Similarly, Dist.M and Dist.L indicate city-pairs of distance ranging from 250 to 600 km and from 600 to 1000 km. The interaction coefficients of connect with the range dummy variables suggest the extent to which HSR connection increases the number of co-patents between city-pairs within those ranges. The coefficient of Connect × Dist.S(0–250km) is significant and positive at 0.261, meaning that HSR is effective in increasing innovation collaborations between connected city-pairs within 250 km by 26.1%. From 250 to 600 km, though insignificant, HSR connection accounts for a 13.1% increase in co-patents, while from 600 to 1000 km, the effect is negative. Therefore, the results of model 2 support Hypothesis 1a, indicating that HSR connection increases intercity innovation collaborations. The HSR effect is most significant within 250 km, but diminishes by longer distances.

Model 3 additionally controls for the interaction term: Connect × (Distance + SameProv), which is to examine whether being in the same province, conditional on geographical distance, moderates the effect of HSR connection. The interaction shows a positive and significant effect at 0.301, suggesting that HSR connection increases by 30.1% for cities that are within the same province after controlling for geographical distance. That is, institutional proximity complements HSR connection in facilitating intercity co-patents. In other words, the HSR effect is moderated by institutional distance. Hence, it supports Hypothesis 2.

Innovation quality and partnerships
To test the effect of HSR connection on quality of co-patent value and the number of collaborative partnerships, we use the same regression specifications as in models 2 and 3 from Table 3, but with different dependent variables.

Table 3. Baseline regression of negative binomial regressions.

|                  | (1)           | (2)           | (3)           |
|------------------|---------------|---------------|---------------|
| Connect          | 1.972***      | 0.009         | 0.042         |
|                  | (0.059)       | (0.056)       | (0.071)       |
| Distance         | −0.001***     | −0.001***     | −0.001***     |
|                  | (0.0001)      | (0.0001)      | (0.00004)     |
| SameProv         | −0.121        | 1.312***      | 1.431***      |
|                  | (0.087)       | (0.057)       | (0.050)       |
| Dist.S(0–250km)  |               | 0.344***      |               |
|                  |               | (0.097)       |               |
| Dist.M(250–600km)| −0.004        |               |               |
|                  | (0.078)       |               |               |
| Dist.L(600–1000km)| −0.021       |               |               |
|                  | (0.060)       |               |               |
| Connect × Dist.S(0–250km)| 0.261*** |               |               |
|                  | (0.075)       |               |               |
| Connect × Dist.M(250–600km)| 0.131*   |               |               |
|                  | (0.073)       |               |               |
| Connect × Dist.L(600–1000km)| −0.051 |               |               |
|                  | (0.069)       |               |               |
| Connect × Distance |                   | −0.00002     |               |
|                  |               | (0.00005)                       |
| Connect × SameProv |               | 0.301***     |               |
|                  |               | (0.077)       |               |
| CityPairgtControls | No            | Yes           | Yes           |
| CityYeargtFE     | No            | Yes           | Yes           |
| Observations     | 79,058        | 78,834        | 78,834        |
| Akaike information criterion (AIC) | 201,818.700 | 170,575.300  | 170,666.300  |

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Robust standard errors are clustered at the city level.
In Table 4, columns (1) and (3) report the interaction effect within different distance ranges, and columns (2) and (4) report the interaction with institutional proximity. For regressions on value-weighted co-patents, column (1) suggests that HSR connect increases value-weighted co-patents by 28.5% within 250 km, greater than 26.1% in the simple patent count in the baseline estimation. The difference suggests that the increased innovation collaborations as a result of HSR connection also tend to produce higher innovation value, hence confirming Hypothesis 1b.

Column (2) shows that HSR connection increases value-weighted co-patents by 30.4%, which is similar to 30.1% in the baseline estimate, suggesting that co-patent quality increases for collaboration between cities of geographical proximity but not institutional proximity.

Regressions on co-patent partnerships in columns (3) and (4) reveal slightly differential HSR effects on the dimension of geographical and institutional proximity. On one hand, there is a weak significant HSR effect on forming collaborative partners for cities of geographical proximity, indicating that most of the increased co-patents are from the deepening of existing collaborative partnerships. On the other, HSR connection increases collaborative partners for cities of institutional proximity (being in the same province). Hence, there are both deepening and widening effects on collaboration between cities of the same province. A further analysis, which breaks down the types of innovators, reveals a more nuanced pattern of how HSR affects intercity collaborative partnerships where Hypothesis 1c can be partially supported.

Comparison between URI and II collaborations
Table 5 presents the regression results that explore the difference between URI and II co-patents. Columns (1) and (5) are regressions on co-patent counts, which shows that HSR connection increases II co-patents by 27.6% and URI co-patents by 18.9%, suggesting the greater effect of HSR connection on II than URI collaboration within 250 km. Between 250 and 600 km, the coefficient is also greater for II than URI collaborations, but both effects are insignificant. Columns (3) and (7) are regressions on collaborative partnerships, which show that HSR connection has a positive and significant effect on the formation of innovation partnerships for II but not for URI collaborations. Together, the results confirm our Hypothesis 3a, indicating that HSR connection has a greater effect on II than URI collaborations so far as geographical dimension is concerned.

For columns (2), (4), (6) and (8), the regressions explore the effect of institutional dimension on HSR connection and both II and URI collaborations. The coefficients on Connect × SameProv suggest that HSR connection increases within provincial innovation collaborations more for URI and II types. For II co-patents, the HSR effect on co-patent count is weakly significant at 0.198 and the effect on the formation of partnerships

| Table 4. Regressions on value weighted co-patents and partnerships. | Value-weighted | Partnerships |
|---------------------------------|----------------|-------------|
| Dependent variable              | (1)            | (2)         | (3)          | (4)          |
| Connect                        | 0.009          | 0.063       | −0.007       | −0.102**     |
|                                | (0.083)        | (0.097)     | (0.048)      | (0.043)      |
| Distance                       | −0.001***      | −0.001***   | −0.001***    | −0.001***    |
|                                | (0.0001)       | (0.0001)    | (0.0001)     | (0.0001)     |
| SameProv                       | 1.415***       | 1.543***    | 1.264***     | 1.487***     |
|                                | (0.134)        | (0.122)     | (0.104)      | (0.087)      |
| Connect × Dist.S(0−250km)       | 0.285***       | 0.102*      |             |             |
|                                | (0.105)        | (0.061)     |             |             |
| Connect × Dist.M(250−600km)     | 0.164          | −0.049      |             |             |
|                                | (0.125)        | (0.056)     |             |             |
| Connect × Dist.L(600−1000km)    | −0.058         | −0.031      |             |             |
|                                | (0.146)        | (0.064)     |             |             |
| Connect × Distance             | −0.00003       |             | 0.0001      |             |
|                                | (0.0001)       |             | (0.00005)   |             |
| Connect × SameProv             | 0.304***       | 0.223***    |
|                                | (0.114)        | (0.064)     |
| Control Variables              | Yes            | Yes         | Yes         | Yes         |
| City, Year FE                  | Yes            | Yes         | Yes         | Yes         |
| Observations                   | 78,834         | 78,834      | 78,834      | 78,834      |
| Akaike information criterion (AIC) | 156,252.300   | 156,339.900 | 138,473.600 | 138,896.600 |

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Standard errors are clustered at the city level. The odd-numbered columns follow model 3, and the even-numbered columns follows model 4 in the baseline results of Table 3. Some non-essential variables are omitted.
Table 5. Regression comparison between university and research institute collaboration (URI) and inter-firm collaboration (II) co-patents.

|                  | II co-patents | II partnerships | URI co-patents | URI partnerships |
|------------------|---------------|----------------|----------------|-----------------|
|                  | (1)           | (2)            | (3)            | (4)             | (5)           | (6)            | (7)            | (8)             |
| **Dependent variable: Co-patent count and partnerships** |               |                |                |                 |               |                |                |                 |
| Connect          | -0.043        | 0.039          | -0.029         | -0.067          | 0.042         | -0.157*       | 0.010          | -0.257***       |
|                  | (0.094)       | (0.102)        | (0.074)        | (0.073)         | (0.047)       | (0.083)        | (0.047)        | (0.063)         |
| Distance         | -0.001***     | -0.001***      | -0.001***      | -0.001***       | -0.0005***    | -0.001***     | -0.0005***     | -0.001***       |
|                  | (0.0001)      | (0.0001)       | (0.0001)       | (0.0001)        | (0.0001)      | (0.0001)       | (0.0001)       | (0.0001)        |
| SameProv         | 1.201***      | 1.272***       | 0.993***       | 1.156***        | 1.640***      | 1.870***      | 1.629***       | 1.863***        |
|                  | (0.138)       | (0.101)        | (0.115)        | (0.085)         | (0.096)       | (0.110)        | (0.096)        | (0.100)         |
| Connect × Dist, S (0–250 km) | 0.276**      | 0.165*         | 0.189**        | -0.013          |
|                  | (0.122)       | (0.094)        | (0.086)        | (0.086)         |
| Connect × Dist, M (250–600 km) | 0.149        | -0.002         | -0.053         | -0.134**        |
|                  | (0.135)       | (0.076)        | (0.067)        | (0.067)         |
| Connect × Dist, L (600–1000 km) | -0.013       | 0.033          | -0.165***      | -0.136**        |
|                  | (0.150)       | (0.090)        | (0.053)        | (0.053)         |
| Connect × Distance | -0.00004    | 0.00004        | 0.0001*        | 0.0002***       |
|                  | (0.0001)      | (0.0001)       | (0.0001)       | (0.0001)        |
| Connect × SameProv | 0.198        | 0.209***       | 0.435***       | 0.311***        |
|                  | (0.147)       | (0.074)        | (0.134)        | (0.083)         |
| City gt Controls | Yes           | Yes            | Yes            | Yes             |
| City, Year gt FE | Yes           | Yes            | Yes            | Yes             |
| Observations     | 79,058        | 79,058         | 79,058         | 79,058          |
| Akaike criterion (AIC) | 119,984.300  | 120,002.100  | 94,778.950     | 94,926.000      |
|                  | 105,502.600   | 105,744.000   | 93,477.190     | 93,894.030      |

Note: *p < 0.1; **p < 0.05; ***p < 0.01. Standard errors are clustered at the city level. Some non-essential variables are omitted.
is 0.209. Both numbers are noticeably smaller than the coefficient for URI at 0.435 and 0.311. Thus, comparison of HSR effect on II and URI collaborations again confirms Hypothesis 3b.

**ROBUSTNESS CHECK**

**Test of parallel trend assumption**

The validity of DID method hinges on two key assumptions, namely the parallel trend assumption and exogeneity of HSR connections. We graphically illustrate the trends of co-patents between different types of city-pairs. Figure 2 illustrates the trend comparison of different city-pairs. The left panel compares the trend of connected and non-connected city-pairs, and the right panel illustrates the trends of connected city-pairs of different distance ranges. The pre-treatment lines followed a relatively close trend with each other. However, when inspecting more closely, there is a slight divergence visible at year $t - 1$, which also shows up in Figure 3 that there is a slight jump in year $t - 1$. Given that all variables are random, a slight divergence is acceptable. Nonetheless, we argue that the slight divergence in parallel trends is likely the result of the 'step ahead' effect. That is, organizations may move ahead to carry out more collaborative activities in anticipation of HSR connection. As R&D and innovations are long-term projects, in the foresight of better connectivity and more convenient travel, firms may move ahead to increase collaboration with partners that are more accessible in the near future. We argue that the 'step-ahead' effect differs from a secular divergent effect and does not carry into the following periods.

To test whether the slight divergence is within statistical tolerance, we construct an event study to investigate the dynamics of HSR effects in specific years before and after connection. Figure 3 illustrates those coefficient estimates for individual years. On the left panel, the blue dots indicate the estimated coefficients of HSR connection on co-patents for city-pairs within 250 km. The three years before connection had a weak effect, while the effect of post-connection started to pick up and peaked in year $t + 2$, two years after connection. The right panel shows the effect on collaborative partnerships. Similar to the effect on co-patents, the pre-connection years showed little effect, while intercity collaborative partnerships started to establish from $t + 0$ to $t + 3$. The results of the event study again confirm that the parallel trend assumption holds and that the increased intercity co-patents and partnerships occur likely as a result of the construction of the HSR network.

Furthermore, we conduct a placebo test to check whether the identified HSR effect could be contaminated by the pre-treatment trend. The results (see the Appendix in the supplemental data online) confirm that the parallel trend assumption required in the DID method is satisfied.

**Test of endogeneity**

Another potential identification issue is endogeneity resulting from reverse causality. If planning of HSR routes were to be based on the expectation that some cities tend to engage more in collaborative activities, then HSR connection would be endogenous due to reverse causality. Following Dong et al. (2019), we use historical rail connection and city-level elevation as instrumental variables (IVs) and adopt a two-stage least square (2SLS) approach to test endogeneity. The results (see the Appendix in the supplemental data online) confirm that findings from our DID analysis are robust. Taken together, the results of trends comparison, event study and IV regression all confirm that our DID analysis satisfies its key assumptions and that our findings are robust.

**Marketization index (MI) as a measure of institutions**

In this study, we follow Hong and Su (2013) and use a province-border measure of institutional distance. We find evidence to confirm that institutional distance moderates the positive effect of HSR connection on innovation collaboration between connected city-pairs. To test the robustness of our results, we use the MI as an alternative

---

**Figure 2.** Trends comparisons before and after the high-speed rail (HSR) openings. Note: Since HSR connections took place in different years over time, we align the opening year at time $t$. The dashed line indicates the event of an HSR connection dividing the pre- and post-periods. The left panel compares the connected and unconnected city-pairs. The right panel compares the connected city-pairs of different distances in between.
In this paper we investigate the effect of HSR on intercity innovation collaboration. Using the massive construction of the HSR network in China as a quasi-natural experiment with a variety of econometric approaches, we find that HSR connections increase collaboration quantity (co-patents) and quality (value-weighted co-patents) by 26.1% and 28.6%, respectively, for city-pairs within 250 km. The HSR effect diminishes after 250 km and disappears beyond 600 km. Institutional proximity positively moderates the HSR effect, that is, the HSR effect for city-pairs of the same provinces is stronger than for those across different provinces. Further analyses show that the HSR effect on II co-patents is greater than URI co-patents within 250 km, while the HSR effect is stronger for URI than II co-patents within the same province. Our research contributes to the literature on collaboration in several ways.

First, our research provides a more nuanced understanding of the HSR effect on intercity innovation collaboration. Our research investigates inter-organizational collaborations and finds a similar HSR effect on the quantity and quality of co-patenting between connected city-pairs, thus enriching the evidence base of the HSR and innovation collaboration literature. Moreover, extant research is inconsistent on whether HSR has the ‘polarized-effect’ or ‘levelling-up effect’. For example, some studies found that HSR enhances the economy of core cities or large cities at the expense of smaller cities (e.g., Ke et al., 2017; Monzón et al., 2013; Vickerman, 2018). On the contrary, some studies found that HSR creates new locational advantages for small cities (e.g., Chen & Haynes, 2017; Sasaki et al., 1997). Dong et al.’s (2019) research on academic co-publication found that HSR increases co-authors’ productivity and cooperation among authors from central and secondary cities. The evidence from our research on inter-organizational collaboration does not dismiss the ‘polarized effect’ of HSR as we find the significant HSR effect on innovation collaboration is only found for large–large city-pairs within 250 km. However, our findings lean more towards the ‘levelling-up’ effect as we find that the HSR effect is more pronounced on co-patenting between major–small

**Heterogenous effect on cities of different sizes**

To test the heterogeneity of treatment effect among cities of different sizes, we further conduct a subsample analysis on cities of different sizes. We split Chinese cities into 49 major and 236 small cities. To disentangle the effect of HSR, we separate the observations into three types: major–major, major–small and small–small. We run subsample regressions of the three types of city-pairs (detailed in the Appendix in the supplemental data online). Table A4 online reports the HSR effects on city-pairs of different size combinations. The analysis reveals three additional intriguing results. First, the most significant HSR effect comes from between major–small city-pairs. HSR connection increases the number of co-patents by 22.6% within 250 km, 13.9% within 600 km and 33% within the same province for the major–small city-pairs. Second, for large–large city-pairs, the significant HSR effect (20.8% increase) is only found within 250 km, and the effect is more pronounced (29.4% increase) within the same province. Third, HSR connection has insignificant effect on innovation collaborations in small–small city-pairs.

**DISCUSSION AND CONCLUSIONS**

In this paper we investigate the effect of HSR on intercity innovation collaboration. Using the massive construction of the HSR network in China as a quasi-natural measure of institutional distance. The MI is a frequently used measure of institutional quality in China (Li et al., 2006). It measures the extent to which Chinese provinces progress toward a fully fledged market economy under economic reform (Bin et al., 2020). Using MI as an alternative measure of institutional distance, we find consistent evidence (see the Appendix in the supplemental data online) with one that uses province-border as a measure of institutional distance. The additional analysis thus confirms that our main thesis regarding the effect of institutional distance remains robust.

**Figure 3.** Dynamics of high-speed rail (HSR) effects: co-patents and partnerships.

Note: We run the same fixed-effects regression model with indicator variables corresponding to three years before and six years after the HSR connection. The left panel reports the regression coefficients of co-patents. The right panel reports the coefficients of the partnerships. Vertical bars represent the 90% confidence internal.
city-pairs up to the geographical distance of 600 km. While our research observes the tendency of innovation collaboration between large cities, we nonetheless find that HSR networks help reconfigure the NIS by expanding the system’s reach to smaller cities.

Second, our research contributes to the understanding of the complementary effect between geographical distance and institutional distance on innovation collaboration. In the limited literature that adopts the definition of institutional proximity at the macro-level, Hong and Su’s (2013) research implies a substitution effect between geographical proximity and institutional proximity. Marek et al.’s (2017) research a ‘U’-shaped effect of institutional distance on interregional innovation collaboration. We extend this line of research to account for changes in the infrastructure on collaboration in the case of HSR. We particularly embed our arguments in the decentralization perspective. We find that institutional friction remains persistent even after cities are connected by HSR. As a result, the positive effect of lower travel costs is weakened by persistent institutional friction, suggesting that HSR benefits are moderated among cities across different provinces. Our research thus implies that the impact of HSR on intercity innovation collaboration would be greater if local protectionism in the decentralization system were to be weakened.

Third, we enrich the literature of transportation infrastructure and intercity collaboration with empirical evidence that shows that HSR connections affect URI and II collaborations differently.

The findings of this study have practical implications for policymakers. Recent studies suggest intercity collaboration linkages improve cities’ innovation capacity (Cao et al., 2021; Yao et al., 2020). Hence, to best capitalize on the effect of HSR, policymakers need to pursue a synergetic development plan of transportation and intercity collaboration network. Local governments could work together and toward eliminating the persistent presence of institutional friction that weakens the benefits of HSR connection, especially on URI innovation collaboration, freeing up the flow of knowledge and innovative resource to form a more integrated NIS. As HSR connection presents an opportunity for innovators in small cities to form collaborative projects with the major cities where innovative resources are more abundant, the local policymakers should create accommodation and encourage such collaboration with the major cities.

This study has limitations. First, our research only captures the partial effect of HSR connection on innovation collaboration because collaborative innovation activities do not always produce patentable technologies. A comprehensive measurement of collaborative innovation activities, if possible, would improve analyses and estimates. Second, our research does not directly measure the travel cost of HSR relative to other modes of transport. Yet, the question of how HSR saves travel costs to enhance innovation collaboration can be interesting and worth exploring. Third, future work could extend to investigations into how HSR connections affect other forms of intercity open innovation practices, such as intercity technology licensing and joint ventures.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in Mendeley Data at http://doi.org/10.17632/rngz5kyb22.2.

DISCLOSURE STATEMENT

No potential conflict of interest was reported by the authors.

FUNDING

This work was supported by the National Social Science Foundation of China [grant number 20&ZD124].

NOTES

1. URI collaborations include university–university (UU), university–research institutes (UR), university–industry (UI) and research institute–industry (RI). Any co-patents involving universities or research institutes are categorized as URI collaborations.

2. We use a binary dummy variable indicating whether a city is connected to the HSR network for two reasons. First, easier interpretation: using binary variables we can interpret the results as treatment effect. A continuous variable such as volume or number of trains can be difficult to interpret. Second, the HSR effect is unlikely to be linear in relation to the volume or numbers of trains.

ORCID

Li Yao © http://orcid.org/0000-0002-7283-7462
Jun Li © http://orcid.org/0000-0002-6737-2728

REFERENCES

Agrawal, A., & Goldfarb, A. (2008). Restructuring research: Communication costs and the democratization of university innovation. American Economic Review, 98(4), 1578–1590. https://doi.org/10.1257/aer.98.4.1578

Bai, C.-E., Du, Y., Tao, Z., & Tong, S. Y. (2004). Local protectionism and regional specialization: Evidence from China’s industries. Journal of International Economics, 63(2), 397–417. https://doi.org/10.1016/S0022-1996(03)00070-9

Balland, P.-A., Boschma, R., & Frenken, K. (2015). Proximity and innovation: From statics to dynamics. Regional Studies, 49(6), 907–920. https://doi.org/10.1080/00343404.2014.883598

Bercovitz, J., & Feldman, M. (2011). The mechanisms of collaboration in inventive teams: Composition, social networks, and geography. Research Policy, 40(1), 81–93. https://doi.org/10.1016/j.respol.2010.09.008

Bin, P., Chen, X., Fracasso, A., & Tomasi, C. (2020). Firm employment growth in China: The role of marketization and regional economic factors. Growth and Change, 51(1), 402–439. https://doi.org/10.1111/grow.12348
Based on Chinese patent data. Research Policy, 42(2), 454–464. https://doi.org/10.1016/j.respol.2012.05.012

Inoue, H., Nakajima, K., & Saito, Y. U. (2017). The impact of the opening of high-speed rail innovation. RIETI.

Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. The Quarterly Journal of Economics, 108(3), 577–598. https://doi.org/10.2307/2118401

Ke, X., Chen, H., Hong, Y., & Hsiao, C. (2017). Do China’s high-speed-rail projects promote local economy? New evidence from a panel data approach. China Economic Review, 44, 203–226. https://doi.org/10.1016/j.checo.2017.02.008

Lawrence, M., Bullock, R., & Liu, Z. (2019). China’s high-speed rail development. OCLC. 1101272415.

Li, H., Meng, L., & Zhang, J. (2006). Why do entrepreneurs enter politics? Evidence from China. Economic Inquiry, 44(3), 559–578. https://doi.org/10.1093/ei/cbj031

Marek, P., Tirze, M., Fuhrmeister, C., & Blum, U. (2017). R&D collaborations and the role of proximity. Regional Studies, 51(12), 1761–1773. https://doi.org/10.1080/00343404.2016.1242718

Maurseth, P. B., & Verspagen, B. (2002). Knowledge spillovers in Europe: A patent citations analysis. Scandinavian Journal of Economics, 104(4), 531–545. https://doi.org/10.1111/1467-9442.00300

Monzón, A., Ortega, E., & López, E. (2013). Efficiency and spatial equity impacts of high-speed rail extensions in urban areas. Cities, 30, 18–30. https://doi.org/10.1016/j.cities.2011.11.002

Neter, J., Kutner, M. H., Nachtsheim, C. J., & Wasserman, W. (1996). Applied linear statistical models. Irwin.

North, D. C. (1990). Institutions, institutional change and economic performance. Cambridge University Press.

Perkins, D. H. (1988). Reforming China’s economic system. Journal of Economic Literature, 26(2), 601–645.

Picci, L. (2010). The internationalization of inventive activity: A gravity model using patent data. Research Policy, 39(8), 1070–1081. https://doi.org/10.1016/j.respol.2010.05.007

Ponds, R., Van Oort, F., & Frenken, K. (2007). The geographical and institutional proximity of research collaboration. Papers in Regional Science, 86(3), 423–443. https://doi.org/10.1111/j.1435-9957.2007.00126.x

Sasaki, K., Ohashi, T., & Ando, A. (1997). High-speed rail transit impact on regional systems: Does the Shinkansen contribute to dispersion? The Annals of Regional Science, 31(1), 77–98. https://doi.org/10.1007/s001680050040

Scott, W. R. (2013). Institutions and organizations: Ideas, interests, and identities. Sage.

Sun, X., Zhang, Y., & Wandelt, S. (2017). Air transport versus high-speed rail: An overview and research agenda. Journal of Advanced Transportation, 2017. https://doi.org/10.1155/2017/8426926

Torre, A., & Rallet, A. (2005). Proximity and localization. Regional Studies, 39(1), 47–59. https://doi.org/10.1080/00343400520000320842

Vickerman, R. (2018). Can high-speed rail have a transformative effect on the economy? Transport Policy, 62, 31–37. https://doi.org/10.1016/j.tranpol.2017.03.008

Yao, L., Li, J., & Li, J. (2020). Urban innovation and intercity patent collaboration: A network analysis of China’s national innovation system. Technological Forecasting and Social Change, 160, 120185. doi:10.1016/j.techfore.2020.120185