Have Higher Speeds Lowered Forest Impacts for Railways?

Construction & Operation of Considerable Chinese HSR Investment

Can Hu [a], Wumeng He [b*], Alexander Pfaff [c]

[a] *College of Land Management, Huazhong Agriculture University, China*
[b] *The Environmental Research Center, Duke Kunshan University, China*
[c] *Sanford School of Public Policy, Duke University, United States*

* Corresponding author. Wumeng He

Address: 8 Duke Avenue, Kunshan, Jiangsu, China. Email: wumeng.he@dukekunshan.edu.cn

Other Authors' emails: mshucan@outlook.com (C. Hu), alex.pfaff@duke.edu (A. Pfaff).

Acknowledgements

We thank Binbin Li for providing comparative data and helpful feedback. We are grateful to Jingbo Cui for comments on an early version of this work presented at a workshop at Duke Kunshan University. We also appreciate the valuable feedback from Roger von Haegen, Junjie Zhang, Nicolai Kuminoff, and other participants at a subsequent DKU workshop. Finally, we thank Anlu Zhang for useful suggestions and continued support throughout the development of this project.

Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Have Higher Speeds Lowered Forest Impacts for Railways?

Construction & Operation of Considerable Chinese HSR Investment

Abstract (125 words)

Transport infrastructure for moving people and goods is critical for economic development. As such investments also affect land use by raising access to remote areas, environmental tradeoffs may arise. We examine forest impacts for China's world-leading investments in high-speed rail (HSR), 2004-2022, exploiting HSR's exogeneity between the cities driving HSR routes. We find that construction raised forest loss by a little (one fifth of one percent per year), on average. Importantly, HSR operation — which does not increase local access along rail segments — did not raise loss on average. It did so only for sites with the highest prior infrastructure, while lowering forest loss for sites with no prior highways or railways. Targeted HSR siting with care during construction can yield high speeds with low impacts.

Keywords: High-speed rail; Infrastructure; Deforestation; China

1. Introduction

Transport infrastructure has long been recognized as an important driver of economic development and cause of forest losses, including within economically remote and ecologically sensitive regions (Araujo et al. 2023; Asher et al. 2020; Baehr et al. 2023; Barber et al. 2014; da Silva et al. 2023; Foster and Rosenzweig 2003; Souza-Rodrigues 2019). Most conceptual frameworks and empirical studies of transport focus on roads − ranging from secondary dirt tracks to highways. They rightly emphasize multiple roles for roads in facilitating frontier forest access for transport of agricultural outputs and inputs including labor (Busch and Ferretti-Gallon 2017; Kaczan 2020; Pfaff et al. 2018). In doing so, the existing literature (implicitly) treats increases in local accessibility as, essentially universally, the primary channel by which transport infrastructure affects forest outcomes – even if development contexts can affect the directions and magnitudes of those effects (Pfaff et al. 2018).

Rail infrastructure often is assumed to generate impacts similar to roads, by reducing transport costs and enabling agricultural expansion and urbanization (De Soyres et al. 2019; Losos et al. 2019). However, the local accessibility provided by rail is highly concentrated at stations. Rail segments between stations typically do not provide local entry or exit points (Vessali 1996; Wu et al. 2022). Thus, spatial concentration of access may imply that railways offer a more ecologically sensitive alternative for forest-rich regions, such as the Amazon (das Neves et al. 2021; Viana et al. 2008). Navalkar et al. (2023), e.g., have found zero net forest impact associated with the Konkan Railway.

Yet, it may be too optimistic to assume that, without other governance, railways leave surrounding forests undisturbed. Studies report substantial forest-cover loss along railway corridors (Santos et al. 2020; Xiao et al. 2024), with descriptive evidence (rather than causal analysis). For example, Xiao et al. (Xiao et al. 2024) report roughly a 3% forest loss along the China-Laos Railway since 2017, with a large share attributed to construction activities. Given the limited number of rigorous empirical studies, and the mixed nature of existing evidence, the forest impacts of rail infrastructure remain an important concern. Most fundamentally, the existing literature has not directly examined whether transport infrastructure will affect forests in the absence of improved local accessibility.

This paper addresses that question by using rail segments away from stations as a setting to examine transport infrastructure that does not directly increase local accessibility. Railways provide a natural opportunity to test whether accessibility is necessary for new transport to affect forest outcomes. Impacts along segments that lack access points would suggest the presence of alternative channels, while their absence would reinforce the centrality of accessibility (as in roads-focused literatures).

Among all types of rail systems, high-speed rail (HSR) is particularly well suited for this purpose. Operating at speed exceeding 200 km/hour (Blanco et al. 2011; Zheng et al. 2019), HSR networks feature relatively few stations and, outside station areas, virtually no opportunities for local access. Unlike common freight-oriented railways, HSR primarily transports passengers and is constructed expressly to connect major cities (Zheng et al. 2019). Thus, the non-urban areas traversed by HSR do not drive 'treatment', making segments between stations suitable for isolating transport effects.

We study China's high-speed rail network. It had an unprecedented expansion over the last 15 years (Niu et al. 2021) and, by the end of 2022, had exceeded 42,000 km, the largest in the world (Xinhua News 2023a). This followed a centrally planned blueprint (Fig. 1a), approved in 2016, to connect over 95% of cities with populations above 500,000 by raising HSR mileage to 50,000 km by 2025 (State Council of China 2021). As HSR's routes are primarily determined by inter-city connectivity, not the rural conditions they traverse, there exists plausibly exogenous exposure to HSR for forested regions between stations. We combined detailed data on HSR construction and operation with high-resolution land-use data back to 1995, i.e., well before the introduction of China's first HSR line.

While HSR routes can be influenced by geographic and environmental constraints such as terrain, water bodies, and protected areas (see Table A2 in Appendix A), HSR planning strongly prioritizes straight and direct alignments to meet engineering requirements and to minimize construction costs. We exploit this feature by using straight-line paths between HSR stations as instrumental variables that capture the underlying routing logic. Because stations' locations are influenced by pre-existing economic conditions, development plans, and land-use patterns — and because stations have been shown to affect economic activities (Zheng et al. 2019), land use (Wenner and Thierstein 2022), and land values (Chang and Murakami 2019) — we deliberately exclude areas near stations from our analysis. This exclusion is not only an identification strategy but also a design choice consistent with our focus upon transport-infrastructure segments that do not directly alter local accessibility.

Our distinct results for construction versus operations phases align with this conceptual framework. Pooling phases, we find new HSR drive a small average rise in deforestation (just under one tenth of a percent per year). Separating phases, though, reveals that this impact is driven by construction. On average, construction raises forest loss by about one fifth of a percent per year − and that is not purely mechanical clearing of the right-of-way (as impacts remain excluding 1km from new lines).

Importantly, however, operation of new HSR does not raise forest loss relative to background trends, on average. It does so only for the locations with the highest prior transport infrastructure (i.e., both railways and highways), while lowering forest loss for sites with no railway or highway in our data. These patterns are consistent with indirect channels not observed in our data set, such as temporary or permanent new access generated by HSR. Feeder roads cannot be systematically measured but they may well explain forest impact from HSR construction despite little new access from operation. Outmigration may explain lowered forest loss where new HSR offered increased urban connections.

We contribute in showing that HSR operations which lack local access do not generate forest loss, while HSR construction can do so through indirect disturbance. These results complement studies of railway impacts on specific habitats (Losos et al. 2019) and of broader forest transition dynamics (Asher et al. 2020; Baehr et al. 2023). We also document that both HSR construction and operations impacts vary across broad contexts that differ in pre-existing transport infrastructure. This informs future targeting, as countries invest more in transport infrastructure. China is expanding HSR not only domestically but also beyond its borders, within its Belt and Road Initiative (Wang et al. 2020). Our results guide balancing of economic investment with conservation and sustainable land use.

The remainder of this manuscript is as follows. Section 2 provides background concerning China's varied HSR and forest landscapes. Section 3 describes all of our data. Section 4 details our empirical approach. Section 5 provides our suite of results on HSR's forest impacts, then Section 6 concludes.

2. China's HSR & Forests

2.1 Mainland HSR Network

The Beijing-Tianjin Intercity Railway, in use since 2008 and with a design speed of 350 kilometers per hour, marked the dawn of China's HSR era, starting a substantial expansion of China's network. Building upon this initial effort, the State Council approved an ambitious HSR development plan in 2016. It envisioned eight horizontal (east-west) as well as eight vertical (north-south) corridors reaching across the country to form a nationwide grid − and, quite specifically, to connect most of China's major cities by 2025 (National Development and Reform Commission of China 2016).

The planned HSR network spans almost all of China's provinces, from the coast to the heartland, linking provincial capitals with prefectural cities and counties (Fig. 1a). Yet, regional imbalances in HSR's coverage clearly exist. The densely populated southeast has quite extensive HSR coverage, while the sparsely populated and geographically challenging northwest region has much less HSR. Generally, ongoing infrastructure advancements continue to drive HSR expansions within China.

China's HSR has distinct physical characteristics. First, routes are frequently elevated on viaducts, or routed through tunnels, via engineering efforts to reduce land acquisition, minimize resettlement, and preserve high-quality agricultural land (Lawrence et al. 2019; People's Daily 2019). Second, in contrast to many European systems, a large share of HSR stations in China are newly constructed and often located on urban peripheries rather than within historical city centers (Dong et al. 2021).

During construction, access roads are used to transport materials and equipment. These "feeder roads" may include newly built roads, widened local roads, or explicitly temporary roads (Fig. A1). While some may persist, depending on local decisions[1], they are not systematically documented in any available data. As a result, feeder roads cannot be directly incorporated into empirical analysis and are treated here only as contextual background for understanding construction-phase impacts.

In endeavoring to link cities across the country, in order to facilitate movements of people, China's HSR network traverses diverse terrain − including plains, mountains, plateaus, and arid regions. Experience with complex geological conditions, such as soft soil, permafrost, and extreme climates, has contributed to advances in HSR construction techniques and engineering capacities (Dreher et al. 2022). Taken together, the scale, routing logic, and construction practices of China's HSR system imply widespread exposures, for forested landscapes, to rail segments that do not directly provide local access − creating a distinctive setting for evaluating forest impact of transport infrastructure.

[1] The manager of one HSR design team indicated feeder roads were in fact planned to not only facilitate construction but also benefit communities. After HSR construction, the fates of temporary feeder roads often are left to local actors (residents, governments). If they find those roads useful, they may continue using them. Else, they may abandon them.

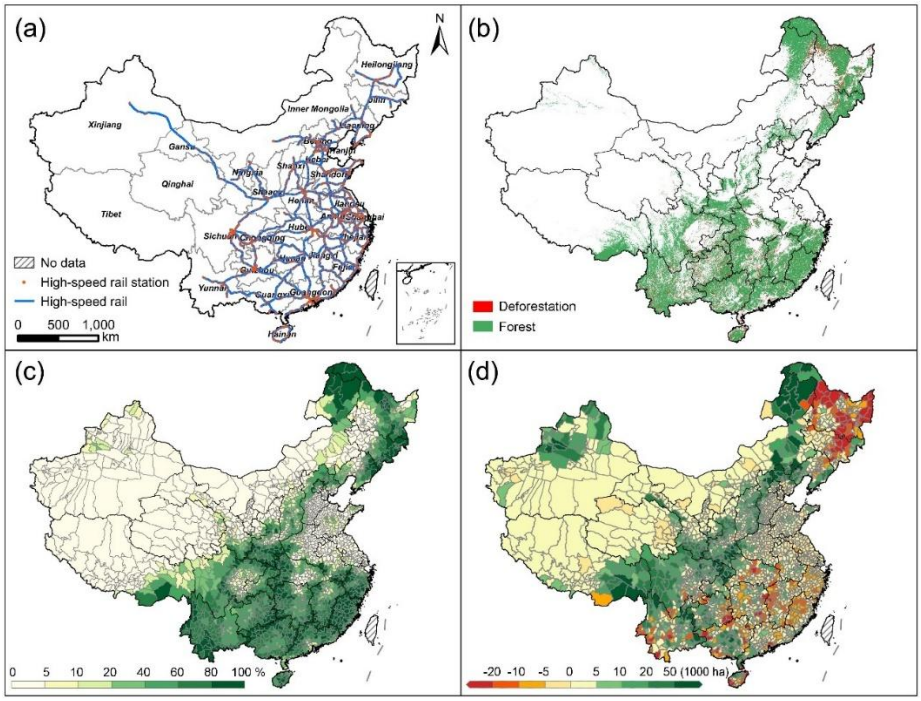


Fig. 1. HSR and forest in mainland China. Note: There are 31 provinces and 2843 counties in mainland China in 2022. (a) Distribution of HSR and their stations in 2022. (b) Forest cover in 1990 and deforestation between 1990 and 2022, 30-meter resolution. (c) Forest cover by county in 2022. (d) Forest area changes at county level between 1990 and 2022.

2.2 Forest Cover in Mainland China

2.2.1 Comparing Data Sets in terms of Temporal Forest-Cover Change

Multiple forest datasets are available for mainland China, differing in forest transition, spatial resolution, temporal coverage, and data sources. Table E1 summarizes key characteristics of widely used global and national datasets, including CLCD (Yang and Huang 2021), CATCD (Cai et al. 2024), National Forest Inventory of China (NFI), CNLUCC (Xu et al. 2018), GLC_FC30D (Zhang et al. 2023), GLAD (Potapov et al. 2022), ESA CCI-LC (ESA 2017), GFC (Hansen et al. 2013), and ESA WorldCover (Zanaga et al. 2021). Figure 2 compares forest cover across these sources.

Despite differing definitions and classifications, datasets exhibit broadly consistent temporal trends: most show a modest increase in forest cover over 2000-2020. While estimates of forest extent vary, spatial distributions are similar: forests are concentrated in Northeast China, southern mountainous regions, and parts of the southeastern coastal areas (Fig. E1). Differences across datasets primarily reflect their differences in forest definitions, spatial resolutions, and classification algorithms.

Among them, CLCD provides annual land-cover data at 30-meter resolution from 1990 to 2022 (Yang and Huang 2021), covering the full period of China's HSR expansion and offering substantial pre-treatment observations. With a reported forest classification accuracy of 85.49% ± 1.30% and extensive recent use (Hou et al. 2022; Xu et al. 2022), CLCD serves as our primary dataset. As a robustness check, we also make use of CATCD, which provides tree-cover data from 1985 to 2023.

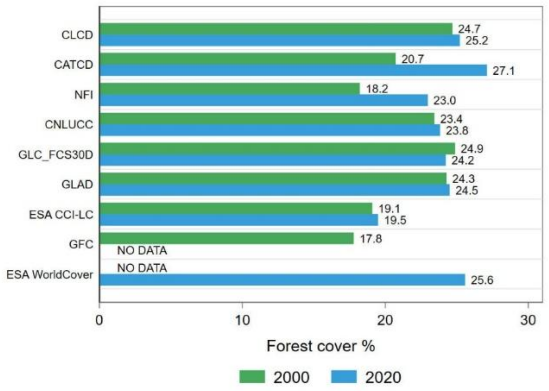


Fig. 2. Comparing forest cover 2000 and 2020 from different sources. NFI data are official forest inventories, conducted by the Chinese government to cover mainland China, Hong Kong, Macao, and Taiwan. In contrast, the other datasets are derived from remote sensing and cover only mainland China. CATCD and GFC provide 30-meter resolution tree cover datasets. We defined "forest" in our study as those pixels with ⩾30% tree cover.

2.2.2 Heterogeneity in the Landscape: documenting an uneven distribution of forest cover

CLCD shows substantial spatial heterogeneity in forest cover across mainland China, 1990-2022. Forest cover in China increased modestly, from about 230 million hectares in 1990 to 240 million in 2022, revealing an average annual growth rate of roughly 0.12% (Fig. A2). Given this slow aggregate change, even small localized impacts from infrastructure investment can be meaningful.

Forest cover has a pronounced southeast-northwest gradient (Fig. 1b). Provinces in the southeast and southwest, such as Fujian (81%), Guangxi (71%), and Yunnan (69%), exhibit high forest cover, while northern and northwestern regions, including Shanghai, Qinghai, Xinjiang, Ningxia, Jiangsu, and Tianjin, have forest cover below 5% (Fig. A3). Forest area is also unevenly distributed: a small number of provinces, including Yunnan (11%), Heilongjiang (9%), Sichuan (8%), Inner Mongolia (8%), Guangxi (6%), and Hunan (5%), account for nearly half of China's total forest area.

Heterogeneity is even more pronounced at finer spatial scales. At the county level, a small fraction of counties (1.3%) holds a disproportionately large share (19%) of forest, while within-province variation exists (Fig. 1c). Similar unevenness is observed across river basins. The Yangtze (35%), Songliao (22%), Pearl (16%), and Southwestern river basins (11%) together contains more than 80% of China's forests, with that forest cover concentrated in a limited number of sub-basins (Fig. A4).

Forest-cover change over time also has varied substantially across the regions of China (Fig. 1d). Net forest gains across this period were concentrated in parts of Sichuan, Shaanxi, Inner Mongolia, Hebei, and Shanxi (increases of 1.79, 1.48, 1.08, 1.05, 0.98, and 0.97 m ha, respectively), reflecting a combination of large-scale ecological restoration programs and just regional land-use dynamics. In contrast, some northeastern and southern areas experienced net forest loss in the study period. These spatial heterogeneities in both the baseline forest conditions and forests' dynamics suggest that impacts of HSR construction and operation could reflect interactions with local conditions.

3. Data

3.1. Forest Cover

CLCD was processed into a binary forest/non-forest classification. To reduce noise from fine-scale classification errors, we aggregated 30-meter pixels to 900-meter cells. For each aggregated cell, we calculated forest-cover share and annual deforestation. The same procedure then was applied to the alternative tree-cover dataset as well, defining forest as tree cover greater than or equal to 30%.

Because China's first HSR construction began in 2004, we restrict the analysis period to 1995-2022. Our primary dependent variable is pixels' annual deforestation rate. We limit the sample to locations with at least 5% forest cover throughout the study period. This avoids extreme deforestation rates driven by small denominators and focuses on sites where forest loss is meaningful and observable.

3.2. High-Speed Rail

We collected information on China's HSRs which commenced operation between 2008 and 2022. High-speed rail is defined by operating speed (Campos and De Rus 2009; Wang et al. 2016); Blanco et al. (2011). China's national railway plan classifies HSR as newly constructed lines exceeding 250km/h or upgraded lines reaching at least 200 km/h. In practice, operational speeds vary due to safety, noise, and engineering constraints. We therefore use 200km/h as the baseline threshold for defining HSR and treat the Hefei-Nanjing Railway as the starting point of China's HSR system[2].

For spatial and temporal information on HSR development, we combined multiple data sources. HSR lines' shapefiles were obtained from OpenStreetMap (OSM) and cross-validated using Google map, Gaode map, and Baidu Map. Information on construction and operation timing was collected from the *China Railway Yearbook,* while stations' coordinates were obtained from Gaode map.

3.3. Conventional Rail and Highway Infrastructure

Pre-existing transport infrastructure may influence or confound the effects of HSR on deforestation (Pfaff et al. 2007). To account for this within controls, and consider heterogeneities of HSR impacts, we check whether each segment runs alongside parallel transport infrastructure, railway or highway that largely overlap the HSR and lie within 5 km. Using this criterion, HSR segments are grouped into four categories: Neither (no parallel rail or highway); Both (both parallel rail and highway); Rail (only parallel railway); and Highway (only parallel highway). Classification was conducted through visual inspection using Gaode Map with overlaid buffer zones (last accessed July 4, 2025).

Within the Neither group, we also identify alternative routes (railways or highways within 10 km with similar directions) and intercepting routes (railways or highways crossing HSR). Excluding segments with either yields an "Empty" group spatially isolated from other transport infrastructure.

[2] Built in 2005 and used since 2008, the Beijing-Tianjin Intercity Railway is considered a benchmark for HSR in China. Yet, we use the Hefei-Nanjing Railway as a starting point. Both commenced operation in 2008 but the Hefei-Nanjing Railway, designed for 250 km/h, was constructed earlier, in 2004, and it aligns with our HSR definition ($\geqslant$200 km/h). Our data include all of the HSR which meet this speed criterion that were operational after the Hefei-Nanjing Railway.

3.4. Control Variables

Our segment fixed-effects models use controls for geographic, climatic, human, and infrastructure characteristics that studies identify as drivers of forest change (Pfaff et al. 2007; Zheng et al. 2019). Geographic controls include elevation and slope from NASA's Digital Elevation Model, aggregated to 900 meters. Hydrological features are drawn from the 1:100 million basic geographic dataset of China's Ministry of Natural Resources. Climatic conditions are captured by 30-year average annual precipitation (1991-2020) provided by the Fine Resolution Mapping of Mountain Environment of China based on WorldClim V2.1. To proxy for human activities, and local development intensity, we calculate distances from each observation to the nearest county center, defined by the location of county governments (Wurm et al. 2021). Pre-treatment transport infrastructure is measured using the distances to conventional railroads and to roads as of 1990, from Baum-Snow et al. (2017)[3].

3.5. Sample(s)

We excluded transport segments with fewer than 50 aggregated spatial units in either the treatment or control group to improve stability. The final dataset consists of 11,937,520 unit-year observations from 426,340 distinct spatial units over a 28-year period. Table 1 presents our descriptive statistics for the variables in the data set with information on HSR segments, initial forest, and forest losses.

Table 1 Full Sample Statistics

	Description	Obs.	Mean	Std. dev.	Min	Max
Elevation	Elevation (meter)	11,937,520	539	524	-2	3843
Slope	Slope (degree)	11,937,520	15	7	0	45
Precip	Average annual rainfall 1990-2020 (mm)	11,937,520	1303	394	295	2302
Distriver	Distance to the nearest river (meter)	11,937,520	1591	1279	0	13104
Distcounty	Distance to county government (meter)	11,937,520	18468	9554	123	75442
Distrail	Distance to railroad before 1990 (meter)	11,937,520	45242	49404	0	218584
Distroad	Distance to national road before 1990 (meter)	11,937,520	8325	6193	0	41475
Segment	Rail portion between two stations (meter)	480	41271	18373	11708	166179
initial stock	Proportion of forest cover in 1995 (%)	426,340	71.26	30.03	5.00	100.00
deforestation rate	Annual deforestation rate in 1995-2022 (%)	11,511,180	-0.23	5.66	-648.33	90.42

Forest cover along HSR routes varies substantially (Fig. 1). To capture this, we calculated average forest share for each segment in 2000, prior to initial HSR construction in 2004. Segments are then put into three groups: low forest share (0 – 30%); medium forest share (30 – 70%); and high forest share (70 – 100%). Most observations fall in high (58%) and medium forest (40%), with relatively

[3] Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics* 99 (3): 435-448. made shapefiles of China's railroad and road networks by digitizing maps from SinoMaps Press (1990). The road network includes national roads ("gonglu") and major urban roads ("dadao", typically six- to eight-lane corridors ranked below national roads).

few in low forest (2%). Table 2 summarizes associated geographic and economic characteristics. Higher forest cover is generally associated with lower elevation, steeper slopes, higher precipitation, greater distance from county centers, and lower proximity to conventional transport infrastructure.

Table 2 Statistics by Forest Cover

	Low (0 – 30%)	Medium (30 – 70%)	High (70 – 100%)
Panel A: Mean			
elevation	823	574	505
slope	13	13	17
precip	681	1218	1385
distriver	2614	1743	1448
distcounty	16702	17409	19264
distrail	18138	36665	52185
distroad	7920	8020	8550
segment	42846	40891	41474
initial stock	23.68	56.00	83.59
deforestation rate	-2.09	-0.55	0.06
Panel B: Number of observations			
Controls	267,932	4,752,272	6,917,316
segment	25	226	229
initial stock	9,569	169,724	247,047
deforest rate	258,363	4,582,548	6,670,269

Note: Controls in Panel B refer to *elevation, slope, precip, distriver, distcounty, distrail, and distroad*, as in Section 3.4.

Since our focus is impacts of new HSR, Table 3 alternative breaks down our full sample according to observations' levels of prior transport infrastructure. These describe distinct contexts, as shown by for instance Panel A's top and bottow rows, e.g., far higher loss rates with high infrastructure.

Table 3 Statistics by Prior Infrastructure

	Empty	Neither	Rail	Highway	Both
Panel A: Mean					
elevation	776	533	409	506	726
slope	18	16	14	15	15
precip	1,318	1,318	1,298	1,332	1,203
distriver	1,692	1,607	1,508	1,577	1,631
distcounty	24,619	19,243	17,478	18,606	16,423
distrail	61,694	44,276	36,353	59,338	30,774
distroad	11,276	9,038	7,881	7,495	7,764
segment	50,869	45,355	39,346	36,966	39,986
initial stock	75.54	72.96	70.92	71.47	65.49
deforestation rate	-0.13	-0.18	-0.23	-0.18	-0.48
Panel B: Number of observations					
Controls	1,056,356	5,773,292	1,402,660	3,043,516	1,718,052
segment	28	201	62	140	77
initial stock	37,727	206,189	50,095	108,697	61,359
deforestation rate	1,018,629	5,567,103	1,352,565	2,934,819	1,656,693

HSR segments also differ in construction and operations time. As in Fig. 3, 90% of HSR segments had at least 5 years of construction and 63% have been in operation for least 5 years. For segments with high forest cover, 94% had construction for at least 5 years (87% for medium, 92% for low) while 70% with high forest were in operation for at least 5 years (60% in medium, 24% in low).

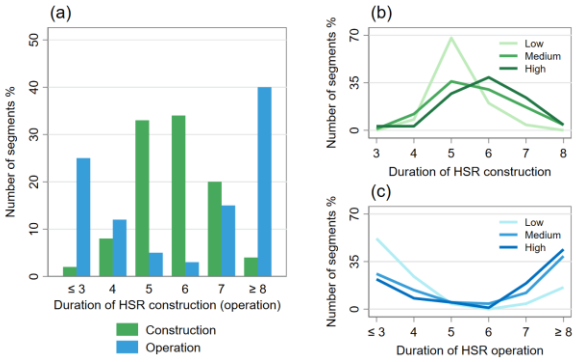


Fig. 3. HSR construction and operation durations. The vertical axis is shares of segments: (a) construction and operation durations for the full sample; (b) construction duration for different groups; and (c) operation duration for different groups.

4. Empirical Methodology

4.1. Sampling Design

We focus on forest outcomes close to HSR lines, where impacts of HSR construction or operation are most likely to be observed. Farther away from transport infrastructure, forest impacts gradually diminish (Barber et al. 2014). We define treatment areas as locations within 5 kilometers of HSR lines. Areas located 5 to 15 kilometers from HSR lines are designated as control zones, reflecting locations sufficiently distance from HSR to serve as a comparison group (Fig. 4). Under this design, our sample includes 3,804,416 treatment observations and 8,133,104 control observations.

As HSR stations are associated with changes in local economic activity and land use, we exclude areas surrounding stations (Fig. 4). Previous research links HSR stations with economic activities and land use in a 4-kilometer radius (Niu et al. 2021; Zheng et al. 2019), an area of roughly 50 km^2 − comparable to the average town or county in China. As the right threshold may vary for stations in provincial capitals versus counties or prefectural cities, we use different criteria: all observations in a 20-km radius for provincial capitals; or in a 4-km radius for the prefectural cities or counties.

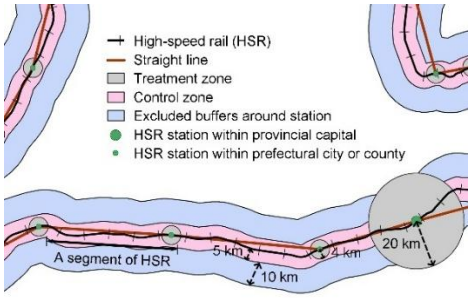


Fig. 4. Sampling process. Note: A segment of HSR refers to a rail portion between two HSR stations.

4.2. Straight Line IV

We constructed our instrumental variable using straight lines between neighboring HSR stations. As the goals are to minimize construction costs and transport times, and as engineering constraints favor direct routes, HSR typically follow straight paths (Losos et al. 2019). Instruments based on straight-line geometry therefore help to mitigate any geographic and social influences on routes. Because rural areas which lie between connected cities are not primary drivers of HSR investments, exposure to new HSR routes along these segments can plausibly be considered to be exogenous.

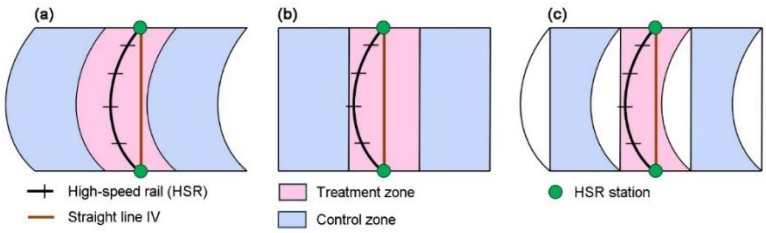


Fig. 5. Construction of the straight-line instrument and screening process, with (a) treatment and control zones around the realized HSR segment, (b) the straight-line proxy between adjacent stations and its buffer zone, and (c) the treatment and control zones for screening process (retaining only aligned locations and excluding mismatched locations).

Fig. 5 depicts our use of straight lines as a benchmark for a screening process that defines treatment and control zones for HSR segments and straight-line proxies. Fig. 5a shows buffers around HSR segments defining treatment and control zones. Fig. 5b applies the same buffer rules to the straight-line connections between adjacent HSR stations to obtain corresponding straight-line buffer zones. Locations in the treatment (or control) zone defined by actual HSR segments and the corresponding straight-line buffer are classified as "exogenous" locations (shaded and hatched area in Fig. 5c). Conversely, locations in the HSR-based treatment (control) but outside the straight-line buffer zone are "endogenous" (blank areas in Fig. 5c) and excluded prior to matching and regression analysis. This screening process yields a smaller but cleaner sample as excluded observations are more likely to reflect route deviations induced by geographic, environmental, or other local constraints. Both treatment and the instrument — i.e., exposure to a straight-line-based HSR segment — are binary. Our binary treatment definition facilitates a direct interpretation of estimated effects as average differences in forest outcomes between areas exposed to HSR and comparable unexposed areas.

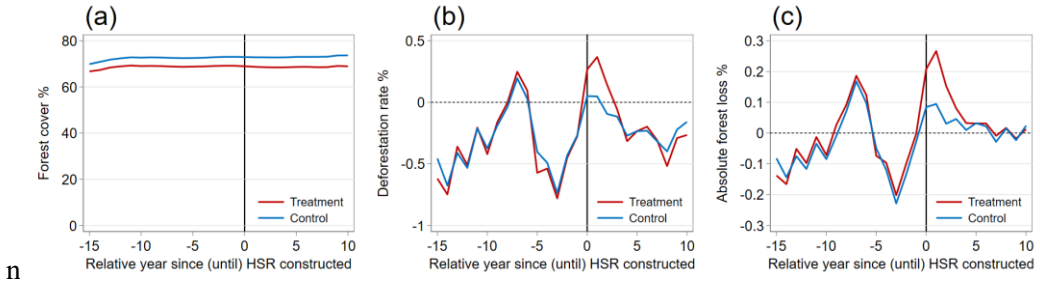
Appendix B provides a standard two-stage least squares (2SLS) specification with the straight-line indicator as an instrument and Table B1's estimates are highly consistent across the two approaches. Its Column 1 reports our 2SLS 1st-stage coefficient of 0.70, statistically significant at 0.1% level with an F-statistic of 320,000 indicating a strong relationship between the instrument and HSR sites. We also calculate, by segment, the share of treatment units located within the straight-line IV zone. The average share is 0.83 for treatment and 0.85 for control, supporting our straight-line benchmark. For the 27 segments for which under 50% of treatment observations overlap the straight-line zone, three-quarters of deviations follow geographical barriers such as mountains, rivers, lakes, or coasts, while the other quarter reflect environmental or economic considerations including protected areas, densely populated regions, and/or existing railway corridors. Table A2 presents examples of both.

4.3. Matching

4.3.1 Approach

Spatial proximity of treatments and controls does not guarantee comparable pre-treatment trends, particularly with heterogeneous landscapes and histories. In Fig. 6a, treatment and controls exhibit similar forest cover, though treated areas are consistently slightly lower. Figs. 6b and 6c indicate broadly parallel trends in forest loss, with some deviations 4 and 5 years prior to HSR construction.

To improve comparability, reducing imbalance, we employ matching methods to construct samples near the new HSR lines. Matching was conducted within each HSR segment, ensuring that treated observations were paired with control units in the same segment. This mitigates the influence of segment-specific factors, such as differences in construction timelines or baseline forest dynamics.



n

Fig. 6. Comparison between treatment and control zones in terms of (a) forest cover trends, (b) relative deforestation rates (comparing to amount of forest stock), and (c) absolute forest loss (albeit as a function of the area in the segment).

Matching was done after screening to ensure that treatment and control groups were drawn from the same exogenously defined sample. We tested multiple approaches, including propensity score matching (PSM) and Mahalanobis distance matching (MDM), both with and without replacement, and calipers. While matching with replacement can improve balance, it may induce dependencies between matched observations that are not fully addressed by sandwich or robust standard errors (Abadie and Spiess 2022). To avoid these issues and achieve better post-matching balance, while minimizing exclusions, we focused upon 1-to-1 PSM, without replacement, using a caliper of 0.1.

We explored many matching covariates including forest stock and loss, elevation, slope, population, and prior infrastructure. Our baseline specification sample was matched on prior forest loss during the last six years before the construction of any new HSR segment, a directly relevant measure that, because it is outcomes-based, captures the combined influences of multiple forest-relevant drivers. When using a strict caliper, not all the treated observations get matched. In the final matched sample, then, we ended up with 2,660,252 observations within each of the treatment and control groups, with ~9% of treatment and ~58% of control observations having been excluded. Table A3 presents the resulting covariate balances both before and after matching. Table A4 reports summary statistics for the matched sample for our baseline specifications. In addition, Appendix D presents robustness checks which make use of alternative samples based upon other matching covariates and methods.

4.3.2 Post-Matching Regressions

Our design allows us to estimate the effects of new HSR on deforestation by comparing forest cover in treated and control zones using a difference-in-differences approach. We did so after completing our screening and matching. While our approach is consistent with prior quasi-experimental studies of transportation infrastructure and land use (Asher et al. 2020; Chen et al. 2021), identification rests upon the assumption that forested areas sufficiently far from HSR stations are exposed to HSR construction exogenously. Our baseline specification is a two-way fixed-effect model (TWFE):

$$Y_{it} = \alpha_1 + \beta_1 start_{it} \times ivnear_i + \gamma_k + \delta_i + \mu_t + \varepsilon_{it} \quad (1)$$

where Y_{it} is the annual deforestation rate for location i in year t, a net percentage decrease in forest cover over the year. The binary indicator $ivnear_i$ equals one if location i falls in the treatment zone by HSR alignment and the straight-line proxy. The indicator $start_{it}$ switches on after the start of HSR construction. The γ_k are event-time fixed effects or indicators for years before and after the start of HSR construction, allowing forest dynamics to vary flexibly over the construction timeline. The terms δ_i and μ_t are location (cell) fixed effects and calendar-year fixed effects, respectively. The error term is ε_{it}. The coefficient β_1 is the average treatment effect of HSR exposure during the post-construction period. Then, however, to distinguish mechanisms through which HSR affects deforestation, we divide the treatment period into construction and operation phases and estimate:

$$Y_{it} = \alpha_1 + \beta_1 start_{it} \times near_i + \beta_2 open_{it} \times near_i + \gamma_k + \delta_i + \mu_t + \varepsilon_{it} \quad (2)$$

where binary $open_{it}$ equals one after HSR operations begin. As operations start after construction, β_1 captures the average treatment effect during construction phases. As HSR is not yet operational at that stage, a significant β_1 is impacts operating through channels other than transport services. Associated indirect channels, such as temporary or permanent access created during construction, offer plausible interpretations. The coefficient β_2 measures the difference between construction and operations effects. As operations' effects on local conditions may be marginal, β_2 also sheds light on whether any construction-related impacts persist, attenuate, or reverse once construction ends.

4.4. Segment Fixed Effects

As a robustness check, we estimate a segment-fixed-effect model using within-segment variation in HSR exposure. This controls for time-invariant characteristics by HSR-segment, incorporating additional geographic, climatic, and other factors. The segment fixed-effect model is specified as:

$$Y_{ijt} = \alpha_1 + \beta_1 start_{it} \times near_i + \theta_1 near_i + \vartheta_1 X_{ij} + \gamma_k + \tau_j + \mu_t + \varepsilon_{ijt} \quad (3)$$

where Y_{ijt} is forest loss for location i in segment j and year t. The vector X_{ij} includes geographic and climatic characteristics, as well as measures of prior investment in other forms of transport infrastructure relevant for segment j. The term τ_j represents segment fixed effect, which control for any time-invariant characteristics specific to each HSR segment. As in the baseline analysis, the sample is first screened using the straight-line instrumental-variable procedure. We further extend this specification by partitioning the post-treatment period into construction and operation phases:

$$Y_{ijt} = \alpha_1 + \beta_1 start_{it} \times near_i + \beta_2 open_{it} \times near_i + \theta_1 near_i + \vartheta_1 X_{ij} + \gamma_k + \tau_j + \mu_t + \varepsilon_{ijt} \quad (4)$$

where $open_{it}$ equals one after HSR operations begin. As before, β_1 captures the average effects in construction, while β_2 measures the incremental change in effects once HSR operations commence.

4.5. Parallel-Trends Identification Assumption

4.5.1 Testing Trends

A key identification assumption for DID designs is parallel trends, i.e., that treatments and controls follow similar trends in the outcome prior to treatment. We assess this in an event-study framework with our matched dataset. Following Autor (2003), we estimated a dynamic specification with leads and lags relative to the start of construction: *Year 7 and earlier*; each of 6 years prior to construction; *Year 0* to *Year 9* after construction; and *Year 10 and later*. That yields the following specifications:

$$Y_{it} = \alpha_2 + \sum_{p=1}^{7} \varphi_{-p} D_{i,t-p} + \sum_{p=0}^{10} \varphi_p D_{i,t+p} + \gamma_{it} + \delta_i + \mu_t + \varepsilon_{it} \quad (5)$$

$$Y_{ijt} = \alpha_2 + \sum_{p=1}^{7} \varphi_{-p} D_{i,t-p} + \sum_{p=0}^{10} \varphi_p D_{i,t+p} + \theta_1 near_i + \vartheta_1 X_{ij} + \gamma_{it} + \tau_j + \mu_t + \varepsilon_{ijt} \quad (6)$$

with $D_{i,t-p}$ lag and $D_{i,t+p}$ lead dummies. *Year 7 and earlier* equals one for at least seven years before construction, while *Year 10 and later* equals one for all years from the tenth year after construction. Each of the other indicators equals one only in the corresponding (relative-time) year. Equation (5) includes individual (cell) fixed effects, while Equation (6) replaces these with segment fixed effects.

4.5.2 Trends-Testing Results

Fig. 7 shows event-study estimates for the full matched sample for individual and segment fixed-effects models. Fig. A5 reports analogs for the individual fixed-effect model, by prior infrastructure. Fig. A6 further extends all the same analysis to the groups defined by baseline forest-cover levels, again using the individual fixed-effects model. For the full sample and all subsamples, the estimated coefficients for the periods prior to HSR construction are statistically indistinguishable from zero. Thus, our results indicate no evidence of differential pre-treatment trends between the treatment and control zones, supporting the parallel-trends assumption for our screened and matched samples.

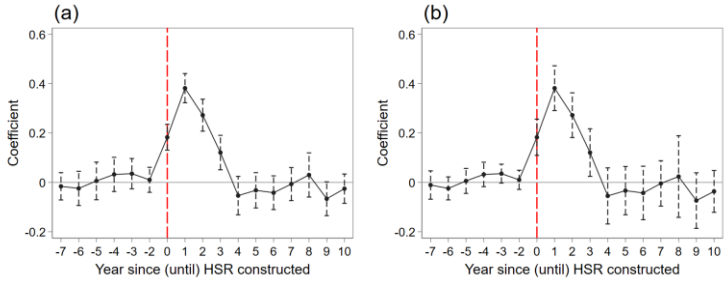


Fig. 7. Parallel trends for the full sample using our matched dataset. The dots stand for the regression coefficients, while the vertical bands represent 95% confidence intervals, for: (a) individual fixed-effect; and (b) segment fixed-effect.

5. Results

5.1. Average Impacts (pooling versus separating Construction & Operation)

Pooling construction and operations in Table 4's Columns 1-3, for average post-treatment effects of HSR on deforestation relative to control zones 5 – 15 km from HSR lines, the areas in 0 – 5 km buffers experienced an average increase of 0.07 percentage points of deforestation per year. Yet we focus on Columns 4-6, which separate construction from operation and show construction impacts.

Table 4's Column 4 shows deforestation rates within the 0-5 km buffer increased by approximately 0.2 percentage points per year relative to control areas during the construction phases of new HSR. In contrast, during operations deforestation rates in treated areas are statistically indistinguishable from those in control areas, as a negative operations coefficient cancels the elevated deforestation rate during construction, which does not persist. Columns 5 and 6 show that this temporal pattern holds for Medium and for High forest cover. Thus, HSR's average impact on forest loss appears to be driven by construction-related disturbances, rather than by the operation of the rail system itself.[4]

Table 4 Baseline estimates of HSR's impact on deforestation.

Dependent variable: deforestation rate	(1) Full Sample	(2) Medium Forest	(3) High Forest	(4) Full Sample	(5) Medium Forest	(6) High Forest
start×ivnear	0.069***	0.021	0.082***	0.199***	0.260***	0.129***
	(0.014)	(0.029)	(0.013)	(0.017)	(0.036)	(0.015)
open×ivnear				-0.262***	-0.477***	-0.096***
				(0.018)	(0.039)	(0.015)
Constant	-0.346***	-0.354***	-0.326***	-0.347***	-0.357***	-0.326***
	(0.020)	(0.043)	(0.017)	(0.020)	(0.043)	(0.017)
Observations	5130486	1983798	3048462	5130486	1983798	3048462
R-squared	0.069	0.065	0.080	0.069	0.066	0.080

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

5.2. Impacts by Prior Infrastructure

We ask whether HSR's forest impact varies with levels of prior transport infrastructure. Columns 1 to 5 in Table 5, from Empty (no prior highways or railways) to Both (prior highways and railways) confirm the small average effect of HSR. But we do not want to pool construction and operations.

Thus, for each level of prior infrastructure, Columns 6 to 10 separate construction from operation. Columns 7-9, which contain the great majority of the sample, unsurprisingly confirm our core result with forest impacts during construction which vanish during operations. Yet, for the Empty context, there was no rise in forest loss during construction and, if anything, operations lower forest losses. For Both contexts, with existing transport to substitute in facilitating construction but also to react to greater flows of people, no effect of construction is followed by a rise in loss during operations.

[4] Breaking these impacts down year-by-year, in Table A5, HSR's impact is larger for the earlier years of construction.

Table 5 Heterogeneous effects across infrastructure groups.

Dependent variable: deforestation rate	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both
start×ivnear	-0.044	0.062***	-0.012	0.062*	0.181***
	(0.038)	(0.019)	(0.044)	(0.030)	(0.041)
open×ivnear					
Constant	-0.043	-0.427***	-1.019***	-0.068	-0.272***
	(0.070)	(0.029)	(0.061)	(0.038)	(0.064)
Observations	468504	2561598	581256	1321866	665766
R-squared	0.071	0.070	0.072	0.073	0.079

Dependent variable: deforestation rate	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
start×ivnear	0.009	0.216***	0.174**	0.220***	0.046
	(0.043)	(0.022)	(0.054)	(0.036)	(0.052)
open×ivnear	-0.117**	-0.309***	-0.375***	-0.319***	0.282***
	(0.045)	(0.023)	(0.062)	(0.035)	(0.061)
Constant	-0.044	-0.429***	-1.020***	-0.067	-0.271***
	(0.070)	(0.029)	(0.061)	(0.038)	(0.064)
Observations	468504	2561598	581256	1321866	665766
R-squared	0.071	0.070	0.073	0.073	0.079

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

5.3 Robustness

5.3.1. Beyond Mechanical Clearing of Right-of-Way?

HSR installation may mechanically imply forest clearing of rights-of-way. To test impacts clearly beyond this, we separate treated units located within 1 km of HSR and their corresponding controls. Table 6's coefficients, in Column 1 and our focus in Column 2, are smaller than above but confirm significant impacts during HSR construction which, just as above, do not continue during operation (Columns 3 and 4 likely contain not only mechanical losses but also disturbances from feeder roads).

Table 6 Effects inside and outside the 1-km track buffer.

Dependent variable: deforestation rate	Excluding the 0-1 km track buffer		The 0-1 km track buffer	
	(1)	(2)	(3)	(4)
start×ivnear	0.046**	0.115***	0.155***	0.507***
	(0.016)	(0.019)	(0.032)	(0.037)
open×ivnear		-0.139***		-0.716***
		(0.020)		(0.040)
Constant	-0.374***	-0.374***	-0.244***	-0.246***
	(0.023)	(0.023)	(0.045)	(0.045)
Observations	4034286	4034286	1096200	1096200
R-squared	0.069	0.069	0.071	0.071

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

5.3.2. Segment Fixed Effects

Table 7 displays results for segment fixed effects, incorporating various covariates. In Column 4, after Columns 1-3 confirmed robust results from above when blending construction and operation, the estimated treatment effects indicating a positive impact of HSR construction on deforestation (moreover, the inclusion of segment fixed effects absorbs some of the impacts of control variables). Column 4 bolsters the conclusion that HSR's impact on deforestation is during construction years.

Table A7 checks it by prior infrastructure, showing generally consistent results with Table 5. Again the Empty, Neither, Rail, and Highway groups show construction impacts reversed during operation. Similarly, in Both, loss impacts persist across construction and operation (Table A7, columns 6-10).

Table 7 Robustness checks for the full sample with segment fixed effects.

Dependent variable: deforestation rate	(1)	(2)	(3)	(4)
start×ivnear	0.057*	0.058*	0.057*	0.191***
	(0.028)	(0.028)	(0.028)	(0.044)
open×ivnear				-0.255***
				(0.067)
ivnear	0.022	-0.005	0.002	0.002
	(0.015)	(0.017)	(0.016)	(0.016)
elevation		-0.000***	0.000	0.000
		(0.000)	(0.000)	(0.000)
slope		0.028***	0.022***	0.022***
		(0.005)	(0.004)	(0.004)
precip		0.000***	-0.000	-0.000
		(0.000)	(0.001)	(0.001)
lndistriver		-0.033***	-0.004	-0.004
		(0.009)	(0.007)	(0.007)
lndistcounty		-0.081	-0.118***	-0.118***
		(0.047)	(0.028)	(0.028)
lndistrail		-0.011	-0.005	-0.005
		(0.020)	(0.017)	(0.017)
lndistroad		-0.042**	-0.021*	-0.021*
		(0.016)	(0.009)	(0.009)
Constant	-0.699***	0.259	1.316	1.321
	(0.192)	(0.473)	(0.771)	(0.772)
Segment fixed effect	No	No	Yes	Yes
Observations	5130486	5130486	5130486	5130486
R-squared	0.005	0.007	0.021	0.021

Notes: All regressions include time fixed effect. Standard errors are clustered at the segment level and stated in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

5.3.3. Segment-level DID Estimates

Recent DID literature shows TWFE may be biased under staggered adoption (De Chaisemartin and d'Haultfoeuille 2020). TWFE coefficients are weighted sums, of group-by-period treatment effects, and weights may be negative (De Chaisemartin and d'Haultfoeuille 2020; Sun and Abraham 2021). In our setting, segments are treated at different times, which may induce negative weights in TWFE.

To address this concern, we estimate treatment effects separately at a segment level, then aggregate using three different weighting rules. Specifically, we estimate the baseline specification separately for each segment to obtain a segment-specific point estimate and standard error. We then aggregate the segment-level estimates usings: simple-average weighting, sample-size weighting, and inverse-variance weighting. Each yields an overall coefficient, standard error, and corresponding p-value. This avoids the negative-weight issue of staggered TWFE, with clean and well-defined estimates by segment, and the aggregation is transparent and mechanically simple in terms of the weighting.

Table 8 Aggregated segment-level estimates under different weighting rules.

Dependent variable:	Simple-average weighting		Sample-size weighting		Inverse-variance weighting	
deforestation rate	(1)	(2)	(3)	(4)	(5)	(6)
start×ivnear	0.080***	0.209***	0.075***	0.170***	0.037***	0.071***
	(0.021)	(0.027)	(0.013)	(0.016)	(0.005)	(0.006)
open×ivnear		-0.258***		-0.172***		-0.061***
		(0.030)		(0.018)		(0.007)

Notes: Standard errors are reported in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table 8 reports the segment-level estimates aggregated using three different weighting rules. While magnitudes vary with weights, results are consistent with our main findings. Columns 1, 3, and 5 support the conclusion that HSR leads to deforestation in treatment zones relative to control zones. Columns 2, 4, and 6 confirm that loss occurs during construction, with no effect during operations.

5.3.4. Alternative Outcome Variables

We check with absolute forest loss as the outcome variable in Table 9's Columns 1 and 2, Table A8, and Table A12 (plus see further robustness checks in Appendix C per relative-loss computations). Table 9's Columns 1 and 2 show forest impacts during construction, consistent with earlier findings, then Table A8's Columns 6-10, by infrastructure, support Table 5: for Neither, Rail, and Highway losses rise during construction but not operations; for Both, HSR impacts occur during operations. Table 9's Columns 3 and 4 employ a different forest data set. They again reveal higher losses during construction that vanish for operations. (Tables A9 and A13 report related heterogeneous impacts.)

Table 9 Robustness checks for the full sample using alternative outcome variables

Dependent variable:	absolute forest loss		deforestation rate	
	(1)	(2)	(3)	(4)
start×ivnear	0.043***	0.103***	0.082***	0.242***
	(0.005)	(0.006)	(0.020)	(0.021)
open×ivnear		-0.122***		-0.324***
		(0.006)		(0.017)
Constant	-0.031***	-0.031***	-4.173***	-4.174***
	(0.007)	(0.007)	(0.036)	(0.036)
Observations	5130486	5130486	4620618	4620618
R-squared	0.078	0.079	0.139	0.139

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

6. Conclusion

Confronted with the intertwined challenges of advancing economic development and safeguarding ecosystems, transportation policymakers in developing countries have compelling reasons to seek a balance between objectives in deciding on which transport infrastructure to implement and where. Informed deliberation required an understanding of how diverse transport infrastructure shapes not only economic development outcomes, such as within the 'end-point' cities that may well motivate most such investments, but also alterations in land use, such as within the forests in between cities.

A significant existing literature emphasizes potential effects on land use, and in particular forests, from investment in roads. Rather few studies differentiate such effects across the distinct contexts within large landscapes, which are central within the 'new economic geography' that consider how facilitating transport can encourage economic growth. Even fewer studies examine rail effects and, of those, very few study high-speed rail. We closely examined the high-speed-rail network in China, the world's largest HSR network, in light of common claims that HSR might lower land-use impacts of rail and, even more, might lower land-use impacts of transport relative to easily accessed roads.

We find HSR can increase deforestation, in particular for low-forest areas − but in a specific way. We contribute to the literature by separating HSR construction from operation and showing impact from the former. Thus, we support priors about HSR versus standard rail, or especially versus roads, in showing no average impact from HSR operation. Yet, HSR construction does have forest impacts, and over twice as large for (rare) low-forest areas, the only setting with impacts of HSR operations.

A construction effect is quite understandable given the use of 'feeder' roads to aid in construction. Whether newly constructed or extensions of existing road networks, with which they can integrate, such construction roads can not only involve forest loss when they are coming into being but also can become relevant local thoroughfares, such as to transport agricultural commodities to market. When not near a rail station, where railroads might not substantively augment transportation access, feeder roads clearly could be drivers of land transformation. This challenges perspectives based on HSR operation alone, suggesting a more nuanced difference between impacts of roads and railroads.

That said, 'feeder' roads are an interpretation, not an empirical finding, as our results are limited by a lack of direct measurement of feeder roads, most likely because they are unofficial pathways. Construction roads often are undocumented on maps and may not always remain post-construction. Subsequent research endeavors could, then, address this data gap to enhance precision in attribution. Nonetheless, it remains important to consider not only rail operations but also railway construction, and to distinguish the contexts rails are entering, for best railway planning given multiple objectives.

Appendix A. Supplementary Figures and Tables



(a) (b)

(c)

Fig. A1. Feeder roads. (a) and (b) show an integral project for temporary and permanent road. (c) depicts the construction of temporary roads. These images were captured during fieldwork conducted by one of the authors.

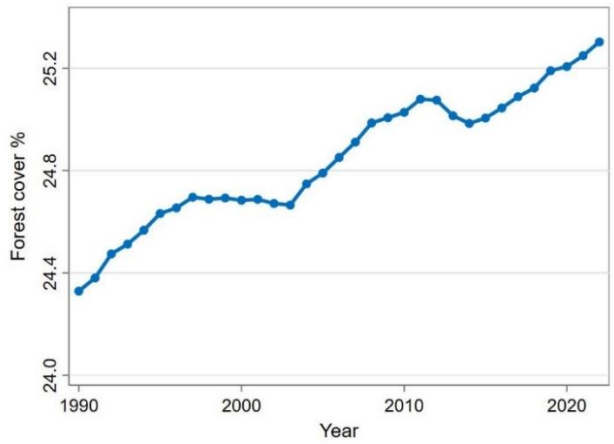


Fig. A2. Forest cover in mainland China from 1990 to 2022.

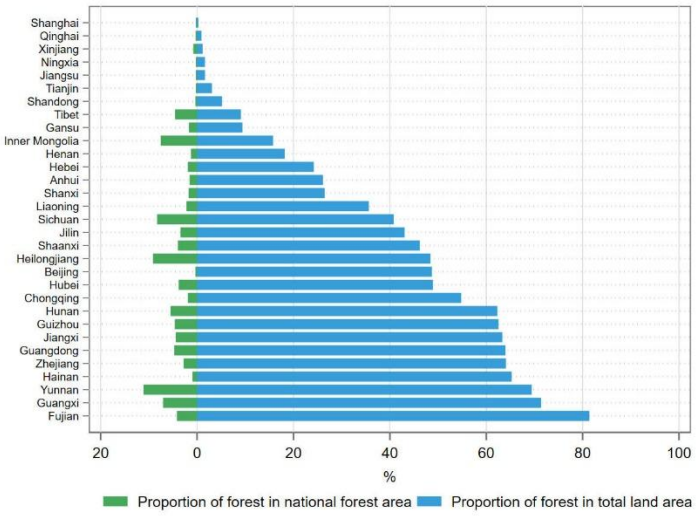


Fig. A3. Proportion of forest area by province in 2022.

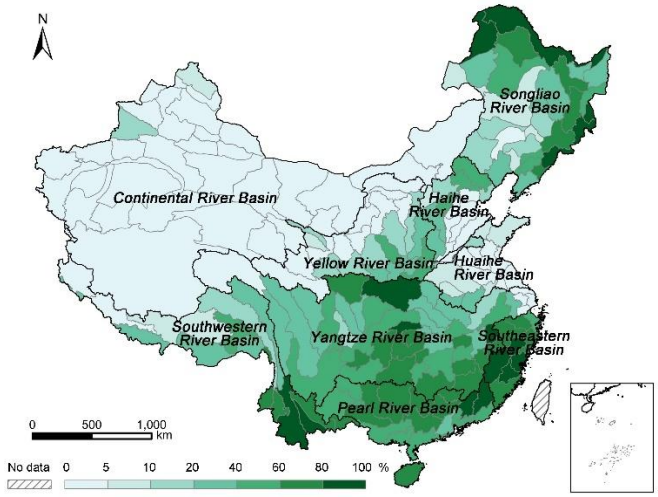


Fig. A4. Forest cover by basin in 2022. Note: Our study includes 9 major basins and 209 sub-basins.

Table A1 Cross-tabulation for the baseline between forest cover and prior-infrastructure groups.

Panel A: Segments

			Prior-infrastructure groups				Total
			Neither	Rail	Highway	Both	
Forest cover group	Low	Count	14	3	6	2	25
		% of Total	2.92	0.63	1.25	0.42	5.21
	Medium	Count	83	29	74	40	226
		% of Total	17.29	6.04	15.42	8.33	47.08
	High	Count	104	30	60	35	229
		% of Total	21.67	6.25	12.5	7.29	47.71
Total		Count	201	62	140	77	480
		% of Total	41.88	12.92	29.17	16.04	100.00

Panel B: Observations

			Prior-infrastructure groups				Total
			Neither	Rail	Highway	Both	
Forest cover groups	Low	Count	5,337	894	2,439	899	9,569
		% of Total	1.25	0.21	0.57	0.21	2.24
	Medium	Count	74,585	16,851	45,330	32,958	169,724
		% of Total	17.49	3.95	10.63	7.73	39.81
	High	Count	126,267	32,350	60,928	27,502	247,047
		% of Total	29.62	7.59	14.29	6.45	57.95
Total		Count	206,189	50,095	108,697	61,359	426,340
		% of Total	48.36	11.75	25.5	14.39	100.00

Table A2 Examples of HSR segments deviating from straight lines.

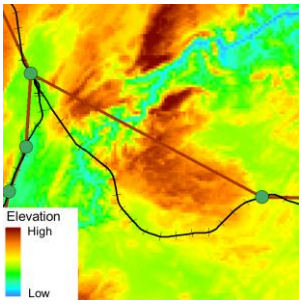
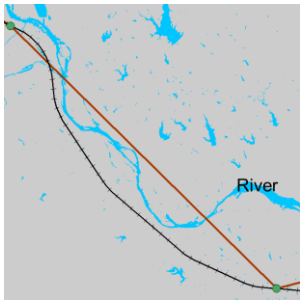
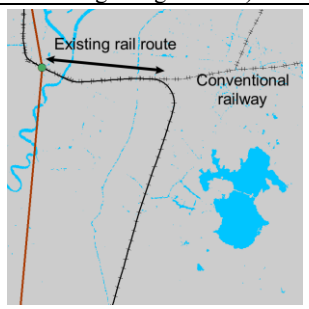
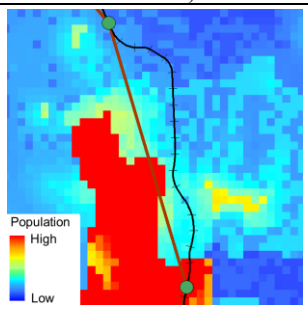
Type	Descriptions		
Geographical barriers			
	Mountain Nanning-Kunming High-speed Railway (Mile-Puzhehei stations)	River Wuhan-Shiyan High-speed Railway (Guchengbei-Longzhong stations)	Ocean Changchun-Hunchun Intercity Railway (Tumenbei-Hunchun stations)
Environmental or economic costs			
	Protected area Shijiazhuang-Taiyuan High-speed Railway (Yangquanbei-Shijiazhuang stations)	Following the existing rail route Ningbo-Taizhou-Wenzhou Railway (Ningbo-Fenghua stations)	Densely populated area Zhengzhou-Taiyuan High-speed Railway (Xiangyuandong-Changzhidong stations)

Table A3 Covariates before and after matching for the baseline.

	Unmatched			Matched			Reduced |bias| %
	Treatment	Control	*p*	Treatment	Control	*p*	
prior_forest_loss1	-0.277	-0.269	0.706	-0.204	-0.228	0.317	-199.7
prior_forest_loss2	-0.451	-0.431	0.349	-0.348	-0.382	0.157	-73.7
prior_forest_loss3	-0.781	-0.738	0.057	-0.628	-0.687	0.022	-38.0
prior_forest_loss4	-0.538	-0.492	0.041	-0.402	-0.458	0.024	-19.7
prior_forest_loss5	-0.573	-0.403	0.000	-0.401	-0.431	0.298	82.4
prior_forest_loss6	0.096	0.029	0.003	0.127	0.127	0.988	99.5

Table A4 Descriptive statistics for matched sample.

	Obs.	Mean	Std. dev.	Min	Max
elevation	5,320,504	512	506	-1	3835
slope	5,320,504	15	7	0	43
precip	5,320,504	1319	387	312	2302
distriver	5,320,504	1551	1213	0	12744
distcounty	5,320,504	17896	9582	123	75240
distrail	5,320,504	44653	49053	0	217318
distroad	5,320,504	7869	6203	0	41173
segment	454	41309	18633	11708	166179
initial stock	190,018	71.42	29.73	5.00	100.00
deforestation rate	5,130,486	-0.18	5.48	-648.33	90.00

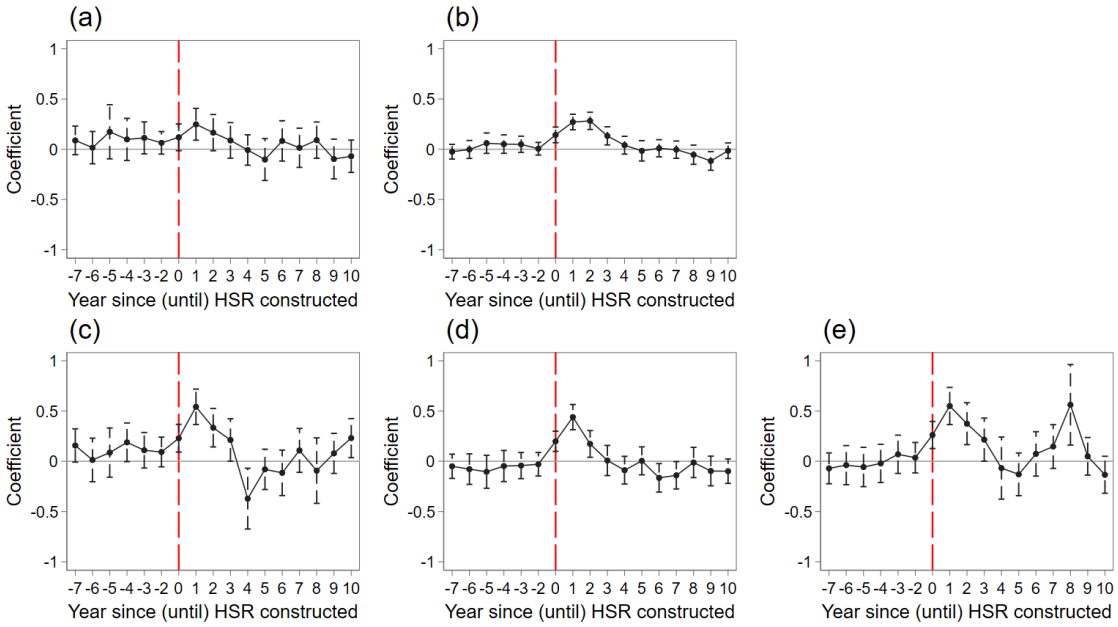


Fig. A5. Parallel trends for infrastructure groups using our matched dataset. The dots stand for the regression coefficients. The vertical bands depict 95% confidence intervals, for: (a) empty group; (b) neither group; (c) rail group; (d) highway group; and (e) both group.

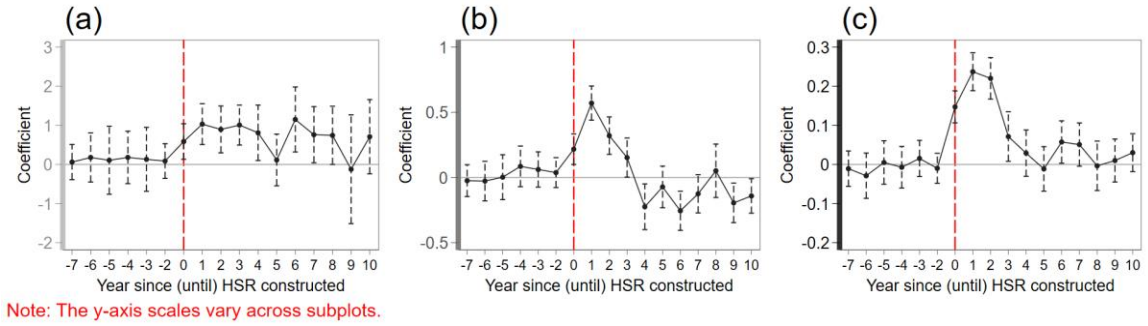


Fig. A6. Parallel trends for forest-cover groups using our matched dataset. The dots stand for the regression coefficients. The vertical bands depict 95% confidence intervals, for: (a) low forest cover; (b) medium forest cover; and (c) high forest cover.

Table A5 Year-by-year estimates of HSR's impact on deforestation.

Dependent variable: deforestation rate	(1)	(2)
start1×ivnear	0.189***	0.175***
	(0.026)	(0.026)
start2×ivnear	0.388***	0.374***
	(0.026)	(0.026)
start3×ivnear	0.278***	0.264***
	(0.024)	(0.024)
start4×ivnear	0.127***	0.119***
	(0.027)	(0.027)
start5×ivnear	-0.047	-0.049
	(0.033)	(0.032)
open1×ivnear		-0.270***
		(0.022)
open2×ivnear		-0.061**
		(0.023)
open3×ivnear		-0.308***
		(0.025)
open4×ivnear		-0.315***
		(0.029)
open5×ivnear		0.057**
		(0.022)
Constant	-0.416***	-0.376***
	(0.022)	(0.022)
Observations	5130486	5130486
R-squared	0.069	0.069

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate significance at the 0.1% level, 1% level, and 5% level, respectively. Column 2 only includes the initial 5 years of construction since HSR segments on average have 5 years for construction. (We only present treatment effect estimates for the initial 5 years of the operational phase because of space limitations.)

Table A6 Effects inside and outside the 1-km track buffer across infrastructure groups.

Dependent variable: deforestation rate	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
Panel A: Excluding the 0-1 km track buffer										
start×ivnear	-0.060	0.046*	-0.034	0.029	0.145**	-0.041	0.139***	0.078	0.135***	-0.050
	(0.043)	(0.021)	(0.049)	(0.033)	(0.046)	(0.049)	(0.024)	(0.060)	(0.040)	(0.058)
open×ivnear						-0.042	-0.186***	-0.224**	-0.214***	0.407***
						(0.051)	(0.025)	(0.069)	(0.039)	(0.069)
Observations	362178	2001402	458028	1044090	530766	362178	2001402	458028	1044090	530766
Panel B: The 0-1 km track buffer										
start×ivnear	0.011	0.119**	0.069	0.186**	0.320***	0.181*	0.493***	0.527***	0.537***	0.422***
	(0.083)	(0.041)	(0.106)	(0.068)	(0.092)	(0.089)	(0.047)	(0.125)	(0.081)	(0.115)
open×ivnear						-0.373***	-0.750***	-0.944***	-0.712***	-0.218
						(0.099)	(0.051)	(0.142)	(0.080)	(0.131)
Observations	106326	560196	123228	277776	135000	106326	560196	123228	277776	135000

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A7 Robustness checks by infrastructure with segment fixed effects.

Dependent variable: deforestation rate	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
start×ivnear	-0.031	0.049	-0.021	0.043	0.181*	0.019	0.207**	0.167	0.213*	0.042
	(0.058)	(0.032)	(0.120)	(0.057)	(0.088)	(0.097)	(0.062)	(0.123)	(0.084)	(0.159)
open×ivnear						-0.105	-0.299**	-0.360**	-0.317*	0.272
						(0.134)	(0.100)	(0.124)	(0.131)	(0.226)
ivnear	0.042	0.000	0.021	0.002	-0.011	0.042	0.000	0.019	0.002	-0.010
	(0.033)	(0.021)	(0.060)	(0.029)	(0.060)	(0.033)	(0.021)	(0.060)	(0.029)	(0.060)
Constant	-0.452	-0.040	3.372	0.917	5.118**	-0.451	-0.035	3.393	0.917	5.105**
	(1.173)	(1.104)	(2.418)	(1.454)	(1.782)	(1.171)	(1.104)	(2.424)	(1.455)	(1.780)
Observations	468504	2561598	581256	1321866	665766	468504	2561598	581256	1321866	665766
R-squared	0.027	0.022	0.026	0.023	0.034	0.027	0.022	0.026	0.023	0.034

Notes: All regressions include control variables, segment fixed effects, and time fixed effects. Standard errors are clustered at the segment level and stated in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A8 Robustness checks by infrastructure using absolute forest loss.

Dependent variable: absolute forest loss	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
start×ivnear	0.015	0.046***	0.023	0.029**	0.073***	0.044**	0.109***	0.112***	0.106***	0.052**
	(0.013)	(0.007)	(0.016)	(0.011)	(0.014)	(0.015)	(0.008)	(0.019)	(0.013)	(0.017)
open×ivnear						-0.065***	-0.126***	-0.181***	-0.155***	0.043*
						(0.017)	(0.008)	(0.021)	(0.014)	(0.019)
Constant	0.116***	-0.065***	-0.276***	0.065***	0.013	0.116***	-0.066***	-0.276***	0.065***	0.013
	(0.022)	(0.010)	(0.018)	(0.013)	(0.018)	(0.022)	(0.010)	(0.018)	(0.013)	(0.018)
Observations	468504	2561598	581256	1321866	665766	468504	2561598	581256	1321866	665766
R-squared	0.079	0.080	0.080	0.082	0.094	0.079	0.080	0.081	0.082	0.094

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and stated in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A9 Robustness checks by infrastructure using different forest dataset.

Dependent variable: deforestation rate	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
start×ivnear	-0.141*	0.022	0.119*	0.109**	0.234***	-0.073	0.146***	0.299***	0.301***	0.445***
	(0.056)	(0.029)	(0.053)	(0.039)	(0.060)	(0.058)	(0.030)	(0.056)	(0.040)	(0.061)
open×ivnear						-0.148***	-0.247***	-0.369***	-0.391***	-0.440***
						(0.045)	(0.022)	(0.051)	(0.037)	(0.049)
Constant	-1.594***	-3.373***	-4.014***	-5.283***	-4.877***	-1.594***	-3.374***	-4.015***	-5.281***	-4.879***
	(0.177)	(0.057)	(0.080)	(0.066)	(0.092)	(0.177)	(0.057)	(0.080)	(0.066)	(0.092)
Observations	422442	2324592	519156	1202202	574668	422442	2324592	519156	1202202	574668
R-squared	0.125	0.120	0.174	0.184	0.176	0.125	0.120	0.174	0.184	0.176

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A10 Effects inside and outside the 1-km track buffer across forest cover groups.

Dependent variable: deforestation rate	(1) Medium	(2) High	(3) Medium	(4) High
Panel A: Excluding the 0-1 km track buffer				
start×ivnear	-0.006	0.064***	0.139***	0.071***
	(0.033)	(0.015)	(0.040)	(0.017)
open×ivnear			-0.289***	-0.014
			(0.043)	(0.017)
Observations	1570860	2386368	1570860	2386368
Panel B: The 0-1 km track buffer				
start×ivnear	0.125	0.145***	0.718***	0.338***
	(0.067)	(0.030)	(0.081)	(0.034)
open×ivnear			-1.197***	-0.392***
			(0.091)	(0.032)
Observations	412938	662094	412938	662094

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A11 Robustness checks by forest cover with segment fixed effects.

Dependent variable: deforestation rate	(1) Low	(2) Medium	(3) High	(4) Low	(5) Medium	(6) High
start×ivnear	0.610	-0.001	0.078***	0.461	0.251**	0.123***
	(0.303)	(0.061)	(0.023)	(0.545)	(0.096)	(0.029)
open×ivnear				0.327	-0.472**	-0.088**
				(0.710)	(0.147)	(0.033)
ivnear	-0.155	0.051	-0.009	-0.156	0.051	-0.009
	(0.121)	(0.034)	(0.015)	(0.121)	(0.034)	(0.015)
Constant	2.755	-0.503	0.736	2.798	-0.484	0.737
	(5.559)	(1.486)	(0.919)	(5.576)	(1.491)	(0.919)
Observations	98226	1983798	3048462	98226	1983798	3048462
R-squared	0.072	0.022	0.019	0.072	0.022	0.019

Notes: All regressions include control variables, segment fixed effects, and time fixed effects. Standard errors are clustered at the segment level and stated in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A12 Robustness checks by forest cover using absolute forest loss.

Dependent variable: absolute forest loss	(1) Low	(2) Medium	(3) High	(4) Low	(5) Medium	(6) High
start×ivnear	0.156***	0.023*	0.052***	0.146***	0.124***	0.082***
	(0.030)	(0.010)	(0.006)	(0.037)	(0.012)	(0.007)
open×ivnear				0.023	-0.203***	-0.062***
				(0.050)	(0.013)	(0.007)
Constant	-0.350***	0.055***	-0.090***	-0.350***	0.054***	-0.090***
	(0.023)	(0.013)	(0.007)	(0.023)	(0.013)	(0.007)
Observations	98226	1983798	3048462	98226	1983798	3048462
R-squared	0.142	0.076	0.087	0.142	0.076	0.087

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and stated in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table A13 Robustness checks by forest cover using different forest dataset.

Dependent variable: deforestation rate	(1) Low	(2) Medium	(3) High	(4) Low	(5) Medium	(6) High
start×ivnear	0.222	0.054	0.112***	-0.090	0.302***	0.195***
	(0.169)	(0.030)	(0.024)	(0.167)	(0.031)	(0.026)
open×ivnear				0.644***	-0.497***	-0.171***
				(0.142)	(0.026)	(0.019)
Constant	-4.792***	-6.179***	-1.459***	-4.786***	-6.180***	-1.459***
	(0.294)	(0.054)	(0.046)	(0.294)	(0.054)	(0.046)
Observations	80190	2517426	2023002	80190	2517426	2023002
R-squared	0.141	0.156	0.097	0.141	0.156	0.097

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Appendix B. Two-Stage Least Squares (2SLS)

Table B1 reports the estimated average post-treatment effects using 2SLS and screening process. Columns 1 and 2 present the standard 2SLS estimates, with the first stage in Column 1 and the second stage in Column 2. Columns 3 and 4 report the manual 2SLS results. In Column 3, we first obtain fitted values from the first stage and then use it as the treatment variable in the second-stage regression. As expected, this manual approach reproduces the point estimate in Column 2. The fitted values are 0.0518 and 0.7527, corresponding to the predicted values for the control and treated, respectively. In Column 4, we binarize the fitted values (i.e., recode them as 0 and 1) and estimate the second-stage regression. The coefficient is smaller than that in Column 3 because it is reported on a different treatment scale[5]. Column 5 reports the result for the screened sample. The estimate is nearly identical to the corresponding estimates in Columns 2 and 3, indicating that the screening process does not materially affect the estimated effect. Instead, it yields a cleaner sample while preserving the treatment effect implied by 2SLS.

Table B1 The average post-treatment effect for 2SLS and Screening process.

	Standard 2SLS		Manual 2SLS		(5)
Dependent variable:	(1) First Stage start×near	(2) Second Stage def_rate	(3) Second Satge: fitted def_rate	(4) Second Stage: binary fitted def_rate	Screening process def_rate
start×near		0.083***			
		(0.016)			
start×ivnear	0.701***		0.083***	0.058***	0.084***
	(0.001)		(0.016)	(0.011)	(0.012)
Constant			-0.237***	-0.420***	-0.433***
			(0.002)	(0.014)	(0.015)
Observations	10588563	10588563	10588563	10588563	9130401
R-squared			0.069	0.069	0.069
First-stage F	320000				

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

[5] Column 4's should equal Column 3's multiplied by the gap in fitted values, i.e., 0.0582≈0.083*(0.7527-0.0518).

Appendix C. Robustness to Different Deforestation-rate Thresholds

While the 5% threshold used in the main regression helps eliminate extreme deforestation rates driven by small denominators, it may introduce bias as excluded units often experienced more severe deforestation. To assess the sensitivity of our findings, we conduct robustness checks using different thresholds. Table C1 presents statistics on deforestation rate, number of observations, and the number of segments across thresholds. Table C2 and C3 report the regression results for the full sample, using relative and absolute measures of deforestation, under each of different thresholds.

Table C1 Summary statistics across different thresholds.

Panel A: Stats of deforestation rate %			
	forest cover > 0	forest cover ≥ 10%	forest cover ≥ 20%
Min	-22500.01	-372.56	-199.60
Max	99.38	82.35	73.73
Mean	-1.11	-0.11	-0.03

Panel B: Number of observations and segments						
	forest cover > 0		forest cover ≥ 10%		forest cover ≥ 20%	
	Unmatched	Matched	Unmatched	Matched	Unmatched	Matched
Total obs.	11,239,060	6,477,688	8,668,464	4,917,864	7,781,620	4,391,688
Treatment	3,557,736	3,238,844	2,703,540	2,458,932	2,407,580	2,195,844
Control	7,681,324	3,238,844	5,964,924	2,458,932	5,374,040	2,195,844
Dropped obs.	/	42%	/	43%	/	44%
Treatment	/	9%	/	9%	/	9%
Control	/	58%	/	59%	/	59%
Segments	516		427		403	

In Columns 1 and 4 in Tables C2 and C3, we first lower the threshold to include all the units with forest cover greater than zero, i.e., to reintroduce the most severely deforested areas to assess their influence on core results. For the full sample, the coefficient during construction rises to 0.4%, i.e., twice the baseline estimate (Column 4 in Table C2). In contrast, the estimate under the absolute measure remains at 0.1%, identical to baseline (Column 4 in Table C3). This suggests amplification in the relative specification is mechanical.

Heterogeneity results support this interpretation, particularly for Low forest cover (Panels A in Tables C4 and C5). In Low group, the construction effect reaches 0.98% under the relative measure—the highest across groups—but only 0.05% under the absolute measure, the lowest. This striking discrepancy reflects the inclusion of units with forest cover below 5%, where minimal forest loss is exaggerated in relative terms due to tiny denominators. By contrast, Medium and High forest cover show smaller construction effects and little to no impact for operation, consistent with Table 5. Notably, the construction effect for High forest cover is nearly identical to above, highlighting the stability of results for these areas.

In Columns 2 and 5 in Tables C2 and C3, we instead restrict the sample to units with forest cover greater than 10% and see that the full-sample estimates are nearly identical to the baseline results in Table 3. Columns 4 to 6 (Panels B in Tables C4 and C5) indicate positive and significant

construction effects across groups, with the higher impact in Low forest cover — consistent with heterogeneity in Table 5. Notably, the construction effect for Low forest cover reaches 1.64% using the relative measure and 0.61% using the absolute measure. This consistency of increased effect suggests that, after excluding the lowest-cover units, remaining observations capture deforestation impacts (rather than reflecting mechanical inflation).

We further restrict the sample in Columns 3 and 6 (Tables C2 and C3) to include only units with forest cover greater than 20%. As in Panels C (Tables C4 and C5), this entirely removes Low forest cover and greatly reduces the sample size for Medium forest, undermining representativeness and comparability across groups. High forest cover still has a sufficient and stable sample with estimated impact consistent with above: deforestation impacts mostly occur during construction, then become negligible during operation.

Tables C6 and C7 offer infrastructure group results using both relative and absolute measures across different thresholds. As shown in Panels A to C, the results for infrastructure groups remain largely consistent whether deforestation is measured in relative or absolute terms. Moreover, the results under different thresholds remain robust – deforestation continues after HSR start operation for the Both group, while earlier deforestation tends to diminish over time for the other groups.

Table C2 Robustness checks for the full sample using different thresholds (the relative measure).

Dependent variable: deforestation rate	(1) forest cover > 0%	(2) forest cover ≥ 10%	(3) forest cover ≥ 20%	(4) forest cover > 0%	(5) forest cover ≥ 10%	(6) forest cover ≥ 20%
start×ivnear	0.142**	0.067***	0.058***	0.400***	0.169***	0.131***
	(0.044)	(0.012)	(0.009)	(0.058)	(0.014)	(0.011)
open×ivnear				-0.520***	-0.206***	-0.147***
				(0.059)	(0.015)	(0.011)
Constant	-1.469***	-0.181***	-0.113***	-1.471***	-0.181***	-0.113***
	(0.068)	(0.016)	(0.013)	(0.068)	(0.016)	(0.013)
Observations	6246342	4742226	4234842	6246342	4742226	4234842
R-squared	0.049	0.070	0.072	0.049	0.070	0.072

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table C3 Robustness checks for the full sample using different thresholds (the absolute measure).

Dependent variable: absolute forest loss	(1) forest cover > 0%	(2) forest cover ≥ 10%	(3) forest cover ≥ 20%	(4) forest cover > 0%	(5) forest cover ≥ 10%	(6) forest cover ≥ 20%
start×ivnear	0.043***	0.044***	0.042***	0.103***	0.099***	0.090***
	(0.005)	(0.005)	(0.005)	(0.006)	(0.006)	(0.006)
open×ivnear				-0.121***	-0.112***	-0.096***
				(0.006)	(0.006)	(0.006)
Constant	-0.056***	-0.010	-0.012	-0.056***	-0.010	-0.012
	(0.006)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)
Observations	6246342	4742226	4234842	6246342	4742226	4234842
R-squared	0.077	0.078	0.078	0.077	0.078	0.078

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table C4 Robustness checks by forest cover using different thresholds (the relative measure).

Dependent variable: deforestation rate	(1) Low	(2) Medium	(3) High	(4) Low	(5) Medium	(6) High
Panel A: Sample with forest cover > 0%						
start×ivnear	0.859***	0.034	0.089***	0.985**	0.435***	0.139***
	(0.234)	(0.074)	(0.025)	(0.311)	(0.098)	(0.029)
open×ivnear				-0.267	-0.798***	-0.101***
				(0.323)	(0.098)	(0.028)
Observations	678024	2986848	2581470	678024	2986848	2581470
Panel B: Sample with forest cover ≥ 10%						
start×ivnear	0.853**	0.024	0.085***	1.641***	0.229***	0.126***
	(0.309)	(0.027)	(0.011)	(0.420)	(0.032)	(0.013)
open×ivnear				-1.682***	-0.428***	-0.083***
				(0.459)	(0.036)	(0.013)
Observations	15336	1554606	3172284	15336	1554606	3172284
Panel C: Sample with forest cover ≥ 20%						
start×ivnear		-0.011	0.074***		0.081*	0.143***
		(0.028)	(0.009)		(0.033)	(0.011)
open×ivnear					-0.203***	-0.136***
					(0.037)	(0.012)
Observations		797094	3437748		797094	3437748

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table C5 Robustness checks by forest cover using different thresholds (the absolute measure).

Dependent variable: absolute forest loss	(1) Low	(2) Medium	(3) High	(4) Low	(5) Medium	(6) High
Panel A: Sample with forest cover > 0%						
start×ivnear	0.037*	0.031***	0.057***	0.050**	0.119***	0.088***
	(0.015)	(0.008)	(0.006)	(0.018)	(0.010)	(0.007)
open×ivnear				-0.028	-0.176***	-0.063***
				(0.022)	(0.010)	(0.007)
Observations	678024	2986848	2581470	678024	2986848	2581470
Panel B: Sample with forest cover ≥ 10%						
start×ivnear	0.322**	0.018	0.055***	0.606***	0.119***	0.084***
	(0.101)	(0.011)	(0.005)	(0.134)	(0.013)	(0.006)
open×ivnear				-0.606***	-0.210***	-0.059***
				(0.155)	(0.014)	(0.007)
Observations	15336	1554606	3172284	15336	1554606	3172284
Panel C: Sample with forest cover ≥ 20%						
start×ivnear		-0.001	0.052***		0.049**	0.099***
		(0.014)	(0.005)		(0.016)	(0.006)
open×ivnear					-0.112***	-0.093***
					(0.018)	(0.007)
Observations		797094	3437748		797094	3437748

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table C6 Robustness checks by infrastructure using different thresholds (the relative measure).

Dependent variable: deforestation rate	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
Panel A: Sample with forest cover > 0%										
start×ivnear	-0.110	0.107	0.100	0.088	0.400**	-0.099	0.442***	0.333*	0.326***	0.252
	(0.078)	(0.058)	(0.116)	(0.088)	(0.150)	(0.085)	(0.079)	(0.137)	(0.092)	(0.229)
open×ivnear						-0.023	-0.671***	-0.465**	-0.484***	0.301
						(0.089)	(0.078)	(0.156)	(0.082)	(0.237)
Observations	503280	3014658	742014	1618110	871560	503280	3014658	742014	1618110	871560
Panel B: Sample with forest cover ≥ 10%										
start×ivnear	0.031	0.081***	-0.038	0.048	0.144***	0.088*	0.204***	0.080	0.160***	0.098*
	(0.030)	(0.015)	(0.037)	(0.025)	(0.034)	(0.034)	(0.018)	(0.046)	(0.030)	(0.042)
open×ivnear						-0.126***	-0.245***	-0.241***	-0.226***	0.098*
						(0.037)	(0.019)	(0.052)	(0.030)	(0.045)
Observations	449874	2401758	530118	1216134	594216	449874	2401758	530118	1216134	594216
Panel C: Sample with forest cover ≥ 20%										
start×ivnear	0.012	0.064***	0.004	0.049*	0.105***	0.065*	0.144***	0.111**	0.130***	0.082*
	(0.023)	(0.012)	(0.029)	(0.020)	(0.027)	(0.026)	(0.014)	(0.035)	(0.024)	(0.034)
open×ivnear						-0.115***	-0.160***	-0.222***	-0.164***	0.050
						(0.029)	(0.015)	(0.039)	(0.024)	(0.036)
Observations	420660	2171718	473472	1085940	503712	420660	2171718	473472	1085940	503712

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table C7 Robustness checks by infrastructure using different thresholds (the absolute measure).

Dependent variable: absolute forest loss	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
Panel A: Sample with forest cover > 0%										
start×ivnear	0.021	0.056***	0.004	0.024*	0.064***	0.033*	0.119***	0.090***	0.102***	0.046**
	(0.014)	(0.007)	(0.015)	(0.010)	(0.012)	(0.015)	(0.008)	(0.017)	(0.012)	(0.015)
open×ivnear						-0.027	-0.128***	-0.172***	-0.157***	0.036*
						(0.018)	(0.008)	(0.018)	(0.012)	(0.016)
Observations	503280	3014658	742014	1618110	871560	503280	3014658	742014	1618110	871560
Panel B: Sample with forest cover ≥ 10%										
start×ivnear	0.028*	0.051***	-0.000	0.031**	0.079***	0.057***	0.110***	0.080***	0.094***	0.071***
	(0.013)	(0.007)	(0.017)	(0.011)	(0.014)	(0.015)	(0.008)	(0.020)	(0.013)	(0.017)
open×ivnear						-0.063***	-0.119***	-0.164***	-0.128***	0.017
						(0.017)	(0.008)	(0.022)	(0.014)	(0.019)
Observations	449874	2401758	530118	1216134	594216	449874	2401758	530118	1216134	594216
Panel C: Sample with forest cover ≥ 20%										
start×ivnear	0.015	0.048***	0.013	0.029**	0.074***	0.045**	0.096***	0.087***	0.085***	0.071***
	(0.012)	(0.006)	(0.016)	(0.011)	(0.014)	(0.014)	(0.008)	(0.019)	(0.013)	(0.018)
open×ivnear						-0.065***	-0.097***	-0.152***	-0.112***	0.005
						(0.016)	(0.008)	(0.021)	(0.014)	(0.019)
Observations	420660	2171718	473472	1085940	503712	420660	2171718	473472	1085940	503712

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Appendix D. Robustness to Different Matching Design

We first conduct 1-to-1 propensity score matching (PSM) using an alternative set of covariates — including elevation, slope, population, transport infrastructure, and initial forest stock — to assess the robustness of the findings. Population for 2000 (i.e., pre-treatment) at 100 m resolution are obtained from WorldPop and aggregated to 900×900-meter cells.

To examine sensitivity to the matching method, we further adopt 1-to-1 Mahalanobis distance matching (MDM), considering both matching with replacement and without replacement. Consistent with the baseline matching design, we also include prior forest loss over the six years preceding HSR construction.

Table D1 presents covariate balance statistics under different matching designs. Panels A, B, and C illustrate that the covariate differences between treated and control units are significantly diminished after matching. Fig. D1 further demonstrates that the matched samples generated by these alternative designs support the parallel trends assumption.

Table D1 Covariates before and after matching under different matching design.

	Unmatched			Matched			Reduced
	Treatment	Control	*p*	Treatment	Control	*p*	|bias| %
Panel A. PSM with alternative covariates							
elevation	485.090	543.070	0.000	510.130	507.830	0.382	96.0
slope	14.469	15.454	0.000	14.791	14.751	0.293	96.0
population	151.610	113.450	0.000	120.070	121.130	0.406	97.2
distroad	6928.100	8784.900	0.000	7776.500	7691.200	0.010	95.4
Initial stock	68.912	72.540	0.000	70.445	70.086	0.022	90.1
Panel B. MDM with replacement							
prior_forest_loss1	-0.277	-0.269	0.706	-0.132	-0.128	0.830	49.4
prior_forest_loss2	-0.451	-0.431	0.349	-0.280	-0.262	0.355	11.2
prior_forest_loss3	-0.781	-0.738	0.057	-0.573	-0.526	0.017	-7.8
prior_forest_loss4	-0.538	-0.492	0.041	-0.436	-0.399	0.083	20.2
prior_forest_loss5	-0.573	-0.403	0.000	-0.460	-0.409	0.045	70.3
prior_forest_loss6	0.096	0.029	0.003	0.078	0.068	0.582	84.6
Panel C. MDM without replacement							
prior_forest_loss1	-0.277	-0.269	0.709	-0.110	-0.121	0.542	-34.9
prior_forest_loss2	-0.451	-0.431	0.351	-0.251	-0.242	0.596	52.6
prior_forest_loss3	0.781	-0.738	0.057	-0.524	-0.498	0.168	40.2
prior_forest_loss4	-0.538	-0.492	0.041	-0.387	-0.363	0.216	49.0
prior_forest_loss5	-0.573	-0.403	0.000	-0.402	-0.372	0.187	82.4
prior_forest_loss6	0.096	0.029	0.002	0.073	0.058	0.413	78.3

Note: Panel A reports results from matching without replacement, using a caliper of 0.1. In Panels B and C, we set the caliper to 5 and 4, respectively, to achieve better covariate balance and retain more observations. Tighter calipers improve covariate similarity but discard too many observations, which reduces representativeness and statistical power; in our case, overly restrictive calipers also cause the pre-trend test to fail.

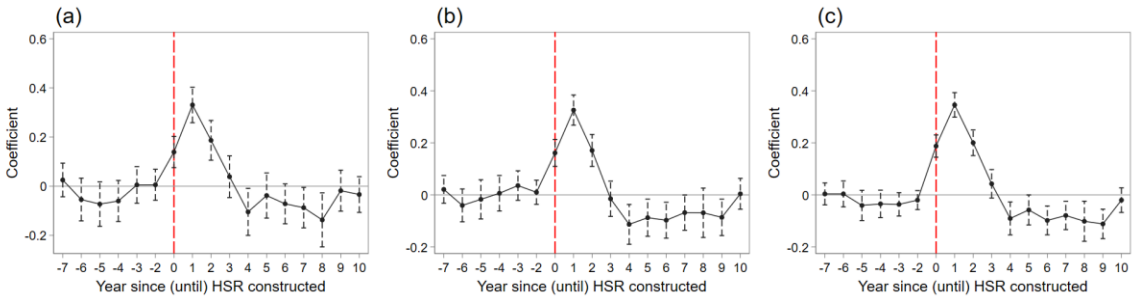


Fig. D1 Parallel trends for full sample under different matching design. The dots stand for the regression coefficients. The vertical bands depict 95% confidence intervals, for: (a) PSM with alternative covariates; (b) MDM with replacement; (c) MDM with replacement.

Our main results remain qualitatively unchanged when we replicate the Table 3 regressions using different matching designs (Table D2). Columns 4-6 show that the positive correlation between construction and deforestation persists, with coefficients ranging from 0.13 to 0.15 percentage points. The deforestation caused by construction is offset once HSR begins operation, which further supports our earlier conclusion that the operational phase has no additional impact on deforestation.

We also check robustness for HSR's effect on deforestation across different pre-treatment forest cover and pre-treatment Infrastructure in Table D3 and Table D4. Table D3 shows that construction is associated with significant forest loss across groups, with larger effect in the Low group. After operations begin, the Medium and High groups exhibit offsetting effects that largely reverse the losses during the construction. Table D4 mirrors our earlier findings across infrastructure groups. For the Empty, Neither, Rail, and Highway groups, construction impacts are largely offset during operation, whereas forest losses persist in the Both group. These results indicate that our heterogeneous effects are robust and do not rely on the matching design.

Table D2 Robustness checks for the full sample under different matching design.

Dependent variable: deforestation rate	(1) PSM with alternative covariates	(2) MDM with replacement	(3) MDM without replacement	(4) PSM with alternative covariates	(5) MDM with replacement	(6) MDM without replacement
start×ivnear	0.017	0.006	0.030*	0.134***	0.116***	0.149***
	(0.017)	(0.015)	(0.012)	(0.020)	(0.017)	(0.014)
open×ivnear				-0.236***	-0.223***	-0.240***
				(0.021)	(0.018)	(0.016)
Constant	-0.335***	-0.324***	-0.303***	-0.336***	-0.325***	-0.304***
	(0.025)	(0.021)	(0.018)	(0.025)	(0.021)	(0.018)
Observations	3982716	4131162	4943430	3982716	4131162	4943430
R-squared	0.067	0.072	0.072	0.067	0.072	0.073

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table D3 Robustness checks by forest cover under different matching design.

Dependent variable: deforestation rate	(1) Low	(2) Medium	(3) High	(4) Low	(5) Medium	(6) High
Panel A. PSM with alternative covariates						
start×ivnear	0.302	-0.038	0.048**	-0.184	0.210***	0.065***
	(0.161)	(0.034)	(0.017)	(0.196)	(0.042)	(0.020)
open×ivnear				1.199***	-0.493***	-0.034
				(0.284)	(0.045)	(0.018)
Observations	60426	1604880	2317410	60426	1604880	2317410
Panel B. MDM with replacement						
start×ivnear	0.439**	-0.082**	0.037**	0.187	0.149***	0.068***
	(0.137)	(0.029)	(0.013)	(0.169)	(0.035)	(0.015)
open×ivnear				0.593*	-0.463***	-0.063***
				(0.242)	(0.039)	(0.014)
Observations	85536	1706481	2339145	85536	1706481	2339145
Panel C. MDM without replacement						
start×ivnear	0.562***	-0.069**	0.080***	0.419**	0.164***	0.116***
	(0.130)	(0.025)	(0.010)	(0.161)	(0.032)	(0.012)
open×ivnear				0.338	-0.466***	-0.074***
				(0.232)	(0.036)	(0.012)
Observations	94716	1931634	2917080	94716	1931634	2917080

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Table D4 Robustness checks by infrastructure under different matching design.

Dependent variable: deforestation rate	(1) Empty	(2) Neither	(3) Rail	(4) Highway	(5) Both	(6) Empty	(7) Neither	(8) Rail	(9) Highway	(10) Both
Panel A. PSM with alternative covariates										
start×ivnear	-0.051	0.080***	-0.218***	-0.117**	0.198***	-0.004	0.244***	-0.080	0.002	0.000
	(0.038)	(0.021)	(0.060)	(0.038)	(0.054)	(0.046)	(0.025)	(0.074)	(0.045)	(0.069)
open×ivnear						-0.102*	-0.326***	-0.281***	-0.241***	0.414***
						(0.049)	(0.026)	(0.083)	(0.045)	(0.079)
Observations	410724	2164482	376542	989226	452466	410724	2164482	376542	989226	452466
Panel B. MDM with replacement										
start×ivnear	0.036	0.017	-0.142**	-0.020	0.116**	0.024	0.146***	0.070	0.127***	-0.041
	(0.044)	(0.020)	(0.046)	(0.030)	(0.041)	(0.047)	(0.023)	(0.056)	(0.036)	(0.051)
open×ivnear						0.027	-0.259***	-0.428***	-0.298***	0.329***
						(0.048)	(0.024)	(0.064)	(0.035)	(0.062)
Observations	351837	2026107	463752	1087965	553338	351837	2026107	463752	1087965	553338
Panel C. MDM without replacement										
start×ivnear	-0.007	0.050**	-0.077*	0.011	0.088*	0.033	0.185***	0.096*	0.149***	0.010
	(0.031)	(0.015)	(0.038)	(0.025)	(0.035)	(0.033)	(0.018)	(0.047)	(0.030)	(0.045)
open×ivnear						-0.089*	-0.271***	-0.351***	-0.279***	0.163**
						(0.038)	(0.020)	(0.058)	(0.031)	(0.055)
Observations	445014	2467530	551880	1280016	644004	445014	2467530	551880	1280016	644004

Notes: All regressions include individual and time fixed effects. Standard errors are clustered at the individual level and in parentheses below estimates. ***, **, *, indicate the significance at the 0.1% level, 1% level, and 5% level, respectively.

Appendix E. Overview of Forest datasets

Table E1 Overview of major high-resolution forest datasets.

Dataset	Resolution	Spatial coverage	Temporal coverage	Pixel value	Forest definition	Reported accuracy	Satellite imagery	Source
The 30-m Annual China Land Cover Dataset (CLCD)	30m	China	1985, 1990–2022	Land use/cover class	Following FROM_GLC; Mappable to FAO and IGBP systems	PA=86.52%, UA=90.06% (2017) F1 score = 85.49 %	Landsat TM, ETM+, OLI	Yang and Huang (2021)
The 30-m China Annual Tree Cover Dataset (CATCD)	30m	China	1985–2023	Tree cover, 0-100 %	Arboreal and bamboo forests (canopy cover > 20%, area > 0.067 ha), following NFI criteria but excluding special shrubs	Validated using the 9th NFI: R^2 = 0.93, RMSE = 6.9%, and MAE = 5.6%	Landsat TM, ETM+, OLI	Cai et al. (2024)
National Forest Inventory of China (NFI)	/	China	5-year intervals	/	Arboreal and bamboo forests (canopy cover > 20%, area > 0.067 ha); Shrubs in regions with annual precipitation < 400 mm (canopy cover > 30%, area > 0.067 ha), referred to as “special shrubs”	/	Statistical data	National Bureau of Statistics of China
China National Land Use/Cover Change (CNLUCC)	30m	China	1980, 1990, 1995, 2000, 2005, 2010, 2015, 2018, 2020, 2023	Land use/cover class	Areas dominated by trees (canopy cover ≥ 30%); Shrubland (canopy > 40%, height ≤ 2 m), open forest (canopy cover 10–30%); Orchards, nurseries, etc.	/	Landsat TM, ETM+, OLI	Resource and Environmental Science and Data Center in China
The Global 30-m Land-cover Dynamic Monitoring Product (GLC_FCS30D)	30m	Global	1985 – 2000: 5-year intervals; 2000 – 2022: yearly	Land use/cover class	Tree cover >15% at a pixel scale	PA=92.83%, UA=86.35% (2020)	Landsat TM, ETM+, OLI	Zhang et al. (2023)
Global Land Cover and Land Use Change (GLAD)	30m	Global	2000, 2020	Non-forest, 0; Forest, 1	Tree height ≥5 m at a pixel scale; Wildland, managed, and planted forests, agroforestry, orchards, and natural tree regrowth	OA=97.4% (2000) OA=97.2% (2020)	Landsat TM, ETM+, OLI	Potapov et al. (2022)
The ESA’s Climate Change Initiative: Land Cover (ESA CCI-LC)	300m	Global	1992–2022	Land use/cover class	Areas dominated by trees (canopy cover > 15%); Mixed forests, mosaic tree–shrub areas, and flooded forests	OA=75.4% (Land cover 2015)	MERIS, AVHRR, SPOT-V, etc.	ESA Climate Change Initiative

Continued

Dataset	Resolution	Spatial coverage	Temporal coverage	Pixel value	Forest definition	Reported accuracy	Satellite imagery	Source
Hansen Global Forest Change (GFC): tree cover 2000; forest change 2001-2024	30m	Global	2000–2024	Tree cover, 0-100 %	Canopy closure of vegetation taller than 5 m	/	Landsat ETM+	Hansen et al. (2013)
ESA WorldCover	10m	Global	2020, 2021	Land use/cover class	Areas dominated by trees (tree cover ≥ 10%); Areas planted with trees for afforestation purposes and plantations; Some tree-covered areas seasonally or permanently flooded with fresh water except for mangroves	PA= 89.9%, UA= 80.8% (2020)	Sentinel-1, Sentinel-2	Zanaga et al. (2021)
GlobeLand30	30m	Global	2000, 2010, 2020	Land use/cover class	Areas dominated by trees (tree cover >30%); Open woodland (tree cover 10%-30%)	UA=83.58% (2010)	Landsat TM, ETM+, OLI; Huan Jing-1; Gaofen-1	Chen et al. (2015)
Finer Resolution Observation of Global Land Cover (FROM-GLC30)	30m	Global	2010, 2015, 2017	Land use/cover class	Tree cover > 15% and tree height > 3 m	PA=76.45%, UA=80.49 (2010)	Landsat TM/ETM+	Yu et al. (2013)
Global Annual Fractional Tree Cover Dataset (GLOBMAP FTC)	250m	China	2000–2022	Tree cover, 0-100 %	Grid cells dominated by trees; Their proportion was used to estimate tree cover	Validated using 62,022 USGS samples (circa 2010): $R^2 = 0.73$, MAE = 10.55%, and RMSE = 17.98%	MOD09Q1, MOD09A1	Liu et al. (2024)

Notes: (1) This table provides an overview of major high-resolution forest datasets covering China, with a few medium-resolution datasets. (2) Most products report accuracy for the forest class, whereas the ESA CCI-LC product provides an overall accuracy of land use and land cover. OA refers to overall accuracy; PA to producer accuracy; and UA to user accuracy. F1 score represents the harmonic mean of PA and UA. (3) Zhang et al. (2020) evaluated the 2000 forest product of GFC in China and reported an overall accuracy of 94.5%.

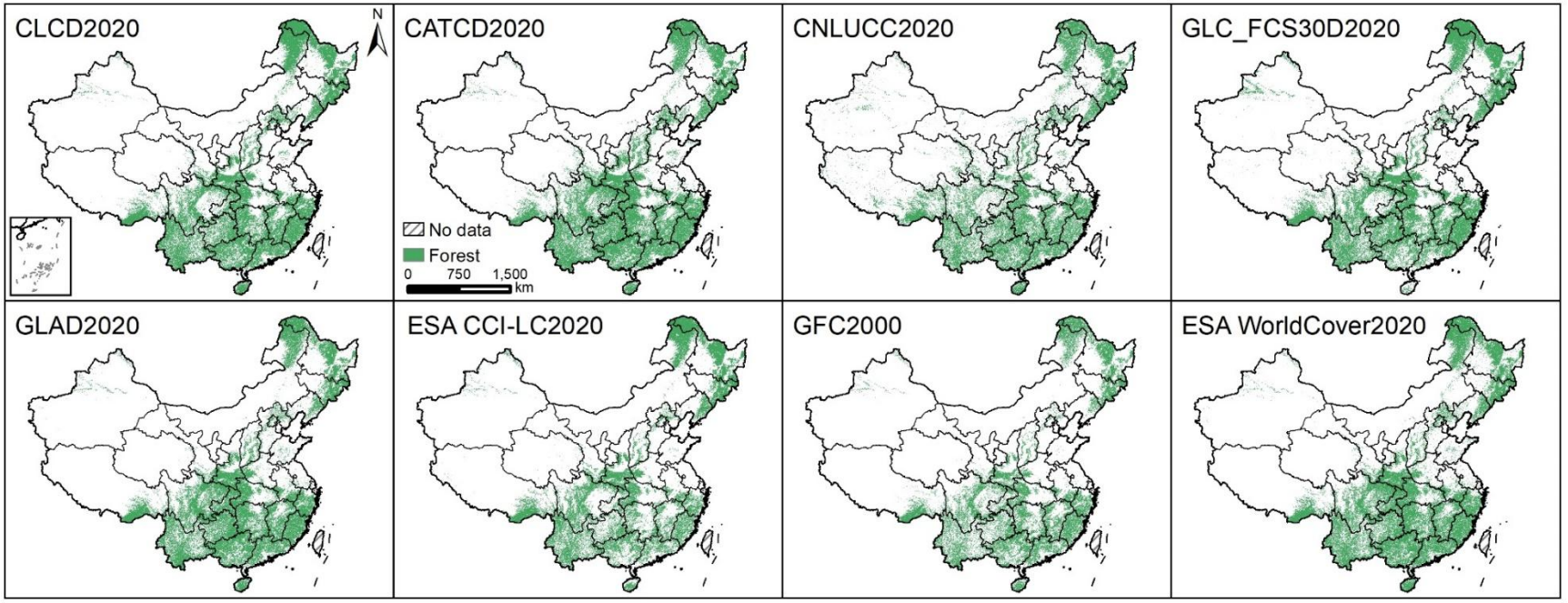


Fig. E1 Distribution of forest across different datasets.

Reference

Abadie, Alberto, and Jann Spiess. 2022. Robust post-matching inference. *Journal of the American Statistical Association* 117 (538): 983-995.

Araujo, Rafael, Juliano Assunção, and Arthur Amorim Bragança. 2023. The effects of transportation infrastructure on deforestation in the amazon: A general equilibrium approach. *World Bank Policy Research Working Paper*.

Asher, Sam, Teevrat Garg, and Paul Novosad. 2020. The ecological impact of transportation infrastructure. *The Economic Journal* 130 (629): 1173-1199.

Autor, David H. 2003. Outsourcing at will: The contribution of unjust dismissal doctrine to the growth of employment outsourcing. *Journal of Labor Economics* 21 (1): 1-42.

Baehr, Christian, Ariel BenYishay, and Brad Parks. 2023. Highway to the forest? Land governance and the siting and environmental impacts of Chinese government-funded road building in Cambodia. *Journal of Environmental Economics and Management* 122: 102898.

Barber, Christopher P., Mark A. Cochrane, Carlos M. Souza Jr, and William F. Laurance. 2014. Roads, deforestation, and the mitigating effect of protected areas in the Amazon. *Biological conservation* 177: 203-209.

Baum-Snow, Nathaniel, Loren Brandt, J Vernon Henderson, Matthew A. Turner, and Qinghua Zhang. 2017. Roads, railroads, and decentralization of Chinese cities. *Review of Economics and Statistics* 99 (3): 435-448.

Blanco, Víctor, Justo Puerto, and Ana B. Ramos. 2011. Expanding the Spanish high-speed railway network. *Omega* 39 (2): 138-150.

Busch, Jonah, and Kalifi Ferretti-Gallon. 2017. What drives deforestation and what stops it? A meta-analysis. *Review of Environmental Economics and Policy* 11 (1): 3-23.

Cai, Yaotong, Xiaocong Xu, Peng Zhu, Sheng Nie, Cheng Wang, Yujiu Xiong, and Xiaoping Liu. 2024. Unveiling spatiotemporal tree cover patterns in China: The first 30 m annual tree cover mapping from 1985 to 2023. *ISPRS Journal of Photogrammetry and Remote Sensing* 216: 240-258.

Campos, Javier, and Gines De Rus. 2009. Some stylized facts about high-speed rail: A review of HSR experiences around the world. *Transport policy* 16 (1): 19-28.

Chang, Zheng, and Jin Murakami. 2019. Transferring land-use rights with transportation infrastructure extensions. *Journal of Transport and Land Use* 12 (1): 1-19.

Chen, Jun, Jin Chen, Anping Liao, Xin Cao, Lijun Chen, Xuehong Chen, Chaoying He, Gang Han, Shu Peng, and Miao Lu. 2015. Global land cover mapping at 30 m resolution: A POK-based operational approach. *ISPRS Journal of Photogrammetry and Remote Sensing* 103: 7-27.

Chen, Zhenhua, Yulong Zhou, and Kingsley E. Haynes. 2021. Change in land use structure in urban China: Does the development of high-speed rail make a difference. *Land Use Policy* 111: 104962.

da Silva, Carlos Fabricio Assunção, Mauricio Oliveira de Andrade, Alex Mota dos Santos, and Silas Nogueira de Melo. 2023. Road network and deforestation of indigenous lands in the Brazilian Amazon. *Transportation Research Part D: Transport and Environment* 119: 103735.

das Neves, Patricia Bittencourt Tavares, Claudio José Cavalcante Blanco, André Augusto Azevedo Montenegro Duarte, Filipe Bittencourt Souza das Neves, Isabela Bittencourt Souza das Neves, and Marcelo Henrique de Paula dos Santos. 2021. Amazon rainforest deforestation influenced by clandestine and regular roadway network. *Land Use Policy* 108: 105510.

De Chaisemartin, Clément, and Xavier d'Haultfoeuille. 2020. Two-way fixed effects estimators with heterogeneous treatment effects. *American economic review* 110 (9): 2964-2996.

De Soyres, Francois, Alen Mulabdic, Siobhan Murray, Nadia Rocha, and Michele Ruta. 2019. How much will the Belt and Road Initiative reduce trade costs? *International Economics* 159: 151-164.

Dong, Lei, Rui Du, Matthew Kahn, Carlo Ratti, and Siqi Zheng. 2021. “Ghost cities” versus boom towns: Do China's high-speed rail new towns thrive? *Regional Science and Urban Economics* 89: 103682.

Dreher, Axel, Andreas Fuchs, Bradley Parks, Austin Strange, and Michael J. Tierney. 2022. Banking on Beijing: The aims and impacts of China's overseas development program. Cambridge University Press, 200 pp.

ESA. 2017. Land Cover CCI Product User Guide Version 2. Technical Report.

Foster, Andrew D., and Mark R. Rosenzweig. 2003. Economic growth and the rise of forests. *The Quarterly Journal of Economics* 118 (2): 601-637.

Hansen, Matthew C., Peter V. Potapov, Rebecca Moore, Matt Hancher, Svetlana A. Turubanova, Alexandra Tyukavina, David Thau, Stephen V. Stehman, Scott J. Goetz, and Thomas R. Loveland. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342 (6160): 850-853.

Hou, Yifeng, Yaning Chen, Jianli Ding, Zhi Li, Yupeng Li, and Fan Sun. 2022. Ecological impacts of land use change in the arid Tarim River Basin of China. *Remote Sensing* 14 (8): 1894.

Jiang, Yang. 2019. Competitive partners in development financing: China and Japan expanding overseas infrastructure investment. *The Pacific Review* 32 (5): 778-808.

Kaczan, David J. 2020. Can roads contribute to forest transitions? *World Development* 129: 104898.

Lawrence, Martha, Richard Bullock, and Ziming Liu. 2019. China's high-speed rail development. World Bank Publications.

Liu, Yang, Ronggao Liu, Jilong Chen, Xuexin Wei, Lin Qi, and Lei Zhao. 2024. A global annual fractional tree cover dataset during 2000–2021 generated from realigned MODIS seasonal data. *Scientific Data* 11 (1): 832.

Losos, Elizabeth Claire, Alexander Pfaff, Lydia Pauline Olander, Sara Mason, and Seth Morgan. 2019. Reducing environmental risks from belt and road initiative investments in transportation infrastructure. *World Bank Policy Research Working Paper* (8718).

National Development and Reform Commission of China. 2016. Plan for Medium- and Long-term Railway Network in China. Retrieved from https://www.ndrc.gov.cn/xxgk/zcfb/ghwb/201607/t20160720_962188.html.

Navalkar, Aniket, Sreeja Jaiswal, M. Pranadh, and T. Jayaraman. 2023. Impact of railways on land use and land cover change: Evidence from India. *Transportation Research Part D: Transport and Environment* 124: 103933.

Niu, Fangqu, Zhongling Xin, and Dongqi Sun. 2021. Urban land use effects of high-speed railway network in China: A spatial spillover perspective. *Land Use Policy* 105: 105417.

People's Daily. 2019. Have you ever wondered why China builds high-speed railway tracks on bridges? (March 21). People's Daily. Retrieved from http://en.people.cn/n3/2019/0321/c90000-9559114.html.

Pfaff, Alexander, Juan Robalino, Eustaquio J. Reis, Robert Walker, Stephen Perz, William Laurance, Claudio Bohrer, Steven Aldrich, Eugenio Arima, and Marcellus Caldas. 2018. Roads & SDGs, tradeoffs and synergies: learning from Brazil’s Amazon in distinguishing frontiers. *Economics* 12 (1): 20180011.

Pfaff, Alexander, Juan Robalino, Robert Walker, Steven Aldrich, Marcellus Caldas, Eustaquio Reis, Stephen Perz, Claudio Bohrer, Eugenio Arima, and William Laurance. 2007. Road investments, spatial spillovers, and deforestation in the Brazilian Amazon. *Journal of Regional Science* 47 (1): 109-123.

Potapov, Peter, Matthew C. Hansen, Amy Pickens, Andres Hernandez-Serna, Alexandra Tyukavina, Svetlana Turubanova, Viviana Zalles, Xinyuan Li, Ahmad Khan, and Fred Stolle. 2022. The global 2000-2020 land cover and land use change dataset derived from the Landsat archive: first results. *Frontiers in Remote Sensing* 3: 856903.

Santos, Diogo Corrêa, Pedro Walfir Martins Souza-Filho, Wilson da Rocha Nascimento Jr, Gustavo Freitas Cardoso, and Jorge Filipe dos Santos. 2020. Land cover change, landscape degradation,

and restoration along a railway line in the Amazon biome, Brazil. *Land Degradation & Development* 31 (15): 2033-2046.
Souza-Rodrigues, Eduardo. 2019. Deforestation in the Amazon: A unified framework for estimation and policy analysis. *The Review of Economic Studies* 86 (6): 2713-2744.
State Council of China. 2021. Fourteenth Five-Year Plan for the Development of a Modern Comprehensive Transportation System (State Decree No. 27). Retrieved from https://xxgk.mot.gov.cn/2020/jigou/zhghs/202201/t20220119_3637245.html.
Sun, Liyang, and Sarah Abraham. 2021. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics* 225 (2): 175-199.
Vessali, Kaveh V. 1996. Land use impacts of rapid transit: A review of the empirical literature. *Berkeley Planning Journal* 11 (1).
Viana, Virgilio M., Mariano C. Cenamo, Mariana N. Pavan, Gabriel C. Carrero, and Matthew D. Quinlan. 2008. Railroads in the Amazon: A key strategy for reducing deforestation. *Carbon & Climate Law Review* 2 (3): 292-299.
Wang, Chao, Ming K. Lim, Xin Yi Zhang, Long Feng Zhao, and Paul Tae-Woo Lee. 2020. Railway and road infrastructure in the Belt and Road Initiative countries: Estimating the impact of transport infrastructure on economic growth. *Transportation Research Part A: Policy and Practice* 134: 288-307.
Wang, Liqin. 2023. China–Japan competition in infrastructure investment in Southeast Asia: A two-level analysis. *Chinese Political Science Review* 8 (4): 527-552.
Wang, Lvhua, Yongxue Liu, Chao Sun, and Yahui Liu. 2016. Accessibility impact of the present and future high-speed rail network: A case study of Jiangsu Province, China. *Journal of Transport Geography* 54: 161-172.
Wenner, Fabian, and Alain Thierstein. 2022. High speed rail as urban generator? An analysis of land use change around European stations. *European Planning Studies* 30 (2): 227-250.
Wu, Rong, Yingcheng Li, and Shaojian Wang. 2022. Will the construction of high-speed rail accelerate urban land expansion? Evidences from Chinese cities. *Land Use Policy* 114: 105920.
Wurm, Michael, Jan Goebel, Gert G. Wagner, Matthias Weigand, Stefan Dech, and Hannes Taubenböck. 2021. Inferring floor area ratio thresholds for the delineation of city centers based on cognitive perception. *Environment and Planning B: Urban Analytics and City Science* 48 (2): 265-279.
Xiao, Chiwei, Yi Wang, Mingyan Yan, and Jeffrey Chiwuikem Chiaka. 2024. Impact of cross-border transportation corridors on changes of land use and landscape pattern: A case study of the China-Laos railway. *Landscape and Urban Planning* 241: 104924.
Xinhua News. 2023a. China focus: China's first HSR designed for 350 km/h delivers 340 mln passenger trips in 15 years (August 1). Xinhua News. Retrieved from https://english.news.cn/20230801/2f48452fa1b84a778704b00cff4ca3e0/c.html.
Xinhua News. 2023b. Jakarta-Bandung high-speed railway outstanding example of Indonesia-China cooperation (September 4). Xinhua News. Retrieved from https://eng.yidaiyilu.gov.cn/p/0RLE2GG9.html.
Xu, Xinliang, Jiyuan liu, Shuwen Zhang, Rendong Li, Changzhen Yan, and Shixin Wu. 2018. China's Multi-Period Land Use Land Cover Remote Sensing Monitoring Dataset (CNLUCC). in: Resource and Environment Data Cloud Platform (Ed.), Beijing, China.
Xu, Yu, Megan Price, Biao Yang, Kai Zhang, Nan Yang, Xiaoxin Tang, Jianghong Ran, Yin Yi, and Bin Wang. 2022. Have China's national forest reserves designated since 1990 conserved forests effectively? *Journal of Environmental Management* 306: 114485.
Yang, Jie, and Xin Huang. 2021. The 30 m annual land cover dataset and its dynamics in China from 1990 to 2019. *Earth System Science Data* 13 (8): 3907-3925.
Yu, Le, Jie Wang, and Peng Gong. 2013. Improving 30 m global land-cover map FROM-GLC with time series MODIS and auxiliary data sets: a segmentation-based approach. *International Journal of Remote Sensing* 34 (16): 5851-5867.

Zanaga, Daniele, Ruben Van De Kerchove, Wanda De Keersmaecker, Niels Souverijns, Carsten Brockmann, Ralf Quast, Jan Wevers, Alex Grosu, Audrey Paccini, Sylvain Vergnaud, Oliver Cartus, Maurizio Santoro, Steffen Fritz, Ivelina Georgieva, Myroslava Lesiv, Sarah Carter, Martin Herold, Lin Lin Li, Nandin Erdene Tsendbazar, Fabrizio Ramoino, and Olivier Arino. 2021. ESA WorldCover 10 m 2020 v100.

Zhang, Di, Hao Wang, Xu Wang, and Zhi Lü. 2020. Accuracy assessment of the global forest watch tree cover 2000 in China. *International Journal of Applied Earth Observation and Geoinformation* 87: 102033.

Zhang, Xiao, Tingting Zhao, Hong Xu, Wendi Liu, Jinqing Wang, Xidong Chen, and Liangyun Liu. 2023. GLC_FCS30D: The first global 30-m land-cover dynamic monitoring product with a fine classification system from 1985 to 2022 using dense time-series Landsat imagery and continuous change-detection method. *Earth System Science Data Discussions* 2023: 1-32.

Zheng, Longfei, Fenjie Long, Zheng Chang, and Jingsong Ye. 2019. Ghost town or city of hope? The spatial spillover effects of high-speed railway stations in China. *Transport Policy* 81: 230-241.

Zhou, Jiabei, Shuang Gao, and Shaojian Wang. 2024. Impact and mechanisms of high-speed rail construction on carbon emissions: A quasi-natural experiment in China. *Applied Geography* 172: 103436.