High speed rail beyond core hubs: evidence from Italy

Anna Bottasso
University of Genoa
Maurizio Conti
University of Genoa
Simone Robbiano
simone.robbiano@economia.unige.it
University of Genoa

Research Article

Keywords: High-Speed Rail (HSR), Regional Development and Innovation, Public Infrastructure Investment, Difference-in-Differences, Synthetic Control Method, Place-Based Policy

Posted Date: December 8th, 2025

DOI: https://doi.org/10.21203/rs.3.rs-8164189/v1

Additional Declarations: No competing interests reported.

High speed rail beyond core hubs: evidence from Italy*

Anonymous

November 2025

Abstract

This paper investigates the causal effects of high-speed rail (HSR) access on local economic performance by focusing on the 2013 opening of the MedioPadana HSR station in Reggio Emilia, Italy—a non-metropolitan area situated between the major nodes of Milan and Bologna. Using a multi-method approach that combines Difference-in-Differences (DiD), Synthetic Control Methods (SCM), and Synthetic Difference-in-Differences (SDID), and leveraging a rich economic/innovation dataset from 1980 to 2023, the study finds that improved HSR connectivity led to a sustained increase in regional Gross Value Added (GVA) of about 9.4%. The effect is particularly pronounced in tradable sectors such as manufacturing and business services, while local non-tradables such as retail and construction experienced relative declines. Mechanism analysis suggests that these gains were driven by increases in labour productivity and scale expansion (rather than capital deepening), as well as in the quality (quantity) of innovation—as measured by forward citation-weighted (unweighted) patent counts. Spillover effects are observed in nearby provinces, thus suggesting localised positive externalities rather than displacement. By focusing on a politically motivated, quasi-random station placement in an intermediate region, the paper provides novel evidence that HSR can foster growth outside of core metropolitan hubs—provided that local capabilities and institutional conditions are in place. The findings offer policy-relevant insights into the design of spatially balanced infrastructure strategies.

Keywords: High-Speed Rail (HSR); Regional Development and Innovation; Public Infrastructure Investment; Difference-in-Differences; Synthetic Control Method; Place-Based Policy.

JEL codes: C21; H54; O18; O31; R11; R42.

1 Introduction

High-speed railway (HSR) has become a flagship instrument of spatial and industrial policy across Europe, with profound implications for regional economic development, and routinely justified on the grounds of productivity gains, market integration and reductions in generalised transport costs. Indeed, HSR reduces travel time and costs, facilitating the flow of labour, capital, and knowledge between regions, thereby potentially reshaping economic geography through agglomeration and spillover effects.

*Compliance with Ethical Standards: the Authors declare no potential conflicts of interest; this work does not involve human participants and/or animals; no informed consent is needed. All of the sources of funding for the work described in this publication are acknowledged below: PRIN (Research Projects of National Relevance) - Italian Ministry of Education, University and Research (MIUR) - Project "INFINITE - INfrastructure: Firms and Individual ouTcomEs" - Project ID 20228JRZYZ.

It is therefore unsurprising that investments in these transport infrastructure are central to place-based policies [Neumark and Simpson, 2015], yet economic theory is divided on whether they universally boost local economies or merely shift activity across space, potentially widening disparities [Redding and Turner, 2015]. In particular, Koster et al. [2022] show that, under certain conditions, new HSR transportation links can harm bypassed areas and may not even improve outcomes for connected intermediate regions.

These ambiguous predictions are also confirmed by the empirical literature stemming from economics, regional science and economic geography.[1] Indeed, HSR has attracted growing attention as a transformative infrastructure with potentially far-reaching economic effects. In particular, some scholars suggest that HSR boosts GDP growth [Ke et al., 2017, Jin et al., 2020, Li et al., 2020, Fan et al., 2022], skilled labour mobility [Baltrunaite and Karmaziene, 2020], house prices and residential land values [Le Boennec et al., 2022, Bottasso et al., 2025a], capital mobility [Duan et al., 2021], labour's agglomeration [Wang et al., 2023] and market integration [Liu et al., 2022, Feng et al., 2023], mainly in large or developed cities, and improves urban efficiency over time, promoting local economic performance [Wetwitoo and Kato, 2017], though spillovers are limited. Conversely, Qin [2017] argues that counties located along the upgraded railway lines saw declines in both GDP and GDP per capita after the upgrade, primarily due to a simultaneous decrease in fixed asset investment; Koster et al. [2022], focusing on the Japanese HSR network, suggest instead that 'in-between' municipalities that are connected to the latter witness significant reductions in employment.[2]

Recent studies provide further nuance. For example, Yoo et al. [2023] highlight that HSR investments may have dual effects, boosting growth in core regions while exacerbating regional disparities. Conversely, Li et al. [2023b] show that HSR expansion can accelerate industrial relocation, particularly from central to coastal areas, thus reshaping regional production landscapes. Moreover, Yoo et al. [2024] find that HSR brings strong benefits on income and population to urban areas, while, in contrast, rural regions experience economic decline and increasing population ageing.[3] Technological spillovers are another emerging theme: Wang et al. [2022] demonstrate that cross-city HSR connections significantly foster firm-level innovation, especially in industries with high R&D intensity. What is more, they suggest that megacities benefit most from the HSR network, and secondary cities also benefit from being connected to megacities by HSR lines. Noticeably, Bottasso et al. [2025b] analyse the impact of HSR connection on regional innovation performance in Italy, suggesting meaningfully increases in both quality and quantity of innovative activities. From a spatial perspective, Li et al. [2023a] argue that HSR generates a "space-time compression effect" that reinforces market accessibility, while Sun et al. [2023] show that such improvements also raise urban economic complexity and coordination across city networks, but can increases the economic gap between center cities and non-center ones. Bottasso et al. [2025b] finally argue that HSR boosted inter-regional inventors collaboration and regional embeddedness in the network of innovation.[4]

Yet the empirical debate remains disproportionately anchored to outcomes in major metropolitan

[1] Ferrari et al. [2018] provide a broader survey that also covers more recent works across transport modes and economic outcomes. See also Blanquart and Koning [2017] on HSR.

[2] European studies, such as the ones on the Milan–Bologna corridor [e.g. Di Matteo et al., 2023], find positive economic impacts for connected provinces, with sectoral differences linked to specialisation and station centrality. Employment effects are instead heterogeneous: some European evidence finds no significant HSR impact [Albalate and Fageda, 2016], while Chinese studies highlight productivity impacts that vary by sector and location, with manufacturing often benefiting via spillovers and services experiencing concentration effects, particularly near major cities [Zhou and Zhang, 2021].

[3] Similar results can be found in Zhang and Gibson [2025].

[4] Other dimensions, such as well-being [Chen and Chen, 2023] and environmental co-benefits [Momenitabar et al., 2021], are also gaining prominence in recent work, though findings remain context-specific.

hubs. This focus is understandable—large cities concentrate economic mass and often capture a sizeable share of network flows—but it risks obscuring the welfare and distributional consequences for the vast set of intermediate regions that lie between the principal termini.[5] Among the few studies that have focused specifically on this issue, Ahlfeldt and Feddersen [2018] study the economic impact of the German HSR connecting Cologne and Frankfurt, suggesting 8.5% higher GDP for counties with intermediate stops. Bottasso et al. [2023] concentrate instead on the Italian HSR network, arguing that intermediate HSR stations opening increased local firms' TFP by about 5%, with stronger effects for nearby firms and variation by industry, size, and prior productivity. Finally, Zheng and Wu [2024] analyse the impact of HSR station placements on urban economic activities within remote Chinese cities, suggesting that the latter typically function as new city subcenters, consistently driving the increase of both land supply and economic activities.[6]

Whether HSR amplifies the dominance of core cities or diffuses opportunities along the network is ultimately an empirical question. Answering it requires credible designs that move beyond capital-city case studies and evaluate the local effects of stations sited in medium-sized territories. To this end, this study leverages the quasi-random placement of an intermediate HSR station in northern Italy, located near Reggio Emilia, to causally assess how integration into the HSR network has influenced the regional Gross Value Added (GVA) trajectory. Specifically, the analysis leverages a rich dataset from 1980 to 2023 and is based on a Differences-in-Differences (DiD) research design, where the NUTS-3 region hosting the aforementioned intermediate HSR station belongs to the treatment group, while remaining ones—except those in which there are other HSR stations—represent the control group.[7]

It is worth noting that the case study under analysis raises some interesting issues. First, Reggio Emilia region has been the first in Italy to receive an intermediate station specifically designed for HSR;[8] what is more, Reggio Emilia HSR station is almost on the straight line connecting Bologna and Milan (see Figure 1), hence being a case suitable for the incidental treatment approach to identification suggested by Redding and Turner [2015]. Second, although situated in Italy's manufacturing heartland, this region has a comparatively small population relative to Milan and Bologna, the two regions it is directly connected to via HSR, so that it is a well-designed test for predictions of Koster et al. [2022]'s model, that implies that mere connection to a network does not enhance an intermediate region's performance compared to an unconnected one, unless the connected region already possesses a sufficiently large market.[9] Last but not least, the opening of the intermediate station in Reggio Emilia has been the result of political negotiations, so that the assumption that the Reggio Emilia HSR station is as good as randomly assigned is quite reasonable.[10] The latter issue guarantees strong internal validity to the case study under scrutiny.

Results indicate a sustained and economically meaningful uplift in regional GVA associated to the opening of the HSR station—between 3% and 11% across specifications—robust to an extensive set of predetermined controls (economic structure, population, institutional quality, and multimodal access-

[5] The term termini refers to major transportation hubs.

[6] See also Moyano and Dobruszkes [2017], Bernard et al. [2019].

[7] In this study, both the terms "region" and "NUTS-3 region", as well as "province", will be used interchangeably to indicate the Italian NUTS-3 statistical territorial units.

[8] See Sala [2017] and `https://www.fsitaliane.it/en/innovation/transport-technology/main-hs-stations/reggio-emilia-mediopadana-hs-station.html`.

[9] The region of Reggio Emilia has about half a million inhabitants, while the metropolitan areas of Milan and Bologna have about 3.25 and 1 million inhabitants, respectively.

[10] The Italian Prime Minister at the time, Romano Prodi, in an article from the daily *La Gazzetta di Parma* (2022, April 10) argued that the decision was also taken on the basis of political hesitation on the part of politicians in Parma. See `https://www.gazzettadiparma.it/parma/2022/04/10/news/prodi-merito-mio-la-mediopadana-a-reggio-emilia-638303/`.

ibility) and exhibiting a gradual build-up consistent with standard adjustment dynamics in transport economics. Sectoral evidence points to a reallocation towards manufacturing and tradable business services, with contractions in local non-tradables such as construction and retail, indicating a shift towards higher-productivity activities. The analysis of the transmission mechanism suggests that the gains are not driven by capital deepening but by scale and efficiency: higher hours worked (6.4%), accompanied by increases in labour productivity (2.6%) and compensation (1.8%), as well as—most notably—by an improvement in innovation quality (84%), alongside a rise in quantity (34.4%), as measured by forward-citation-weighted and unweighted patent counts, respectively, rather than by a larger inventor headcount. Robustness checks reinforce these conclusions: synthetic control method (SCM) and synthetic difference-in-differences (SDID) estimates corroborate the timing and magnitude of the effects; moreover, placebo exercises do not generate spurious impacts and a "donuts-hole" design—that removes adjacent provinces from the control group—yields larger estimates, a pattern compatible with localised positive externalities to nearby areas rather than displacement.

From a policy perspective, these findings speak directly to ongoing debates around the spatial allocation of infrastructure investment. Much of the literature and policy discourse has focused on HSR's role in reinforcing the dominance of large metropolitan hubs. However, this paper's evidence clearly shows that strategically located stations in intermediate regions can generate economically meaningful and persistent returns—provided that complementary conditions are met. In this case, the gains observed in Reggio Emilia were not automatic consequences of physical infrastructure alone, but emerged from its interaction with local capabilities, existing firm structures, and institutional readiness.

This paper contributes to the HSR literature along different margins. First, it shifts the focus from major metropolitan hubs to an intermediate territory and, exploiting the quasi-random siting of the Reggio Emilia Mediopadana station, delivers credible causal evidence on local economic development and innovation outside the usual core-city settings. Second, it strengthens identification by combining Differences-in-Differences with Synthetic Control and Synthetic DiD, complemented by event-study dynamics and honest post-estimation inference [Abadie et al., 2015, Arkhangelsky et al., 2021, Rambachan and Roth, 2023]. Third, it broadens the outcome space: beyond aggregate GVA, it documents sectoral reallocation towards tradables and an upgrading of innovation quality (forward-citation–weighted patents) without a parallel rise in inventor headcount—consistent with denser knowledge flows and effective appropriation—while also highlighting possible localised spillovers through a "donuts-hole" design. Taken together, the results speak to ongoing debates on whether HSR reinforces core-periphery divides or diffuses opportunities: in this intermediate setting, the evidence aligns with meaningful upgrading when local capabilities and institutional complements are in place, and is of comparable or larger magnitude than effects reported in related European contexts [Ahlfeldt and Feddersen, 2018, Bottasso et al., 2023]. Finally, by providing transparent effect sizes and adjustment dynamics that can be mapped into time profiles and spatial reach, the study offers a practical benchmark for cost–benefit appraisal. Policymakers can juxtapose these empirics with capital and operating costs to inform where and how to prioritise comparable investments, especially when integrating transport upgrades with land-use, skills, and business-support policies in intermediate regions

The remainder of the paper is structured as follows. Section 2 describes the data sources and the empirical strategy, respectively. Section 3 presents the main results, including robustness checks and heterogeneity analyses. Section 4 concludes with a discussion of policy implications and directions for future research.

2 Material and Methods

2.1 Data

The empirical analysis relies on a balanced panel dataset covering Italian provinces (NUTS-3 level) over the period 1980–2023 (4,224 observations).[11] The primary source of economic data is the ARDECO repository provided by the European Commission's Joint Research Centre (JRC), which offers harmonised regional accounts for Gross Value Added (GVA), capital stock, employment, labour compensation, and other key production-side indicators.[12] These data allow for a consistent evaluation of economic dynamics across provinces and over time, and are well suited to long-run, policy-oriented regional analysis. In particular, the dataset used in this work includes: Gross Value Added (GVA, total and by sector), GDP, labour productivity (per hour and per person), unit labour costs, compensation per hour worked, wage and salary earners, total and sectoral employment, hours worked, capital stock, gross fixed capital formation (GFCF), and consumption of fixed capital. Demographic variables include population, population change, and net migration. These variables enable decomposition of regional growth into labour, capital, and productivity components, while sectoral and demographic data help identify structural shifts and spatial responses linked to high-speed rail access.

In addition to economic performance metrics, the study also incorporates detailed information on local innovation outcomes, namely a set of fractional-count patent indicators, including both unweighted and forward citation-weighted counts, as well as the number of inventors patenting in the region. These measures are constructed following established methods in the innovation literature and account for both the intensity and the technological significance of local inventive activity.

In particular, the analysis primarily draws on fractional patent counts, weighted (and unweighted) by forward citations, as a proxy for innovation output.[13] The economic literature widely acknowledges patents as key instruments for the appropriation of innovative activity; furthermore, technologies that exert a greater impact on welfare and economic development are more likely to be patented [Pakes and Griliches, 1980]. Within the field of innovation studies, forward citations have been identified as an indirect indicator of an invention's value, since the number of citations received by a patent is generally associated with its influence on subsequent technological developments [Pakes and Griliches, 1980, Trajtenberg, 1990, Hall et al., 2005]. Nonetheless, existing scholarship highlights that patent data represent a useful measure of regional technological capacity. To this end, annual patent data are drawn from the European Patent Office (EPO) repository (EPO–Patstat), which provides bibliographical and legal status information for patents across EU regions. The dataset covers the period 1980–2023, incorporating information on patent applications filed directly under the European Patent Convention, as well as on patent applications filed under the Patent Co-operation Treaty and designating the EPO (Euro-PCT). The database includes detailed information such as the number of forward citations, the

[11] Italy has 107 NUTS-3 regions; in the analysis we drop Milan, Turin, Bologna, Florence, Rome and Naples, given that they host another HSR stations; moreover, we exclude Sardinian provinces, given that they do not have railway continuity with the continent

[12] ARDECO is the Annual Regional Database compiled by the European Commission's Directorate-General for Regional and Urban Policy and regularly updated by the Joint Research Centre. It provides extensive time-series data and indicators for EU regions, along with selected regions in EFTA and candidate countries, across different statistical levels (NUTS-1, NUTS-2, NUTS-3 and metropolitan areas). The information is structured into chapters covering themes such as population, gross domestic product, employment, labour costs, labour productivity, capital formation, capital stock and households. See `https://territorial.ec.europa.eu/ardeco/explorer?lng=en`.

[13] The geographical allocation of patent applications follows the "inventor criterion," whereby each patent is attributed to the place of residence of its inventor(s). In cases involving multiple inventors, the application is equally fractionalised across all inventors and, accordingly, across their respective NUTS-3 regions of residence. A one has been added to all patent and inventor count variables before taking the log to include observations with values of zero.

identity and characteristics of applicants and inventors, and the technological classification of patents according to the International Patent Classification (IPC) and the NACE-2 classification of economic activity. Raw patent data are "regionalised" on the basis of inventors' addresses (NUTS-3 codes). Data are available up to 2023, given that figures for the two most recent years may underestimate application counts due to the typical 18–24 month delay in patent publication.[14]

Last but not least, the dataset includes a measure of institutional quality, namely the Institutional Quality Index (IQI) as developed by Nifo and Vecchione [2014], that relies on five groups of elementary indexes (evaluating corruption, governance, regulation, law enforcement and social participation) whose aim is to measures institutional quality at the provincial and regional levels for the period 2004-2019.

Finally, the dataset also includes an indicator for potential multimodal accessibility, in order to capture the absolute change in accessibility potential between regions over time. In particular, ESPON's potential multimodal accessibility is a gravity–type index that summarises how easily a region can interact with population or economic mass across the European transport system. It combines road, rail and air networks (including intermodal transfers) into a single, normalised measure at NUTS–3 level, reflecting both the size of reachable destinations and the generalised travel times required to reach them. Higher values indicate regions that can access larger markets more quickly, while lower values signal relative isolation.[15]

Table 1 reports summary statistics for the full sample, the treated unit (Reggio Emilia), and the control provinces, as well as for the pre- and post-treatment periods of the treated unit. From the one hand, treated and control provinces differ in levels, as expected. Reggio Emilia exhibits markedly higher values in GVA, labour productivity, compensation per hour, capital per hour worked, and patent-related indicators compared to the average control regions. These gaps reflect the relatively advanced industrial structure of Reggio Emilia prior to HSR connection. On the other hand, when comparing pre- and post-treatment periods within the treated unit, it is possible to observe substantial increases in key economic variables, including GVA, labour productivity, and capital per hour worked, as well as in patent counts and forward citation-weighted innovation metrics. These patterns are consistent with a structural break around the opening of the HSR station. Finally, it is important to note that observed differences in levels across groups and periods do not undermine a DiD identification strategy, provided that parallel pre-treatment trends assumption holds. This key assumption will be supported by estimates of an event-study design and further reinforced by robustness analyses employing both synthetic control and synthetic difference-in-differences methods, which closely match the treated unit's pre-treatment trajectory and validate the main findings, and described in the next Sub-section.

2.2 Method

To estimate the causal effect of HSR access on regional GVA, the paper implements a difference-in-differences (DiD) model. The most parsimonious Equation, estimated by OLS, is as follows:

$$\log GVA_{it} = \beta \cdot (Treated_i \cdot Post_t) + \alpha_i + \delta_t + \lambda_{jt} + \varepsilon_{it} \tag{1}$$

where GVA_{it} is the dependent variable, namely (log) regional GVA for NUTS-3 region i in year t, while $Treated_i$ and $Post_t$ are dummies equal to 1 for the Reggio Emilia province and for years after the

[14] See Bronzini and Piselli [2016].

[15] The index is reported as a comparable score (EU baseline set to 100), allowing straightforward benchmarking across places and over time. See https://database.espon.eu/indicator/1541/#metadata-download for methodological details.

Table 1: Summary statistics

Variables	Mean	SD	Min	Median	Max
	Full Sample				
GVA (Million EURO)	7,892.36	6,391.57	365.10	5,921.75	47,068.80
Labour productivity per hour worked (EURO)	27.22	10.23	4.71	27.76	56.26
Compensation per hour worked (EURO)	16.59	5.40	3.90	16.70	29.70
Hours worked (Thousands)	313,344.90	195,247.80	45,411.50	248,734.50	1,069,959.00
Capital per hour worked (Thousands EURO per hour worked)	0.11	0.04	0.04	0.11	0.26
Patent fractional count (#)	21.63	33.08	0.00	8.03	222.11
Forward citations weighted patent fractional count (#)	179.84	306.05	0.00	57.78	3,530.19
Forward citations (#)	8.97	8.86	0.00	7.60	249.14
Number of inventors (#)	36.07	55.43	0.00	13.00	492.00
	Treated Sample ($Treated_i = 1$)				
GVA (Million EURO)	17,366.32	1,743.38	15,260.80	17,065.50	20,950.20
Labour productivity per hour worked (EURO)	45.36	3.35	39.88	45.13	50.77
Compensation per hour worked (EURO)	25.55	1.30	23.60	25.30	27.50
Hours worked (Thousands)	405,211.70	31,130.63	349,384.60	412,407.20	471,731.30
Capital per hour worked (Thousands EURO per hour worked)	0.12	0.02	0.09	0.12	0.16
Patent fractional count (#)	101.13	17.41	78.09	96.65	124.88
Forward citations weighted patent fractional count (#)	420.82	195.15	205.52	410.63	759.77
Forward citations (#)	4.44	2.47	1.65	4.20	8.43
Number of inventors (#)	139.43	11.18	129.00	132.00	152.00
	Control Sample ($Treated_i = 0$)				
GVA (Million EURO)	7,867.63	6,380.95	365.10	5,908.10	47,068.80
Labour productivity per hour worked (EURO)	27.17	10.20	4.71	27.72	56.26
Compensation per hour worked (EURO)	16.57	5.39	3.90	16.70	29.70
Hours worked (Thousands)	312,377.90	196,018.70	45,411.50	247,193.00	1,069,959.00
Capital per hour worked (Thousands EURO per hour worked)	0.11	0.04	0.04	0.11	0.26
Patent fractional count (#)	21.49	32.92	0.00	8.00	222.11
Forward citations weighted patent fractional count (#)	179.40	306.06	0.00	57.49	3,530.19
Forward citations (#)	8.98	8.87	0.00	7.63	249.14
Number of inventors (#)	35.88	55.30	0.00	13.00	492.00
	Pre-treatment Period of Treated Sample ($Treated_i = 1$ and $Post_t = 0$)				
GVA (Million EURO)	9,930.17	3,972.17	3,393.30	9,449.50	16,285.80
Labour productivity per hour worked (EURO)	26.76	9.26	10.22	28.18	39.82
Compensation per hour worked (EURO)	15.85	4.65	7.30	15.70	23.60
Hours worked (Thousands)	398,993.10	32,431.92	349,384.60	391,695.10	471,731.30
Capital per hour worked (Thousands EURO per hour worked)	0.11	0.02	0.09	0.11	0.14
Patent fractional count (#)	46.38	33.75	3.58	37.12	101.49
Forward citations weighted patent fractional count (#)	387.80	274.86	32.33	337.43	844.04
Forward citations (#)	8.55	2.55	5.25	7.89	18.37
Number of inventors (#)	61.64	47.36	6.00	46.00	147.00
	Post-treatment Period of Treated Sample ($Treated_i = 1$ and $Post_t = 1$)				
GVA (Million EURO)	17,366.32	1,743.38	15,260.80	17,065.50	20,950.20
Labour productivity per hour worked (EURO)	45.36	3.35	39.88	45.13	50.77
Compensation per hour worked (EURO)	25.55	1.30	23.60	25.30	27.50
Hours worked (Thousands)	423,867.30	17,055.76	383,152.60	424,458.80	451,389.90
Capital per hour worked (Thousands EURO per hour worked)	0.14	0.01	0.13	0.14	0.15
Patent fractional count (#)	101.13	17.41	78.09	96.65	124.88
Forward citations weighted patent fractional count (#)	420.82	195.15	205.52	410.63	759.77
Forward citations (#)	4.44	2.47	1.65	4.20	8.43
Number of inventors (#)	139.43	11.18	129.00	132.00	152.00

opening of the station in 2013, respectively (and zero otherwise).[16] In turn, the β coefficient represents the average treatment effect on the treated (ATT) of the opening of the Reggio Emilia MedioPadana station in the post-treatment period, while ε_{it} is a stochastic error term.[17] Finally, the model includes a full set of both year (δ_t) and NUTS-3 regions (α_i) fixed effects, as well as NUTS-1-region-by-Year fixed effects (λ_{jt});[18] province fixed effects absorb all time-invariant characteristics and unobserved hetero-

[16] It is worth noting that regions different from Reggio Emilia hosting other HSR stations, namely Turin, Milan, Bologna, Florence, Rome and Naples, have been excluded from the control group.

[17] Robust standard errors are clustered at NUTS-3 region level.

[18] In the EU's statistical system, NUTS-1 is the highest sub-national tier used for broad socio-economic analysis and macro-area comparisons. In Italy there are five NUTS-1 macro-regions: North-West (ITC)—Piemonte, Valle d'Aosta/Vallée d'Aoste, Liguria, Lombardia; North-East (ITH)—Trentino-Alto Adige/Südtirol, Veneto, Friuli-Venezia Giulia, Emilia-Romagna; Center

geneity of each territory that could be correlated with the included regressors and that could also drive the dependent variable, while year fixed effects capture possible shocks common to all provinces in a given year. The inclusion of NUTS-1-region-by-Year fixed effects allows for NUTS-1 region j-specific time trends or shocks, thereby ensuring that our estimates of the treatment effect are not biased by differential trends or idiosyncratic events at macro-region level.

One advantage of including both NUTS-3 region and NUTS-1-region-by-Year fixed effects in a DiD research design is also to increase the comparability of the treatment and control groups. In the same vein, we also augments Equation 1 with a full set of control variables; in particular, these augmented versions of our baseline model include a matrix $\mathbf{X}_i'$ incorporating incremental sets of control variables capturing different regional predetermined characteristics: demographic factors (population), labour market indicators (labour productivity per hour, wage and salary earners, employment, nominal unit labour costs, compensation per hour worked, hours worked), capital endowment (capital stock), innovation capacity (patent counts, as described in the previous sub-section, number of inventors patenting in the region), economic structure (GDP, GVA, share of industry on total regional economy), institutional quality (as in Nifo and Vecchione [2014]), and regional potential multimodal accessibility.[19] In particular, following Wooldridge [2025], the more extended version of Equation 1 is as follows:

$$\log GVA_{it} = \beta \cdot \left(Treated_i \times Post_t\right) + \mathbf{X_i'}\theta + \left((Treated_i \times Post_t)\cdot \tilde{\mathbf{X}}_\mathbf{i}\right)'\varsigma + \left(\delta_t \cdot \mathbf{X_i}\right)'\pi + \alpha_i + \delta_t + \lambda_{jt} + \varepsilon_{it}, \quad (2)$$

where $\mathbf{X_i}$ is the aforementioned matrix of baseline covariates, $\delta_t \cdot \mathbf{X_i}$ allows for heterogeneous time trends by $\mathbf{X_i}$, and $(Treated_i \cdot Post_t)\cdot \tilde{\mathbf{X}}_\mathbf{i}$ (with $\tilde{\mathbf{X}}_\mathbf{i} = \mathbf{X_i} - \bar{\mathbf{X}}_{\text{treated}}$) allows the treatment effect to vary with X_i. The coefficient β is thus still the average treatment effect on the treated, but now evaluated at the mean covariate values of treated units.

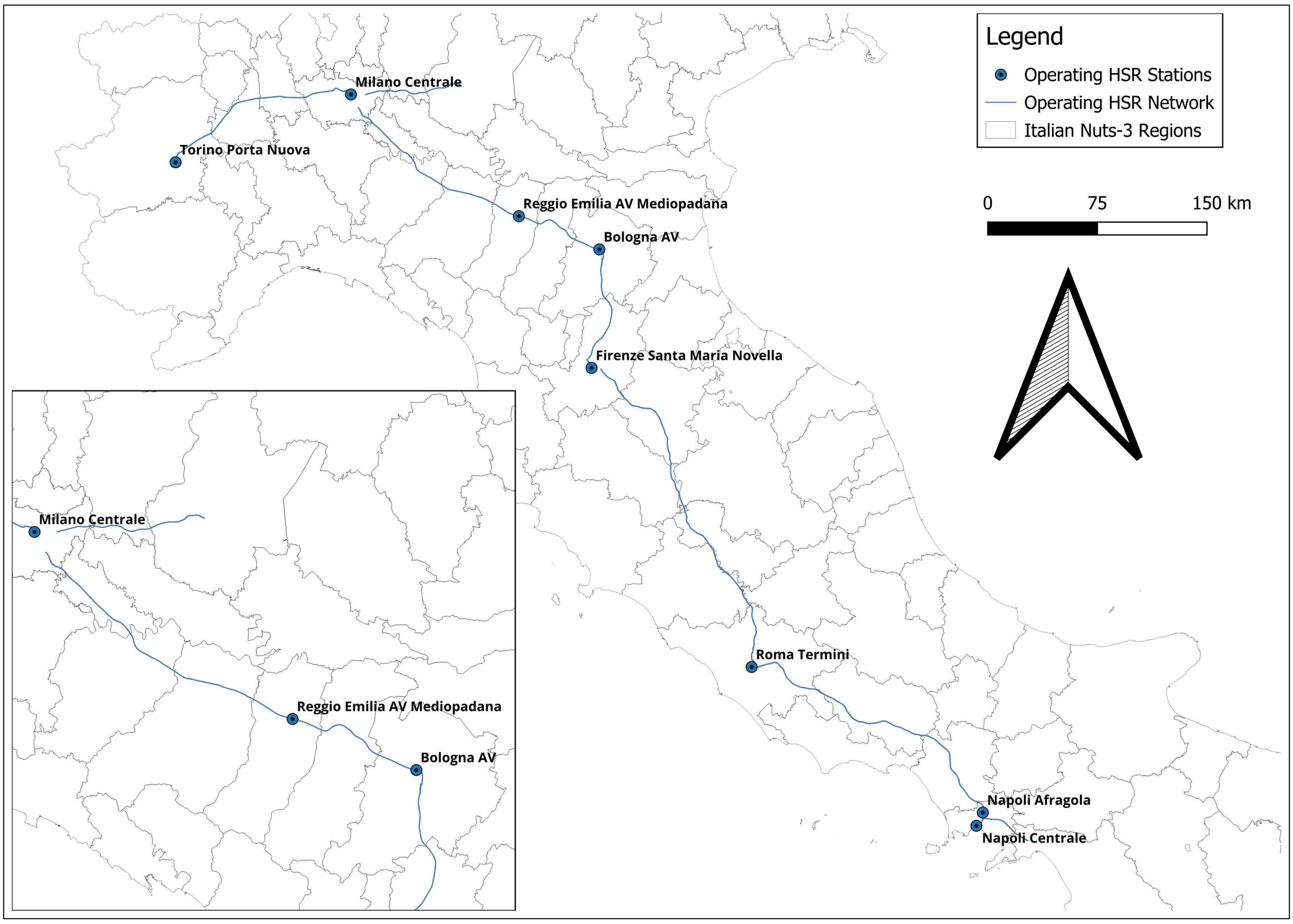
As in all DiD research designs, the identification strategy rests on the assumption of parallel trends in GVA between treated and control units. If that is true, regions located farther from the site of the MedioPadana station serve as an appropriate control group for Reggio Emilia, the treated unit. It is contended that the parallel trends assumption is plausible for several reasons. First, Reggio Emilia does not represent a central node in the Italian HSR network; it lies on the direct corridor between the already operational Bologna and Milan stations (see Figure 1). The opening of the new station thus constitutes a case of incidental treatment approach, a strategy frequently employed in the transport infrastructure evaluation literature [Redding and Turner, 2015, Koster et al., 2022, Bottasso et al., 2023]. Furthermore, the decision to locate the HSR station in Reggio Emilia, rather than in the more likely alternative of Parma, was primarily driven by political considerations. Hence, it is improbable that the siting of the station was determined by pre-existing economic characteristics of the Reggio Emilia area that could otherwise bias the observed evolution of regional GVA. Taken together, these factors suggest that the MedioPadana HSR station's siting is plausibly exogenous—close to random

(ITI)—Toscana, Umbria, Marche, Lazio; South (ITF)—Abruzzo, Molise, Campania, Puglia, Basilicata, Calabria; and Islands (ITG)—Sicilia, Sardegna. Researchers typically use these macro-regions to summarise national patterns and to account for broad geographical heterogeneity—for example, by reporting results by macro-area or including NUTS-1-by-Year fixed effects in panel regressions. Specifically, paper's baseline model includes province fixed effects and NUTS-1-by-Year fixed effects; thus, identification comes from variation between treated and untreated provinces within the same macro-region and year, net of time-invariant province heterogeneity and region-specific shocks. In practice, most within-region-year variation arises in the North-East. To assess sensitivity, the paper also re-estimates the model in Equation 1 without NUTS-1-by-Year fixed effects (using province and year fixed effects), whose results will be shown in column (1) of Table 2.

[19] It is worth noting that, as suggested by Wooldridge [2025], all the aforementioned controls have been taken at the beginning of the sample period, namely 1980, excepted for the institutional quality index and the potential multimodal access, whose first years of observation are 2004 and 2001, respectively.

assignment.

Figure 1: Italian HSR network and detail on Reggio Emilia Mediopadana station.



Notes: Author's elaboration.

Nevertheless, besides these arguments, the paper also conducts conventional tests for pre-trends in DiD research designs by estimating a dynamic version of Equation 1:

$$\log GVA_{it} = \sum_{\xi=-K^-}^{-1} \gamma_\xi \left(Treated_i \cdot Post_{t+\xi}\right) + \sum_{\xi=0}^{M} \varphi_\xi \left(Treated_i \cdot Post_{t+\xi}\right) + \alpha_i + \delta_t + \lambda_{jt} + \varepsilon_{it} \tag{3}$$

where, differently from Equation 1, the K leads of the treatment interaction span from 4 to 1 years before the treatment and the γ's represent pre-trends coefficients. It is worth noting that this specification pools all more distant pre-treatment periods into a bin (period -4 and backward) in order to ensure that the event–study plot reflects treatment dynamics on a common support and to mitigate concerns about spurious pre-trends.[20] Instead, the M lags of treatment variable span from the treatment year ($t = 0$) to 10 years after the treatment and the φ's represent the DiD coefficients of interest, reflecting dynamic average treatment effects on the treated (dynamic ATTs).

From the one hand, it is worth noting that the paper implements the event study with four pre-treatment leads and eleven post-treatment lags, normalising the year immediately before treatment to zero. Leads are specified as -4^-, -3, -2 and -1, where the first aforementioned bin pools all earlier years. This choice balances interpretability and precision. First, binning distant pre-periods reduces

[20] This is in line with Freyaldenhoven et al. [2019], so that we bottom-code distant pre-periods and focus inference on the immediate pre-treatment window to reduce noise and bias.

sampling noise and leverage from thin support far from the intervention date—particularly relevant with a single treated cohort—and avoids inflating Type I error from multiple pre-trend coefficients. Secondly, it mitigates comparability issues stemming from secular changes and measurement breaks in the early decades, while preserving a data-rich test for pre-trends immediately before treatment (2009–2012). In short, the choice defining lead -4^{-} trades a negligible loss in flexibility for substantial gains in statistical power and robustness of the pre-trend assessment [see Freyaldenhoven et al., 2019]. On the other hand, the specification of Equation 3 reports eleven lags, which exhaust the observed post-treatment support from 2013 to 2023. This horizon is long enough to capture short-run adjustment, medium-run diffusion (e.g., supply-chain, labour-market and knowledge spillovers), and convergence towards a post-treatment plateau, without extrapolating beyond the sample. Keeping the post window aligned with observed support also limits composition/weighting changes that can arise at very long horizons and maintains reasonable precision—standard errors typically widen at the tails even with rich fixed effects. Overall, this event-time design prioritises credible pre-trend scrutiny near the intervention and an empirically supported characterisation of dynamic effects over the full post-treatment period.

To alleviate concerns about differential trends between treated and control groups, the paper also implements inference using the procedure developed by Rambachan and Roth [2023], which constructs "honest" confidence intervals for event–study designs in DiD settings.[21] Their approach formalises the intuition that pre-trends are informative about possible violations of parallel trends. Indeed, conventional confidence intervals may undercover when treatment effect heterogeneity is present, since pre-treatment coefficients are typically used for testing the parallel trends assumption while post-treatment coefficients are used for estimation. The authors address this issue by introducing restrictions that bound the magnitude of post-treatment violations relative to those observed in the pre-treatment period. For example, setting $\bar{M} = 1$ assumes that post-treatment violations of parallel trends are no larger than the worst violation observed before treatment, while $\bar{M} = 2$ allows them to be at most twice as large. By combining information from pre-treatment estimates with such bounds, this procedure delivers confidence intervals that remain valid even when treatment effects are heterogeneous across cohorts or over time.[22]

Given that our design features a single treated region, namely Reggio Emilia province, and a rich control sample, the paper also complements baseline DiD of Equation 2 with both the Synthetic Control Method (SCM) proposed by Abadie et al. [2015] and the Synthetic Difference-in-Differences (SDID) estimator [Arkhangelsky et al., 2021] to construct a transparent, pre-trend–matched counterfactual and to localise identification to comparable donors and periods, thereby mitigating concerns about reliance on strict parallel trends. From the one hand, SCM constructs a transparent counterfactual by choosing non-negative unit weights that sum to one so that a weighted average of controls closely reproduces the treated region's pre-treatment trajectory (and covariates), thereby reducing model-dependent extrapolation and making unit contributions explicit; indeed, this design is well suited to comparative case studies with few treated units and emphasises pre-intervention fit as the primary credibility check. On the other hand, SDID combines the strengths of DiD and SCM by learning unit weights that align pre-treatment trends between treated and control units and time weights that balance pre- versus post-treatment periods, and then estimating a weighted two-way fixed-effects regression (TWFE); relative to plain DiD (which implicitly uses uniform unit and time weights), this localisation weakens reliance on

[21] This analysis relies on the `honestdid` Stata command. See Rambachan and Roth [2023] and Bravo et al. [2024] for details.

[22] In particular, the paper sets $0.5 \le \bar{M} \le 2$.

strict parallel trends and can improve precision by down-weighting dissimilar units and non-comparable periods. Formally, "Like SCM, the method reweights and matches pre-exposure trends to weaken the reliance on parallel trend type assumptions. Like DID, the method is invariant to additive unit-level shifts, and allows for valid large-panel inference" [Arkhangelsky et al., 2021, page 4089].[23]

Last but not least, it should be pointed out that, to verify the robustness of our results, we perform a full set of falsification and placebo tests, whose results are discussed in Section 3.1.

3 Results

In this Section, we discuss the main results on the effect of the 2013 opening of the Reggio Emilia MedioPadana HSR station on regional GVA.

Table 2 reports estimates from both Equations 1 and 2 with a full sequence of incremental, predetermined controls. In particular, column (1) reports a sensitivity specification of Equation 1 that excludes NUTS-1-by-Year fixed effects (whilst retaining province and year fixed effects), while column (2) includes the latter, as described in Section 2.2; columns (3)–(9) provide instead estimates of Equation 2 that include NUTS-1-by-Year fixed effects and progressively augmented sets of controls. Across specifications, the coefficient on the treatment interaction is positive and statistically significant at 1% level. In particular, the result in column (1) suggests a 13.77% increase in regional GVA post-treatment, while the inclusion of NUTS-1-by-Time FEs significantly reduces the size of the effect (column 2), with magnitudes that rise from about 3.1% in the bare-bones model to peaks around 11.6% once controls for population, labour/productivity, capital, and innovation; in the most complete specifications the effect stabilises at roughly 9.4%. These patterns indicate a robust post-2013 uplift in GVA for Reggio Emilia relative to comparable provinces. Specifically, relative to the parsimonious specification in Equation 1 (column 2), it is worth noting that the estimated effect is sensitive to the initial inclusion of specific covariates. Indeed, moving from columns (2)–(3) to the enriched specification that adds labour–productivity controls, the coefficient increases from about 0.03 to roughly 0.10 log points (column 4). Beyond that, augmenting the model with capital, innovation, economic size, institutional quality, and multimodal accessibility controls leaves the magnitude broadly stable in the 0.09–0.11 range (columns from 5 to 9), with overlapping confidence intervals even when the sample shrinks in the richest specifications. This pattern suggests that predetermined differential labour/productivity composition accounts for the bulk of the adjustment between Equation 1 and the covariate-rich models, whereas additional controls are not first-order confounders; if anything, richer controls slightly sharpen precision and leave point estimates broadly unchanged, suggesting that overall differential composition and trend heterogeneity captured by $\mathbf{X_i}$ are not driving the results.[24]

As emphasised in the previous Section, the credibility of the DiD design hinges on its identifying assumptions; accordingly, we first assess the absence of anticipation effects and evaluates the parallel-trends assumption using the event-study specification of Equation 3.

Figure 2 shows results of this analysis, highlighting no statistically significant pre-treatment deviations, being the estimated pre-trend coefficients for the K leads of the treatment interaction—spanning

[23] With one treated region, SCM offers a clear, visual and testable counterfactual tailored to the treated unit's pre-path; SDID then provides a complementary estimator that (i) explicitly incorporates unit fixed effects and (ii) uses data-driven unit and time weighting to make the DiD comparison local to the most comparable donors and periods, thereby strengthening credibility when simple parallel trends may be questionable and often improving statistical efficiency.

[24] All specifications in columns from (2) to (9) include NUTS-3 region, year, and NUTS-1-region-by-Year fixed effects, so identification relies on within–province deviations from region–wide shocks and common time effects; the persistence of the estimates under this demanding set of controls reinforces their credibility.

Table 2: DiD estimates of the impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: regional Gross Value Added ($\log GVA_{it}$)									
$Treated_i \cdot Post_t$	0.129***	0.031***	0.033***	0.098***	0.093***	0.110***	0.100***	0.092***	0.090***
	[0.011]	[0.005]	[0.006]	[0.023]	[0.023]	[0.028]	[0.030]	[0.029]	[0.029]
NUTS-3 region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
NUTS-1-by-Time FE	✗	✓	✓	✓	✓	✓	✓	✓	✓
Population CV	✗	✗	✓	✓	✓	✓	✓	✓	✓
Labour & Productivity CV	✗	✗	✗	✓	✓	✓	✓	✓	✓
Capital CV	✗	✗	✗	✗	✓	✓	✓	✓	✓
Innovation CV	✗	✗	✗	✗	✗	✓	✓	✓	✓
Economic Size CV	✗	✗	✗	✗	✗	✗	✓	✓	✓
Institutional Quality Index	✗	✗	✗	✗	✗	✗	✗	✓	✓
Mult. Access Index	✗	✗	✗	✗	✗	✗	✗	✗	✓
Observations	4,224	4,224	4,224	4,224	4,224	4,224	4,224	4,092	4,092

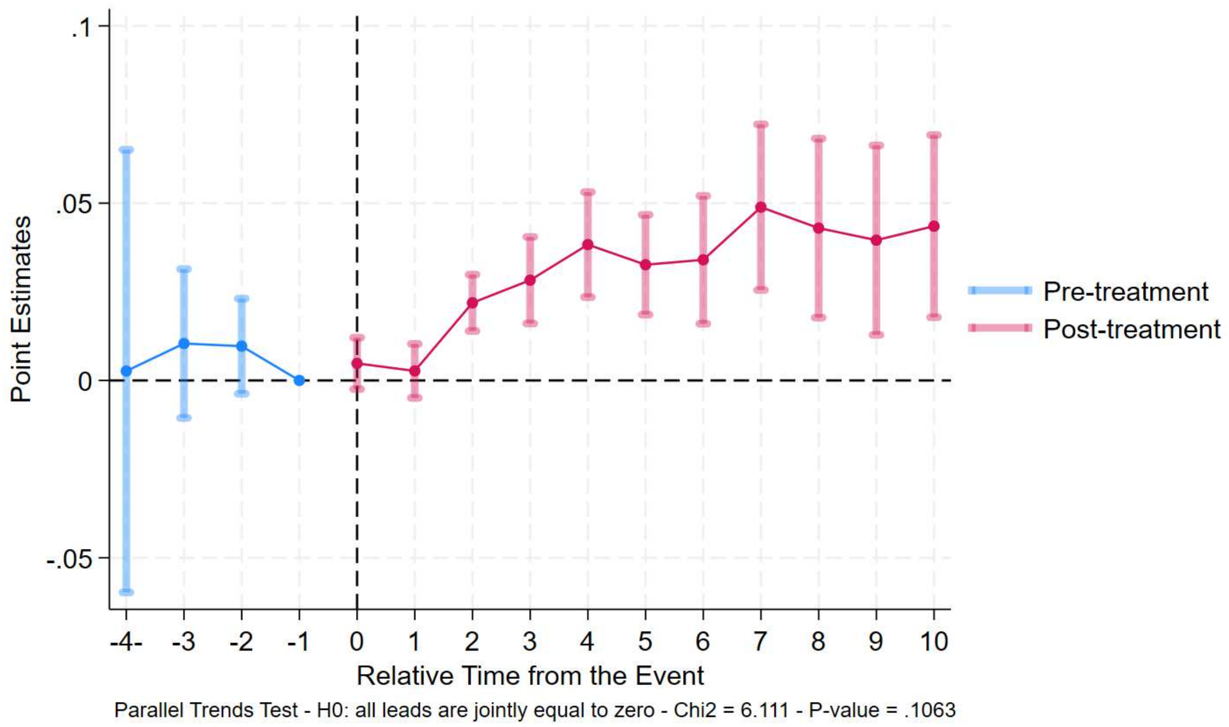
Notes: This Table presents DiD estimates of Equation 1 (columns 1 and 2) and Equation 2 (columns from 3 to 9) on the effect of the opening of Reggio Emilia MedioPadana HSR station on regional GVA. All specifications include both NUTS-3 region and time fixed effects, as well as NUTS-1-region-by-Year ones, with the exception of column (1), which does not include the latter. Controls include predetermined regional characteristics at the beginning of the sample period: demographic factors (population), labour market and productivity indicators (labour productivity per hour, wage and salary earners, employment, nominal unit labour costs, compensation per hour worked, hours worked), capital endowment (capital stock), innovation capacity (Patent fractional counts, number of inventors patenting in the region), economic structure (GDP, GVA, share of industry on total regional economy), institutional quality (as in Nifo and Vecchione [2014]), and regional potential multimodal accessibility. All specifications are estimated by OLS. Robust standard errors, clustered at NUTS-3 region level, are shown in parentheses: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

from 4 (and backward) to 1 year before the treatment—individually and jointly equal to zero. Indeed, a joint test that all leads are simultaneously equal to zero reject the null hypothesis at the 1% significance level ($p-value \approx 0.106$). What is more, the coefficient associated to the lead $-K^- = -4^-$ is statistically not different from zero, suggesting therefore no divergent trends between treated and control units for periods from 4 years prior to treatment and backward. Moreover, the specification enables an assessment of dynamic adjustment—whether the treatment effect accelerates, stabilises, or exhibits mean reversion—by incorporating M lags of the treatment from the implementation year, namely 2013 ($t = 0$), up to ten years thereafter. Post-treatment, coefficients become positive and increase over time, consistent with a gradual build-up of effects rather than a one-off jump. On this issue, consistent with standard adjustment dynamics in transport economics, effects are unlikely to materialise immediately upon opening. In particular, the event–study in Figure 2 shows coefficients that are statistically indistinguishable from zero in 2013–2014, with positive impacts emerging only from 2015 onwards. This pattern is plausible and confirm those found by Bottasso et al. [2023]: firms and households require time to learn about the new connectivity, revise location and sourcing choices, and undertake complementary investments; local authorities and private actors must complete ancillary infrastructure and service adjustments; labour and product markets re-optimise through contracting, hiring, and supply-chain reconfiguration; and innovation spillovers accumulate gradually. Together, these frictions and complementarities imply a short gestation period followed by growing effects, which aligns with the delayed but persistent post-treatment gains that are estimated.[25]

We further address potential violations of the parallel trends assumption through the framework by Rambachan and Roth [2023]; to do so, the paper relies on the *HonestDiD* Stata®package from Bravo et al. [2024] and explores possible violations of the parallel trends by varying the sensitivity parameter $\bar{M}$ over $[0.5, 2]$, where $\bar{M}$ bounds the magnitude of post-treatment violations relative to the largest

[25] Moreover, it should be noted that the number of high-speed trains, namely the amount of Eurostar, Frecciarossa and Frecciargento class trains, serving the Reggio Emilia HSR station increased from about 100 in 2014 to around 160 in 2018.

Figure 2: Event study representation of the impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA.



Notes: OLS estimates of Equation 3 on the effect of the opening of Reggio Emilia MedioPadana HSR station on regional GVA. The specification includes both NUTS-3 region and time fixed effects, as well as NUTS-1-region-by-Year ones. The vertical dashed line identifies the HSR station opening year (2013). The blue dots represent the estimated pre-trend coefficients for the K leads of the treatment interaction, spanning from 4 (and backward) to 1 day before the treatment; the red dots show estimated coefficients of the M lags of the treatment interaction, spanning from the treatment year ($t = 0$) to 10 years after the treatment, reflecting dynamic ATTs. The vertical whiskers represent the respective 99% confidence intervals. The parallel trends test computes p-values of the joint test that all lead coefficients are equal to 0. Robust standard errors are clustered at NUTS-3 region level.

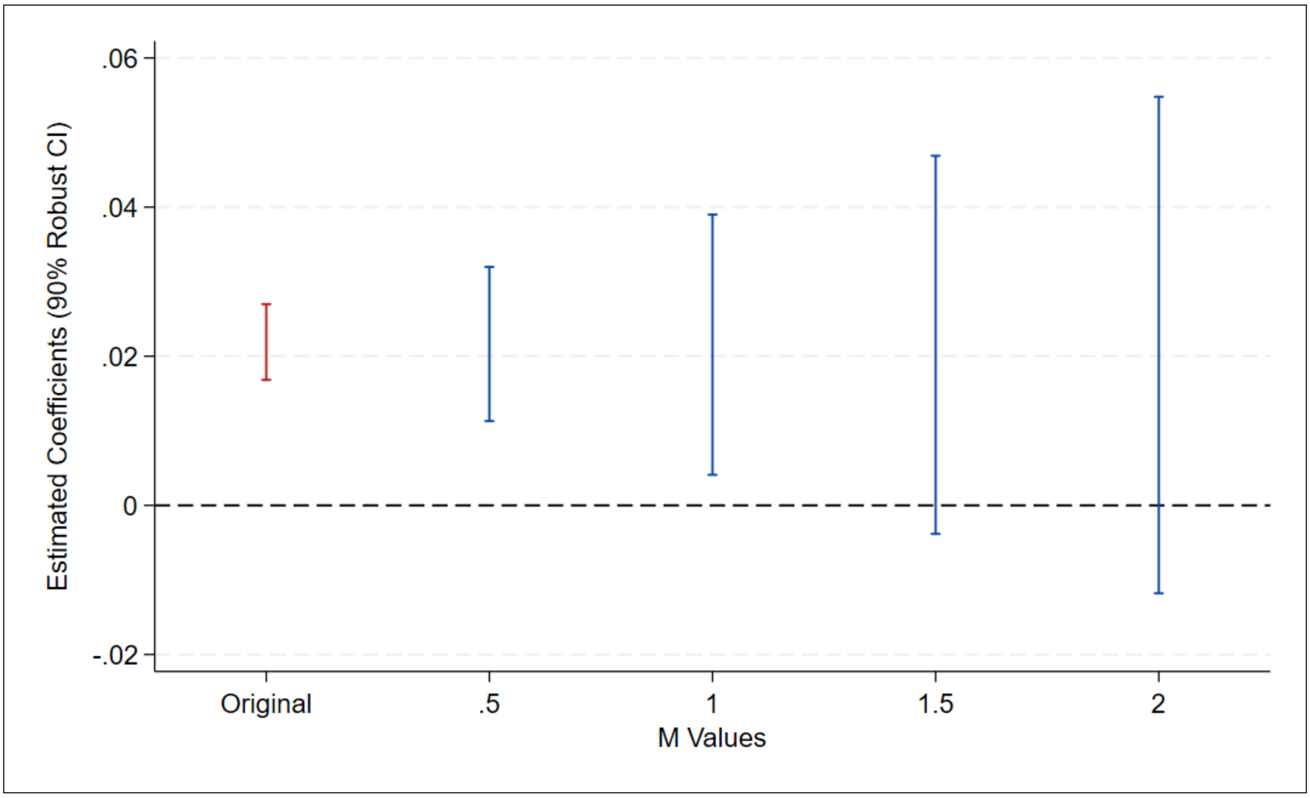
pre-treatment deviation. Estimated intervals shown in Figure 3 indicate that the main effects are reasonably robust to proportional post-treatment violations up to the scale of the largest pre-treatment deviation, i.e. $\bar{M} \leq 1$: under this benchmark, the robust confidence intervals remain strictly above zero. For larger allowances ($\bar{M} > 1$), intervals naturally widen and include zero. Overall, the evidence suggests moderate robustness to deviations from strict parallel trends.

Last but not least, we also explore possible heterogeneous effects by economic sector. Indeed, because Italy's HSR network carries passengers rather than freight, the main channels operate through faster business travel, wider commuter and client catchments, tourism flows, and face-to-face knowledge exchange. Larger gains are therefore expected in knowledge intensive business services (for example professional, ICT, and finance), in manufacturing segments that rely on design, coordination and sales, and in hospitality and cultural services through tourism. Freight oriented activities such as logistics, warehousing and heavy industry should be less directly affected. Construction may display transitory responses linked to ancillary investments. Agriculture and public administration are expected to be less sensitive. Retail trade is theoretically ambiguous, reflecting the balance between market enlargement from additional day trip consumers and potential expenditure reallocation toward larger hubs.

Table 3 provides results of this analysis.[26] As expected, disaggregating GVA by broad NACE sections

[26] Data on GVA by economic sector are not consistently available over the full sample period for all measures; in several

Figure 3: Result of the sensitivity analysis to violation of the parallel trend assumption for the estimation of the impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA.



Notes: *HonestDiD* robust confidence intervals [Rambachan and Roth, 2023, Bravo et al., 2024] for event–study ATT estimates. The procedure uses pre-treatment coefficients to bound possible post-treatment deviations from parallel trends; the sensitivity parameter $\bar{M}$ caps post-treatment violations relative to the largest pre-treatment deviation. We report intervals for $\bar{M} \in \{0.5, 1, 1.5, 2\}$. Intervals excluding zero across $\bar{M}$ indicate robustness.

reveals a reallocation towards tradables and knowledge-intensive activities.[27] The effect is positive and significant in B–E (industry, including manufacturing and utilities; +13.4%, $p \leq 0.05$), in K–N (business services; +9.3%, $p \leq 0.01$), and in R–U (other services; +12.9%, $p \leq 0.05$). By contrast, construction (F) declines (−31%, $p \leq 0.01$), and consumer-facing sectors G–J show a modest contraction (−7.5%, $p \leq 0.05$). Estimates for agriculture (A), real estate (L), professional and administrative services (M–N), and public/health/education (O–Q) are not precisely estimated and not different from zero. Together, these patterns align with a compositional shift towards higher-productivity and tradable activities. Indeed, as aforementioned, HSR intervention lowers the generalised cost of moving people rather than goods, so the main margins of adjustment involve information exchange, client acquisition, labour market integration and leisure mobility. Sectors whose production processes yield high returns to face-to-face interaction are therefore most responsive. Business and market services benefit through denser meeting technology, faster deal cycles and improved access to specialised labour; manufacturing segments that rely on design, prototyping, after sales assistance and supplier audits gain from the wider reach of engineers and sales staff; cultural and hospitality activities expand with additional day visitors and short breaks. By contrast, freight oriented activities and agriculture receive little direct cost relief, so effects are muted; construction can contract after opening once one off works conclude and as capital and

cases sectoral series begins only in 1995, leaving earlier years missing. For this reason, estimates in columns from (4) to (9) rely on fewer observations.

[27]NACE Rev. 2 - Statistical classification of economic activities. See `https://ec.europa.eu/eurostat/web/products-manuals-and-guidelines/-/ks-ra-07-015`.

Table 3: DiD estimates of the heterogeneous impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA by industry.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Dependent variable: regional Gross Value Added ($\log GVA_{it}$) by industry	A	B-E	F	G-J	K-N	L	M-N	O-Q	R-U
$Treated_i \cdot Post_t$	0.188	0.126**	-0.372***	-0.078**	0.089***	0.052	0.028	0.009	0.122**
	[0.151]	[0.063]	[0.099]	[0.037]	[0.034]	[0.058]	[0.040]	[0.027]	[0.050]
NUTS-3 region FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
NUTS-1-by-Time FE	✓	✓	✓	✓	✓	✓	✓	✓	✓
Population CV	✓	✓	✓	✓	✓	✓	✓	✓	✓
Labour & Productivity CV	✓	✓	✓	✓	✓	✓	✓	✓	✓
Capital CV	✓	✓	✓	✓	✓	✓	✓	✓	✓
Innovation CV	✓	✓	✓	✓	✓	✓	✓	✓	✓
Economic Size CV	✓	✓	✓	✓	✓	✓	✓	✓	✓
Institutional Quality Index	✓	✓	✓	✓	✓	✓	✓	✓	✓
Mult. Access Index	✓	✓	✓	✓	✓	✓	✓	✓	✓
Observations	4,092	4,091	4,092	2,697	2,697	2,697	2,697	2,697	2,697

Notes: This Table presents DiD estimates of Equation 2 on the effect of the opening of Reggio Emilia MedioPadana HSR station on regional GVA by industry. All specifications include both NUTS-3 region and time fixed effects, as well as NUTS-1-region-by-Year ones. Controls include predetermined regional characteristics at the beginning of the sample period: demographic factors (population), labour market and productivity indicators (labour productivity per hour, wage and salary earners, employment, nominal unit labour costs, compensation per hour worked, hours worked), capital endowment (capital stock), innovation capacity (Patent fractional counts, number of inventors patenting in the region), economic structure (GDP, GVA, share of industry on total regional economy), institutional quality (as in Nifo and Vecchione [2014]), and regional potential multimodal accessibility. Sector A refers to Agriculture, forestry and fishing; Sectors from B to E refer to Mining and quarrying, Manufacturing, Electricity, gas, steam and air conditioning supply, Water supply, sewerage, waste management and remediation activities; Sector F refers to Construction; Sectors from G to J refer to Wholesale and retail trade, repair of motor vehicles and motorcycles, Transportation and storage, Accommodation and food service activities, Information and communication; Sectors from K to N refer to Financial and insurance activities, Real estate activities (detail on sector L), Professional, scientific and technical activities, Administrative and support service activities (detail on sectors M-N); Sectors from O to Q refer to Public administration and defence, compulsory social security, Education, Human health and social work activities; Sectors from R to U refer to Arts, entertainment and recreation, Other service activities, Activities of households as employers, undifferentiated goods- and services-producing activities of households for own use, Activities of extraterritorial organisations and bodies. All specifications are estimated by OLS. Robust standard errors, clustered at NUTS-3 region level, are shown in parentheses: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

labour reallocate towards higher productivity uses; retail may lose local spending to larger hubs that become reachable within the same travel time budget, while any positive footfall effects can be offset by e-commerce and format competition. Real estate and parts of the public sector may chiefly capitalise benefits into prices or service quality rather than measured value added, which tempers short run GVA responses.

It is worth noting that DiD estimates on GVA summarise the net impact of improved passenger connectivity, but they do not reveal through which channels the gains arise. To open this black box, we re–estimate the more extended specification in Equation 2 on a set of plausibly mediating outcomes that map to the main economic mechanisms suggested by HSR: (i) productivity upgrading and capital–deepening (proxied by labour productivity per hour, capital per hour worked); (ii) skill and rent sharing (compensation per hour worked); (iii) knowledge diffusion and innovation output (forward–citation–weighted patent fractional counts, unweighted patent counts, number of inventors); and (iv) the extensive margin of labour input (hours worked). Under this framework, positive effects on productivity and wages would be consistent with better task matching and rent sharing, while increases in citation–weighted patenting—possibly larger than changes in inventor counts—would indicate improvements in the quality of innovation driven by denser face–to–face interactions and market access rather than mere inventor headcount growth. All regressions retain the same fixed effects and control structure as the main specification, so that estimates are directly comparable with baseline ones.

Table 4 reports Equation 2's estimates on these mediating channels, revealing (i) productivity gains and scale expansion rather than capital deepening, and (ii) an upgrading in the quality of innovation. Labour productivity per hour rises by $\approx 2.6\%$ ($p \leq 0.10$) and compensation per hour worked by $\approx 1.8\%$ ($p \leq 0.05$), while hours worked increase by $\approx 6.4\%$ ($p \leq 0.05$). By contrast, the effect on capital per hour worked is positive but not statistically significant $\approx 4.6\%$ ($p > 0.10$). On the innovation margin, forward–citation–weighted patent fractional counts increase by ≈ 84% ($p \leq 0.01$) and unweighted patent fractional counts by ≈ 34.4% ($p \leq 0.10$); conversely, the estimated coefficient relative to the number of inventors is not statistically different from zero. The ≈ 2.6% gain in productivity plus the ≈ 6.4% increase in hours worked roughly aligns with the (≈ 9.4% regional GVA effect, consistent with gains driven mainly by the extensive margin, complemented by some productivity improvements and by a marked increase in the quality (and quantity) of inventive output rather than increases in inventors headcounts.

Table 4: DiD estimates of the impact of the opening of Reggio Emilia MedioPadana HSR station on proximate channels.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Dependent Variables (log)	Labour productivity per hour	Compensation per hour worked	Hours worked	Capital per hour worked	Unweightet Patent Fractional Count	Forward-citations-weighted Patent Fractional Count	Number of inventors
$Treated_i \cdot Post_t$	0.026*	0.018**	0.062**	0.045	0.296*	0.610***	0.104
	[0.016]	[0.008]	[0.026]	[0.032]	[0.160]	[0.192]	[0.173]
NUTS-3 region FE	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓
NUTS-1-by-Time FE	✓	✓	✓	✓	✓	✓	✓
Population CV	✓	✓	✓	✓	✓	✓	✓
Labour & Productivity CV	✓	✓	✓	✓	✓	✓	✓
Capital CV	✓	✓	✓	✓	✓	✓	✓
Innovation CV	✓	✓	✓	✓	✓	✓	✓
Economic Size CV	✓	✓	✓	✓	✓	✓	✓
Institutional Quality Index	✓	✓	✓	✓	✓	✓	✓
Mult. Access Index	✓	✓	✓	✓	✓	✓	✓
Observations	4,092	4,092	4,092	4,092	3,720	3,720	3,720

Notes: This Table presents DiD estimates of Equation 2 on the effect of the opening of Reggio Emilia MedioPadana HSR station on mediating channels. All specifications include both NUTS-3 region and time fixed effects, as well as NUTS-1-region-by-Year ones. Controls include predetermined regional characteristics at the beginning of the sample period: demographic factors (population), labour market and productivity indicators (labour productivity per hour, wage and salary earners, employment, nominal unit labour costs, compensation per hour worked, hours worked), capital endowment (capital stock), innovation capacity (Patent fractional counts, number of inventors patenting in the region), economic structure (GDP, GVA, share of industry on total regional economy), institutional quality (as in Nifo and Vecchione [2014]), and regional potential multimodal accessibility. A one has been added to all patent and inventor count variables before taking the log to include observations with values of zero. All specifications are estimated by OLS. Robust standard errors, clustered at NUTS-3 region level, are shown in parentheses: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

These results are consistent with a passenger–connectivity shock that enlarges market reach and compresses coordination frictions. Higher expected demand and lower uncertainty plausibly trigger expansion in labour input and capacity utilisation, with economically meaningful productivity gains; wage gains arise via improved matching and rent sharing. The absence of a statistically significant effects on capital per hour worked suggests no robust capital deepening in the short run; rather, the evidence points to scale effects. On the innovation margin, denser face–to–face interaction with customers, suppliers and knowledge partners might foster problem solving, design refinement and the diffusion of tacit know–how.[28] Crucially, the latter translate into private returns only where firms possess sufficient absorptive capacity—pre–existing R&D, skilled human capital, and organisational routines that enable the identification, assimilation and exploitation of external knowledge—and where appropriability conditions (e.g., complementary assets, speed to market, secrecy, and formal IP protection) allow innovators to capture a larger share of the surplus. This combination makes it credible that citation–weighted patent output rises more strongly than inventor headcounts: connectivity improves exposure to valuable ideas, while absorptive capacity turns exposure into higher–quality inventions and appropriability sustains the incentive to invest in them. The delayed onset of effects is also consistent with the time needed for organisational change and capability building required to absorb and appropriate new knowledge flows.

[28] These results align with those in Bottasso et al. [2025b].

3.1 Robustness

This section discusses the empirical results obtained through an in-depth robustness analysis.

First, we re-estimate the more extended specification in Equation 2, dropping the NUTS-1-by-year fixed effects while retaining province and year fixed effects. This follows the rationale—already suggested and illustrated in column (1) of Table 2 without controls—that identification comes from variation between treated and untreated provinces within the same NUTS-1 macro-region and year; hence it is informative to probe sensitivity to removing these interactions. The estimates, reported in column (1) of Table 5, closely align with the baseline and confirm our main findings.

Table 5: Robustness tests and placebo inference.

	(1)	(2)	(3)	(4)	(5)
	No NUTS-1 by Time FEs	SDID	Donuts Hole	*Fake* Treatment	*Fake* Outcome
Dependent Variables	Regional Gross Value Added ($\log GVA_{it}$)				*Fake* regional Gross Value Added ($\log GVA_{it}^{fake}$)
$Treated_i \cdot Post_t$	0.099***	0.068*	0.116***		0.011
	[0.023]	[0.037]	[0.028]		[0.025]
$Treated_i^{fake} \cdot Post_t^{fake}$				-0.001	
				[0.001]	
NUTS-3 region FE	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓
NUTS-1-by-Time FE	✗	✓	✓	✓	✓
Population CV	✓	✓	✓	✓	✓
Labour & Productivity CV	✓	✓	✓	✓	✓
Capital CV	✓	✓	✓	✓	✓
Innovation CV	✓	✓	✓	✓	✓
Economic Size CV	✓	✓	✓	✓	✓
Institutional Quality Index	✓	✓	✓	✓	✓
Mult. Access Index	✓	✓	✓	✓	✓
Observations	4,092	4,092	3,872	4,092	4,092

Notes: This Table presents robustness/sensitivity analyses on the effect of the opening of Reggio Emilia MedioPadana HSR station. All specifications include both NUTS-3 region and time fixed effects, as well as NUTS-1-region-by-Year ones (the latter are not included in the specification of column 1). Controls include predetermined regional characteristics at the beginning of the sample period: demographic factors (population), labour market and productivity indicators (labour productivity per hour, wage and salary earners, employment, nominal unit labour costs, compensation per hour worked, hours worked), capital endowment (capital stock), innovation capacity (Patent fractional counts, number of inventors patenting in the region), economic structure (GDP, GVA, share of industry on total regional economy), institutional quality (as in Nifo and Vecchione [2014]), and regional potential multimodal accessibility. To assess sensitivity, column (1) shows the results from the estimation of the more extended version of Equation 2 without NUTS-1-by-Time fixed effects; column (2) provides Synthetic DiD estimates [Arkhangelsky et al., 2021]; column (3) shows a standard "donut hole" exercise, i.e. an analysis in which first-order contiguous NUTS-3 regions are excluded from the control group, in order to check for possible displacement/spillover effects; column (4) provides placebo estimates by replacing the treatment interaction with a *fake* one, where it is drawn from Bernoulli distributions with parameters t (probability of success), derived from the original sample distributions; column (5) shows results of another placebo test, in which the original dependent variable is replaced with a *fake* one that is drawn from random GVA distributions resembling sample ones (same mean and variance). Robust standard errors, clustered at NUTS-3 region level, are shown in parentheses: ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Second, our design features a single treated region, namely Reggio Emilia, compared with a large set of control ones, potentially differing in levels, trends, composition and exposure to idiosyncratic shocks. These differences need not undermine the design; indeed, while DiD frameworks remain informative as long as parallel trends assumption holds, a single treated unit may raise standard concerns about comparability and inference: estimates could be unduly influenced by the choice of the control group, by imperfect pre-treatment alignment, or by region-specific shocks. These are common issues in place-based evaluations and do not undermine the exercise; rather, they motivate complementary estimators that make the comparison more transparent and local to the most comparable controls. We therefore implement the Synthetic Control Method (SCM), which constructs a convex combination of donors that closely reproduces the treated region's pre-treatment path, and Synthetic DiD (SDID), which learns unit and time weights and then estimates a weighted two-way fixed-effects regression. Together, SCM and SDID provide a stringent pre-trend match and focus the comparison on relevant donors and periods, thereby strengthening the credibility of our research design without relying solely on the uniform-weight DiD benchmark.[29]

Table 6 shows region weights (left panel) and predictors balance (right panel) for the SCM analysis. Specifically, GVA trend in Reggio Emilia region, prior to the opening of the HSR station, is best reproduced by a combination of 6 Italian regions, those to which the SCM delivers positive critical weights. In addition, in the right panel of Table 6, both the GVA values and predictor variables of the treated region and its synthetic counterfactual (over the 34 years before the opening of the HSR station) are provided; comfortingly, the synthetic Reggio Emilia region closely mimics the real one both in terms of GVA and in other predictor variables, thus confirming the goodness of SCM's matching properties.[30] Table 7 depicts the magnitude of the impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA for the whole post-treatment period (2013-2023) leveraging the SCM procedure, while Figure 4 graphically compares trends of the treated region and its synthetic counterfactual.[31]

Table 6: SCM analysis. Predictors balance and region weights.

Donor Region	SCM Weight	Predictors Balance	Reggio Emilia region	Synthetic Reggio Emilia region
Aosta	0.060	(log) population	13.005	13.037
Imperia	0.034	(log) labour productivity per hour	3.218	3.208
Savona	0.094	(log) wage and salary earners	5.000	5.016
Trento	0.321	(log) employment	5.351	5.353
Modena	0.486	(log) unit labour costs	-0.986	-1.037
Siena	0.005	(log) compensation per hour worked	2.717	2.662
		(log) hours worked	12.893	12.903
		(log) capital stock	10.708	10.692
		(log) unweighted patent fractional count	3.499	3.152
		(log) number of inventors	3.743	3.520
		(log) labour productivity per worker	10.760	10.758
		(log) GDP	9.203	9.203
		(log) GVA	9.113	9.112
		(log) GVA (four-year average 1980-2008)	9.037	9.033
		(log) GVA (two-year average 2009-2013)	9.634	9.624

Notes: Predictors balance and region weights for the specification that analyses the impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA. The SCM assigns critical weights in order to build a synthetic control that minimizes the distance from the treated region in terms of innovative capacity and predictors of its subsequent growth. Such predictors are chosen in order to minimize the RMSPE ($RMSPE = 0.022$). A one has been added to all patent and inventor count variables before taking the log to include observations with values of zero.

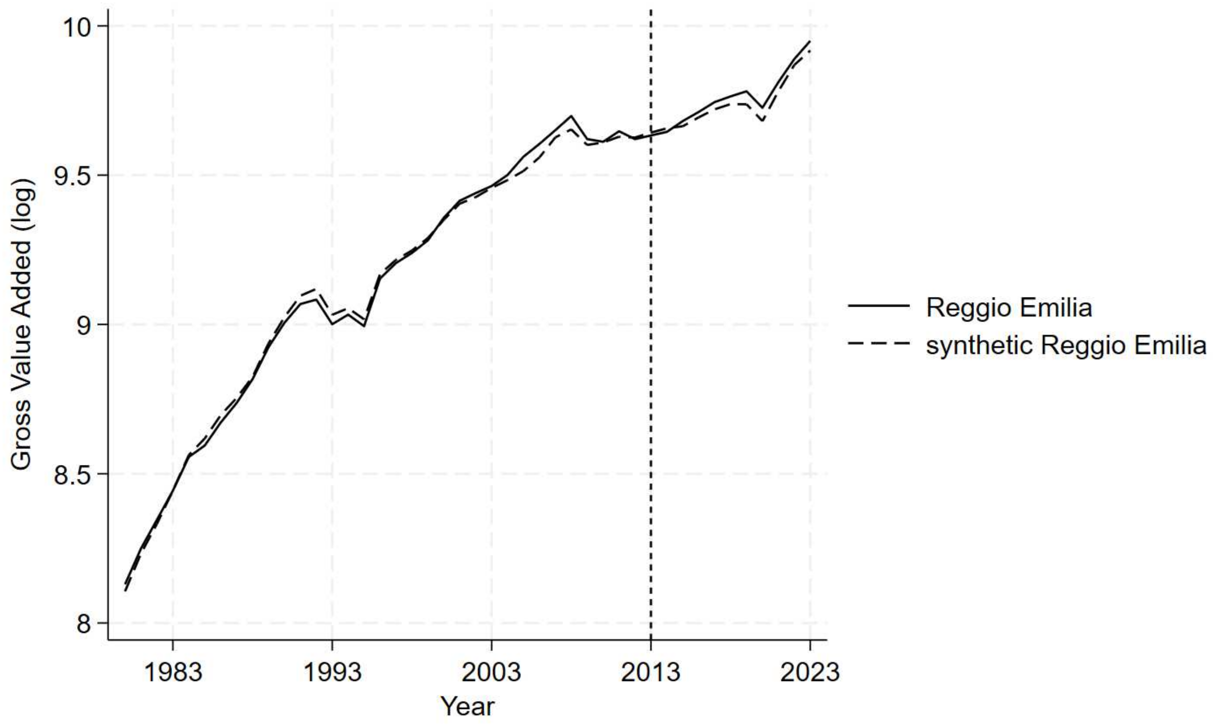
The joint analysis of Figure 4 and Table 7 suggests that, on average, the opening of Reggio Emilia HSR station has positively impacted on regional GVA by about 2% every year, with gaps ranging between +1.76% and +4.57%. Results from Figure 4 suggest instead that the synthetic control traject-

[29] In order to minimize mean squared prediction errors (MSPE), the SCM method has been estimated by relying on a constrained quadratic programming routine that finds the best fitting W-weights conditional on the regression based V-matrix. In particular, it embarks on an fully nested optimization procedure that searches among all (diagonal) positive semidefinite V-matrices and sets of W-weights for the best fitting convex combination of the control units.

[30] It is worth noting that the SCM method prevents the estimation of "extreme counterfactuals", namely those that fall far outside the convex hull of the data [King and Zeng, 2006].

[31] This approach is particularly useful when only a single region receives the treatment. Simple comparisons with neighbouring areas can be biased if those areas differ along unobserved dimensions—geographical, social, political, or economic. Moreover, the small number of treated groups undermines the large-sample approximations that standard inference relies upon [Conley and Taber, 2011].

Figure 4: Impact of the opening of Reggio Emilia MedioPadana HSR station on regional GVA. Trends for Genoa and synthetic control.



Notes: Regional GVA of the treated region (Reggio Emilia) and its synthetic counterfactual. The weights used to build the synthetic control and the predictors balance are shown in Table 6.

Table 7: SCM analysis. Effect estimates.

Year	(log) GVA - Treated	(log) GVA - Synthetic	Difference %
2013	9.6330423	9.6426232	-0.95%
2014	9.6449308	9.6569049	-1.19%
2015	9.6812687	9.6638668	1.76%
2016	9.7117214	9.6935774	1.83%
2017	9.7448139	9.7205307	2.46%
2018	9.7637367	9.7381650	2.59%
2019	9.7808170	9.7374944	4.43%
2020	9.7257471	9.6810175	4.57%
2021	9.8118439	9.7810459	3.13%
2022	9.8889065	9.8693734	1.97%
2023	9.9499035	9.9174156	3.30%

Notes: Regional GVA of the treated region (Reggio Emilia) and its synthetic counterfactual. The weights used to build the synthetic control and the predictors balance are shown in Table 6. Both series are natural logarithms. Let $\Delta_t = GVA_t^{\text{treated}} - GVA_t^{\text{synthetic}}$. Since $\Delta_t = \log(GVA_t^{\text{treated}}/GVA_t^{\text{synthetic}})$, the exact percentage difference in levels is $\%\text{Diff}_t = 100(GVA_t^{\text{treated}}/GVA_t^{\text{synthetic}} - 1) = 100(e^{\Delta_t} - 1)$. This measure is scale-invariant and directly interpretable as the percentage by which the treated level exceeds the synthetic level at time t.

ory closely matches Reggio Emilia's pre-2013 path while, after 2 years from the 2013 opening, regional GVA trend of the treated region and the synthetic control start to significantly diverge, with a sudden increase of Reggio Emilia with respect to its synthetic counterpart. What is more, as already suggested by the event study, also the SCM confirms that the treatment shows its effect after 2 years from the opening of the station and remain positive thereafter, thus suggesting a large impact of improved regional accessibility on regional economy.

Column (2) of Table 5 shows instead SDID results, highlighting the goodness of paper's main findings, being the estimated coefficient of the treatment interaction 0.068 ($p \leq 0.1$), namely a 7% increase in regional GVA after the intervention.[32]

Taken together, results from SCM and SDID estimates strongly alleviate concerns about the existence of one treated region and a large donor pool. The synthetic control results directly address these concerns: the synthetic trajectory closely matches Reggio Emilia's pre-2013 path and diverges only after the station opens, exactly when effects should materialise. This tight pre-treatment fit limits dependence on ad-hoc control selection and makes donor contributions transparent through weights. Complementing this, SDID learns unit and time weights to localise the comparison on the most comparable donors and periods, thereby weakening reliance on uniform parallel trends and reducing sensitivity to the composition of the control group. The concordance across SCM, SDID and DiD strengthens credibility despite the single treated unit, indicating that the main findings are not an artefact of donor choice or weighting.

Third, HSR may reshape economic activity beyond the treated province, either by diffusing gains to nearby areas (knowledge exchange, business travel, commuter flows) or by displacing activity from close competitors. To gauge these spatial equilibrium forces and to check that our counterfactual is not contaminated by neighbours' responses (possibly violating SUTVA assumption), we vary the composition of the control group by excluding first–order contiguous regions. Indeed, if nearby areas benefit, treating them as controls will attenuate the estimated ATT (their outcomes rise with the treated unit), whereas displacement would amplify it (controls fall relative to the treated unit). Comparing estimates across these nested control definitions therefore provides a simple diagnostic of the direction and geographic reach of spillovers, while also reassuring readers that the main results are not an artefact of a particular donor composition.

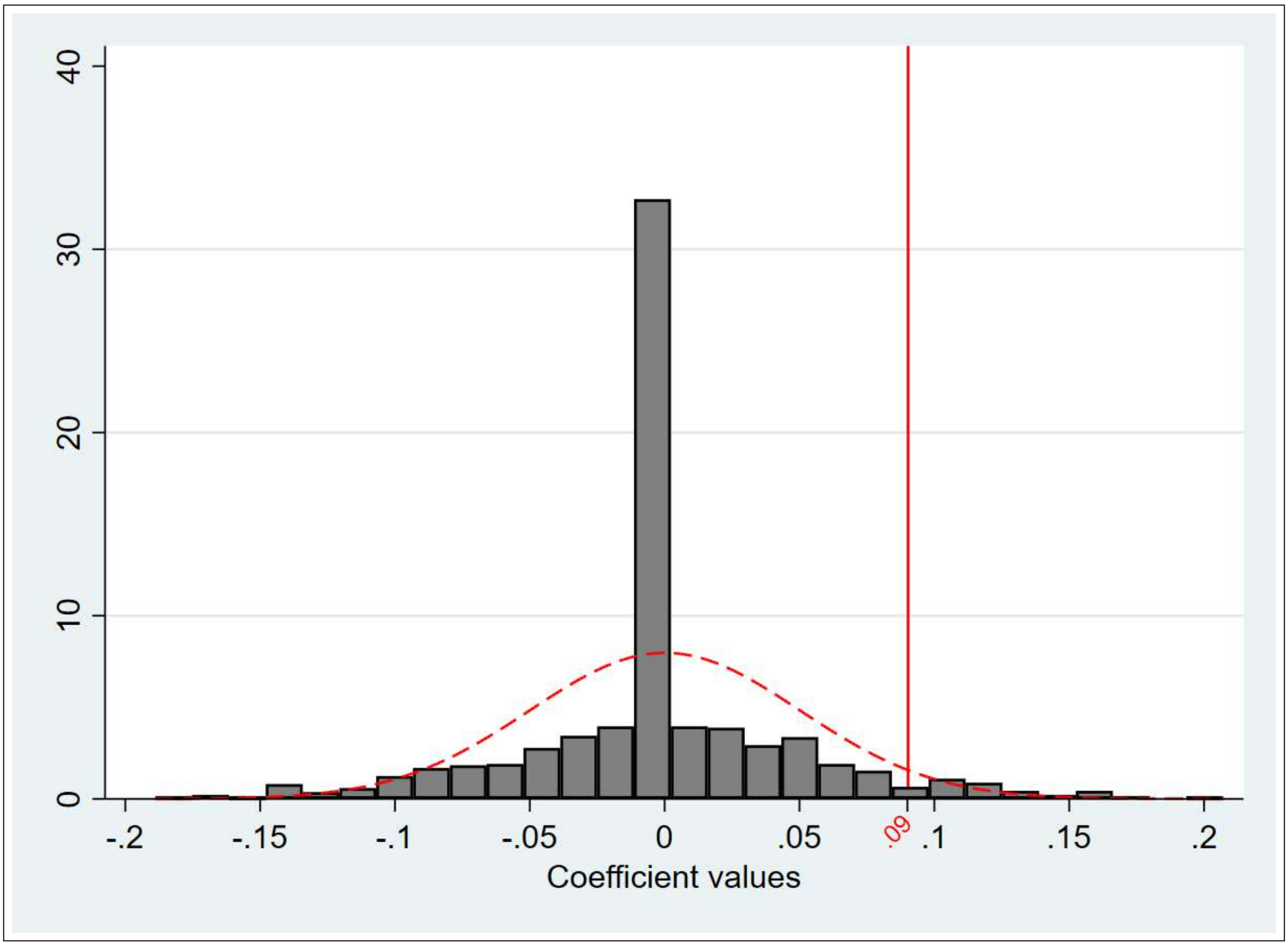
Column (3) of Table 5 shows results of this analysis. In this "donuts-hole" specification that excludes adjacent regions from the control group, the DiD estimate increases from 9.4% of the more extended version of Equation 2 (column 9 of Table 2) to 12.3% ($\approx +31\%$). This pattern is consistent with—though not definitive evidence of—positive spillovers to immediately neighbouring areas: if proximate regions gain via commuting, supplier–client linkages, business travel and knowledge exchange, including them among controls will attenuate the treated–control contrast, i.e. partly netting out gains in the baseline, so removing them yields a larger point estimate that is closer to the direct impact on the treated area. From the one hand, our province-level estimates naturally sit alongside firm-level evidence for Reggio Emilia in Bottasso et al. [2023], which show that firms' productivity gains from the HSR opening are highly localised—largest within 0–10 km, attenuating with distance, and statistically indistinguishable from zero in the 30–50 km band—without signs of displacement at intermediate distances. In our setting, adjacent provinces often contain municipalities that fall within these close-in buffers, so average effects measured at the NUTS-3 scale can reflect the same near-station improvements "bleeding" across administrative borders rather than medium-range spillovers.[33] On the other hand, our findings also

[32]As a sensitivity check, also the SDID model has been re-estimated without NUTS-1-by-Year FEs; comfortingly, results available upon request are unchanged.

[33]First-order contiguous NUTS-3 regions in our work are Parma, Modena, Mantova, Massa-Carrara and Lucca. It is worth noting that, if one considers the geodetic distance based on the centroids of these regions, the average range from Reggio

align with [Murata et al., 2014], which show that knowledge spillovers tend to be localised within 200 km and decay with distance, and that these flows often cross administrative boundaries—implying that designs tied tightly to borders can understate nearby spillovers. One should therefore interpret our 9.4% baseline result as a conservative equilibrium effect in a setting with possible spatially diffuse benefits, while 12.3% is closer to the own-region effect under the assumption of no interference within a two-step neighbourhood; policy assessments should accordingly account for cross-border spillovers when defining counterfactuals.

Figure 5: Placebo Plot Test. Random Treatment Allocation.



Notes: In this test the treatment is randomly assigned by generating simulated values for the treatment interaction $Treated_i \cdot Post_t$. In particular, a *fake* one has been built if random numbers drawn from a uniform distribution [0.1] are greater than the sample treatment probability. The paper then estimates Equation 2 after including the new *fake* treatment interaction, i.e. $Treated_i^{fake} \cdot Post_t^{fake}$, and then iterating such procedure 1,000 times in order to obtain a distribution of placebo coefficients to compare with the actual average estimated value from column (9) of Table 2 ($Treated_i \cdot Post_t = 0.09$). The rationale is that a statistically significant ATT should differ from that obtained with placebo estimates. Dark bars represent the distribution of estimated placebo coefficients. The vertical solid red line represents the actual estimated average treatment effect (0.09). The red dashed line fits a Normal distribution.

Last but not least, in order to validate paper's main findings, columns (4) and (5) of Table 5 provide results of placebo inference. On the one hand, we firstly estimate the more extended specification of Equation 2, including all predetermined control variables, after introducing both *fake* treated regions and *fake* treatment timing. The latter are drawn from Bernoulli distributions, with parameter t (namely the probability of success) derived from the $Treated_i$ and $Post_t$ original sampling distributions. In this

Emilia is about 50 km.

setting, one should not observe any significant effect of the *fake* treatment interaction on regional GVA; comfortingly, results in column (4) confirm this prediction. On the other hand, our main findings are confirmed when estimating Equation 2 after replacing the dependent variable with a *fake* one, where *fake* regional GVA is drawn from specific random distributions resembling sampling ones (same mean and variance). Once again, one should not observe any significant effect of the opening of Reggio Emilia Mediopadana HSR station and, comfortingly, results in columns (5) of Table 5 suggest that this is the case.

Finally, the paper performs a random allocation test by generating a placebo treatment variable taking the value of 1 if random numbers drawn from a uniform distribution [0,1] are higher than the sampling treatment probability.[34] The more extended specification of Equation 2 is then estimated by relying on the new *fake* treatment indicator and iterating this procedure 1,000 times in order to obtain a distribution of placebo average treatment effects to compare with the actual estimated average values shown in column (9) of Table 2 ($Treated_i \cdot Post_t = 0.09$). The rationale is that a statistically significant ATT should be different from that obtained with placebo estimates. Comfortingly, also this test, whose results are shown in Figure 5, confirms the robustness of our findings, given that the actual ATT falls well outside the core distribution of placebo estimates. What is more, the placebo coefficients are almost normally distributed, with estimates centred at zero, thus highlighting no treatment effects under the hypothesis of *fake* treatment assignments.

4 Conclusions and Policy Implications

This paper, leveraging a rich dataset of economic data for Italian NUTS-3 regions in 1980-2023 time-span and DiD, SCM and SDID research designs, provides causal evidence that integrating an intermediate Italian province into the HSR network generated a sustained and economically meaningful uplift in regional performance. The estimated post–treatment gain in GVA is close to one–tenth of a log point, robust to rich sets of controls, stringent fixed effects and alternative estimators. The dynamic profile of the estimated effect is consistent with gradual adjustment rather than an immediate jump, and the results survive robust inference that allows for bounded deviations from strict parallel trends. These findings indicate that the opening of the Reggio Emilia Mediopadana station translated improved passenger connectivity into persistent economic gains.

Beyond aggregate effects, the composition of growth is informative. Treated–control differences are concentrated in tradables and knowledge-intensive business services, while locally oriented sectors such as construction and retail contract—consistent with a shift towards higher-productivity activities that benefit most from face-to-face interaction, client acquisition and thicker labour markets. The mechanism analysis clarifies the channels: not capital deepening, but scale expansion and efficiency—hours worked rise ($\approx 6.4\%$), labour productivity increases ($\approx 2.6\%$) and compensation improves ($\approx 1.8\%$)—alongside an upgrading in innovative output, with large gains in forward-citation-weighted patent counts ($\approx 84\%$) and a rise in unweighted counts ($\approx 34\%$), without a commensurate increase in inventor headcounts. This pattern suggests that improved connectivity amplified knowledge flows that local firms were able to absorb and appropriate.

A further contribution concerns spatial general equilibrium effects. Estimates from a "donuts-hole" design show possible positive, localised spillovers, with adjacent areas that may benefit from the treatment. This pattern both supports a diffusion rather than displacement narrative. Complementary SCM

[34]This placebo analysis has recently been used successfully in Robbiano and Cerruti [2024], Bottasso et al. [2025a,c].

and SDID analyses, which tighten the pre–treatment match and localise identification via data–driven unit and time weights, corroborate the baseline DiD estimates and mitigate concerns raised by the single treated unit.

The paper shifts attention beyond core metropolitan hubs to a politically sited intermediate station, exploiting a plausibly quasi–random placement along a major corridor; this helps separate network–access effects from city–size dominance. Empirically, it leverages a long panel (1980–2023), a demanding fixed–effects structure (including NUTS–1–by–year effects), rich predetermined controls, explicit sectoral heterogeneity, and a mechanism battery that distinguishes the quality of innovation from its sheer volume—dimensions that are often absent or only partially addressed in prior work. These features jointly enhance internal validity and speak directly to external relevance for regions with similar endowments and institutional readiness.

As a single, well–identified case study, external validity should be assessed with care. Future work could replicate the design across other intermediate nodes, study longer–run distributional and welfare consequences (including land and housing markets), and explore firm–level adjustment margins (organisation, supply chains, export dynamics) under alternative identification strategies. Extending the spatial analysis with fine–grained travel–time and service–frequency data would further clarify the decay of spillovers and inform the design of complementary local transport.

Both paper's findings and the evolution in the design of the Reggio Emilia MedioPadana station argue for a place–sensitive infrastructure strategy in which intermediate territories are not bypassed but enabled to translate accessibility into productivity and innovation. Therefore, a broad policy lesson follows. An intermediate HSR station can deliver sizable and enduring local gains; however, the latter are likely to materialise when they are embedded in a coherent local development strategy. This includes investments in last-mile connectivity (e.g., local rail and bus services, active transport links), integrated land-use planning (e.g., zoning for commercial or industrial expansion near the station), business support mechanisms (e.g., incentives for innovation and scale-up), and vocational education and training policies that align with emerging sectoral specialisations. Indeed, Reggio Emilia Mediopadana HSR station illustrates how complementary measures turn an HSR stop into an economic gateway. The latter was purposefully sited about 4 km north of the city, parallel to the A1 motorway, and delivered in June 2013 as the only high-speed stop on the Milan–Bologna trunk—an architecture–infrastructure bundle conceived with strong highway access (Calatrava's three bridges were opened in 2007) rather than dense urban embedment at the outset. Within months, last-mile rail connectivity was layered in via the Reggio Emilia–Guastalla regional railway line, which links the HSR station to the central historical station and city network. From 2020 onward, the municipality and rail manager advanced an intermodal-hub programme (new parking, access and under-viaduct services), culminating—inter alia—in a new east-side atrium opened in November 2024. Public-transport and active-mobility ties have also been upgraded: in February 2024 a limited-stop Minibus "M" began providing a frequent, direct service between the HSR station, the historic centre and the central station; and the "Superciclabile Nord" corridor is being completed to extend high-quality cycle access from MedioPadana station toward Bagnolo in Piano (works scheduled to finish in autumn 2025). In parallel, land-use interventions in the adjoining Mancasale industrial area—most notably the upgrading of Via Filangieri as a strategic link between the HSR station, the business park and the A1 motorway interchange—have reinforced the station's role as a functional gateway to production districts rather than a stand-alone transport asset. Read through the lens of our results, this sequence of complementary measures (PT services, intermodality, and targeted place-making) helps explain why intermediate-region HSR stations can generate benefits that reach beyond the immediate vicinity when they are embedded in a coherent local development strategy; absent such layering, impacts risk remaining spatially narrow and hard to aggregate (see `https://www.reggioemiliawelcome.it/en/reggio-emilia/how-to-get-there/transports/by-train/high-speed-av-mediopadana-railway-station?`). Consistent with transport-economics evidence, such accessibility upgrades tend to translate into wider economic benefits where demand is sufficient and local policy complements are in place (land-use coordination, business support, skills), rather than from infrastructure alone. Without such complementary measures, the benefits of HSR risk remaining spatially and economically limited.

Overall, our results advocate for a more balanced and place-sensitive infrastructure strategy, in which intermediate cities and regions are not simply bypassed, but actively empowered to leverage improved accessibility into sustained gains in productivity and innovation. For countries seeking to reduce regional disparities while supporting industrial upgrading, such integration of transport investment and regional policy offers a promising way forward.

References

A. Abadie, A. Diamond, and J. Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015. doi: https://doi.org/10.1111/ajps.12116. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12116`.

G. M. Ahlfeldt and A. Feddersen. From periphery to core: measuring agglomeration effects using high-speed rail. *Journal of Economic Geography*, 18(2):355–390, 2018.

D. Albalate and X. Fageda. High-technology employment and transportation: evidence from the European regions. *Regional Studies*, 50(9):1564–1578, 2016.

D. Arkhangelsky, S. Athey, D. A. Hirshberg, G. W. Imbens, and S. Wager. Synthetic difference-in-differences. *American Economic Review*, 111(12):4088–4118, 2021.

A. Baltrunaite and E. Karmaziene. Trainspotting: board appointments in private firms. *Bank of Italy Temi di Discussione (Working Paper) No*, 1278, 2020.

A. B. Bernard, A. Moxnes, and Y. U. Saito. Production networks, geography, and firm performance. *Journal of Political Economy*, 127(2):639–688, 2019.

C. Blanquart and M. Koning. The local economic impacts of high-speed railways: theories and facts. *European Transport Research Review*, 9(2):12, 2017.

A. Bottasso, M. Conti, A. R. Ferrara, and S. Robbiano. High-speed railways and firms total factor productivity: evidence from a quasi-natural experiment. In *IZA Discussion Papers Series*, volume 16572, pages 1–27. IZA Institute of Labor Economics, 2023.

A. Bottasso, M. Conti, A. Fiduccia, S. Robbiano, and M. Tartaglia. Does the opening of a high-speed-rail station impact house prices? Evidence from a quasi-natural experiment: Reggio Emilia Mediopadana. In F. Pagliara, editor, *Socioeconomic Impacts of High-Speed Rail Systems*, pages 319–341, Cham, 2025a. Springer Nature Switzerland. ISBN 978-3-031-82528-6.

A. Bottasso, F. Mij, and S. Robbiano. The innovation dividend of high speed rail: evidence from Italy. Available at SSRN - 5590932, 2025b.

A. Bottasso, S. Robbiano, and P. Marocco. Price matching in online retail. *Economic Inquiry*, 63(1): 206–235, 2025c. doi: https://doi.org/10.1111/ecin.13255. URL `https://onlinelibrary.wiley.com/doi/abs/10.1111/ecin.13255`.

M. C. Bravo, J. Roth, and A. Rambachan. Honestdid: Stata module implementing the honestdid r package. Boston College Department of Economics, 2024.

R. Bronzini and P. Piselli. The impact of R&D subsidies on firm innovation. *Research Policy*, 45(2): 442–457, 2016.

F. Chen and Z. Chen. High-speed rail and happiness. *Transportation Research Part A: Policy and Practice*, 170:103635, 2023.

T. G. Conley and C. R. Taber. Inference with "difference in differences" with a small number of policy changes. *Review of Economics and Statistics*, 93(1):113 – 125, 2011. doi: 10.1162/REST_a_00049. URL `https://www.scopus.com/inward/record.uri?eid=2-s2.0-79957504319&doi=10.1162%2fREST_a_00049&partnerID=40&md5=fda8a8c21072fb9c4d7b372d00a0aa5f`.

D. Di Matteo, I. Mariotti, and F. Rossi. Transport infrastructure and economic performance: an evaluation of the Milan-Bologna high-speed rail corridor. *Socio-Economic Planning Sciences*, 85:101304, 2023.

L. Duan, D. Niu, W. Sun, and S. Zheng. Transportation infrastructure and capital mobility: evidence from China's high-speed railways. *The Annals of Regional Science*, 67(3):617–648, 2021.

J. Fan, H. Kato, Z. Yang, and Y. Li. Effects from expanding high-speed railway network on regional accessibility and economic productivity in China. *Transportation Research Record*, 2676(4):145–160, 2022.

Q. Feng, Z. Chen, C. Cheng, and H. Chang. Impact of high-speed rail on high-skilled labor mobility in China. *Transport Policy*, 133:64–74, 2023.

C. Ferrari, A. Bottasso, M. Conti, and A. Tei. *Economic role of transport infrastructure: theory and models*. Elsevier, 2018.

S. Freyaldenhoven, C. Hansen, and J. M. Shapiro. Pre-event trends in the panel event-study design. *American Economic Review*, 109(9):3307–3338, 2019.

B. H. Hall, A. Jaffe, and M. Trajtenberg. Market value and patent citations. *The RAND Journal of Economics*, 36(1):16–38, 2005. ISSN 07416261. URL `http://www.jstor.org/stable/1593752`.

M. Jin, K.-C. Lin, W. Shi, P. T. Lee, and K. X. Li. Impacts of high-speed railways on economic growth and disparity in China. *Transportation Research Part A: Policy and Practice*, 138:158–171, 2020.

X. Ke, H. Chen, Y. Hong, and C. Hsiao. Do China's high-speed-rail projects promote local economy? New evidence from a panel data approach. *China Economic Review*, 44:203–226, 2017.

G. King and L. Zeng. The dangers of extreme counterfactuals. *Political Analysis*, 14(2):131–159, 2006.

H. R. Koster, T. Tabuchi, and J.-F. Thisse. To be connected or not to be connected? The role of long-haul economies. *Journal of Economic Geography*, 22(4):711–753, 2022.

R. Le Boennec, J. Bulteau, and T. Feuillet. The role of commuter rail accessibility in the formation of residential land values: exploring spatial heterogeneity in peri-urban and remote areas. *The Annals of Regional Science*, 69(1):163–186, 2022.

G. Li, K. Pu, and M. Long. High-speed rail connectivity, space-time distance compression, and trans-regional tourism flows: evidence from China's inbound tourism. *Journal of Transport Geography*, 109:103592, 2023a.

Y. Li, Z. Chen, and P. Wang. Impact of high-speed rail on urban economic efficiency in China. *Transport Policy*, 97:220–231, 2020.

Z. Li, Q. Wang, M. Cai, and W.-K. Wong. Impacts of high-speed rail on the industrial developments of non-central cities in China. *Transport Policy*, 134:203–216, 2023b.

Y. Liu, D. Tang, T. Bu, and X. Wang. The spatial employment effect of high-speed railway: quasi-natural experimental evidence from China. *The Annals of Regional Science*, 69(2):333–359, 2022.

M. Momenitabar, R. Bridgelall, Z. Dehdari Ebrahimi, and M. Arani. Literature review of socioeconomic and environmental impacts of high-speed rail in the World. *Sustainability*, 13(21):12231, 2021.

A. Moyano and F. Dobruszkes. Mind the services! High-speed rail cities bypassed by high-speed trains. *Case Studies on Transport Policy*, 5(4):537–548, 2017.

Y. Murata, R. Nakajima, R. Okamoto, and R. Tamura. Localized knowledge spillovers and patent citations: a distance-based approach. *Review of Economics and Statistics*, 96(5):967–985, 2014.

D. Neumark and H. Simpson. Place-based policies. In *Handbook of Regional and Urban Economics*, volume 5, pages 1197–1287. Elsevier, 2015.

A. Nifo and G. Vecchione. Do institutions play a role in skilled migration? The case of Italy. *Regional Studies*, 48(10):1628–1649, 2014.

A. Pakes and Z. Griliches. Patents and R&D at the firm level: a first report. *Economics Letters*, 5(4): 377–381, 1980.

Y. Qin. 'No county left behind?' The distributional impact of high-speed rail upgrades in China. *Journal of Economic Geography*, 17(3):489–520, 2017.

A. Rambachan and J. Roth. A more credible approach to parallel trends. *Review of Economic Studies*, 90(5):2555–2591, 2023.

S. J. Redding and M. A. Turner. Transportation costs and the spatial organization of economic activity. *Handbook of Regional and Urban Economics*, 5:1339–1398, 2015.

S. Robbiano and G. Cerruti. Assessing the impact of price-matching guarantees and price fluctuations on consumer feedback: insights from the online consumer electronics market. ResearchSquare. DOI `https://doi.org/10.21203/rs.3.rs-4800672/v1`, 2024.

M. Sala. *High speed rail stations and secondary cities. Reggio Emilia AV Mediopadana*. PhD thesis, Politecnico di Milano, 2017.

X. Sun, S. Yan, T. Liu, and J. Wang. The impact of high-speed rail on urban economy: synergy with urban agglomeration policy. *Transport Policy*, 130:141–154, 2023.

M. Trajtenberg. A penny for your quotes: patent citations and the value of innovations. *The Rand Journal of Economics*, pages 172–187, 1990.

X. Wang, D. Tang, Y. Liu, and T. Bu. The impact of high-speed railway on labor market between the North and South: evidence from China. *The Annals of Regional Science*, 71(2):487–515, 2023.

Y. Wang, G. Cao, Y. Yan, and J. Wang. Does high-speed rail stimulate cross-city technological innovation collaboration? Evidence from China. *Transport Policy*, 116:119–131, 2022.

J. Wetwitoo and H. Kato. High-speed rail and regional economic productivity through agglomeration and network externality: a case study of inter-regional transportation in Japan. *Case Studies on Transport Policy*, 5(4):549–559, 2017.

J. M. Wooldridge. Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. *Empirical Economics*, pages 1–43, 2025. doi: https://doi.org/10.1007/s00181-025-02807-z.

S. Yoo, J. Kumagai, K. Kawasaki, S. Hong, B. Zhang, T. Shimamura, and S. Managi. Double-edged trains: economic outcomes and regional disparity of high-speed railways. *Transport Policy*, 133:120–133, 2023.

S. Yoo, J. Kumagai, and S. Managi. Urban-rural gap induced by high-speed rail: 35 years of evidence from Japan. *Research in Transportation Business & Management*, 55:101131, 2024.

X. Zhang and J. Gibson. Local economic effects of connecting to China's high-speed rail network: evidence from spatial econometric models. *The Annals of Regional Science*, 74(2):56, 2025.

L. Zheng and S. Wu. Remote high-speed rail stations, urban land supply, and the emergence of new economic activities. *Transportation Research Part A: Policy and Practice*, 189:104226, 2024.

Z. Zhou and A. Zhang. High-speed rail and industrial developments: evidence from house prices and city-level GDP in China. *Transportation Research Part A: Policy and Practice*, 149:98–113, 2021.