Green Gains from Connectivity: Highway Expansion and Forest Quality[†]

Xingjian Ding[a] Yumin Hu[b] Shilei Liu[c] Cong Peng[d*]
Jintao Xu[d] Mingzhi (Jimmy) Xu[a] Qinghua Zhang[e]

Roads are often linked to deforestation in frontier settings, but their effects in managed forests are less well understood. We investigate the intensive-margin response of forest outcomes to roads using China's 2000–2010 expressway rollout as the setting. We construct a unique panel from the National Forest Inventory covering over 18,000 geo-located plots with ground-based measures of standing timber volume, which capture tree-level outcomes that satellite-based measures typically miss. Using long differences and standard instrumental variables, we find that moving 10 km closer to an expressway increases timber volume by approximately 2%, with gains concentrated within 1–20 km of new roads. The implied biomass increase equals 217–552 million tons of CO_2, comparable to Canada's annual emissions at the upper bound. Mechanism evidence and a spatial equilibrium model show that improved downstream market access strengthens incentives for investment and specialization in forestry under strict forest land-use controls, highlighting a role for transport infrastructure in promoting sustainable growth.

JEL Codes: R42, Q54, Q23, H41, R13
Keywords: Transportation Infrastructure; Climate Change; Forest; Market Access; Labor Specialization

[†]This paper's analyses are solely those of the authors and draw in part on National Forest Inventory (NFI) data administered by the National Forestry and Grassland Administration. The views expressed do not necessarily reflect those of the Administration, which had no role in the study and bears no responsibility for the results. We are grateful to seminar participants at the Summer Meeting of the Urban Economics Association of China, the Göttingen Workshop, the SPACE Workshop, and the 2025 Jinan–SMU Conference on Urban and Regional Economics for helpful comments and suggestions. Mingzhi Xu gratefully acknowledges financial support from the National Natural Science Foundation of China (Grant Nos. 72322007 and 72503230).

[a]Institute of New Structural Economics, Peking University, Beijing, China
[b]Center for Transnational Studies and Institute of International Economics, School of Economics, Nankai University, Tianjin, China
[c]School of Ecology and Environment, Renmin University of China, Beijing, China
[d]China Center for Economic Research, Institute of South–South Cooperation and Development, National School of Development, Peking University, Beijing, China
[e]Guanghua School of Management, Peking University, Beijing, China
[*]Corresponding author. Email address: congpeng@nsd.pku.edu.cn

1 Introduction

Road infrastructure expands market access and can stimulate economic development (Faber, 2014; Ghani, Goswami and Kerr, 2015; Baum-Snow et al., 2020; Fretz, Parchet and Robert-Nicoud, 2021; Asher and Novosad, 2020), but its effects depend critically on complementary institutions such as property rights and regulatory enforcement (Burgess et al., 2015; Esfahani and Ramírez, 2003; Banerjee, Duflo and Qian, 2020; Jedwab and Storeygard, 2021; Peng and Wang, 2026). Weak institutions may cause improved connectivity to accelerate resource extraction and deforestation (Burgess et al., 2012). A large literature—largely grounded in frontier settings where forests are weakly managed—shows that roads accelerate deforestation at the extensive margin (Souza-Rodrigues, 2019; Baehr, BenYishay and Parks, 2023; Engert et al., 2024; Gollin and Wolfersberger, 2025; Lundberg et al., 2025).

However, this frontier emphasis captures only part of the global forest landscape. Roughly half of the world's forests are managed: more than two billion hectares are covered by long-term management plans, with coverage rising steeply with development—from about one-fifth of forests in parts of the tropics to near-universal coverage in Europe.[1] In these managed settings, roads lower trade costs in ways that can simultaneously increase incentives for clearing and extraction while also raising returns to sustained forest management. As a result, the net effect of infrastructure expansion on forest outcomes is theoretically ambiguous (Asher, Garg and Novosad, 2020; Kaczan, 2020).

This paper highlights an intensive-margin channel that has received limited attention in the transport–environment literature: the effect of road connectivity on the quality of forests that remain forested under management institutions. This margin is increasingly relevant as land-based carbon sinks remain among the few scalable options for offsetting hard-to-abate emissions, understanding how transport infrastructure shapes the quality of managed forests has broader implications for developing countries that are simultaneously expanding transport networks and extending formal forest management.[2] With limited scope for further expansion of forest area, improvements in forest quality—measured by

[1] Food and Agriculture Organization of the United Nations (2020) reports that 2.05 billion hectares of forest were under long-term management plans in 2020, with coverage rates of 17% in South America, 24% in Africa, and 96% in Europe. The same report estimates total global forest area at 4.06 billion hectares. The Global Forest Review reports that about 41% of global forests are managed natural forests and 11% are planted forests, compared to 48% unmanaged natural forests (Goldman, 2024). The underlying high-resolution global forest management map is described in Lesiv et al. (2022).

[2] Using China as an example, if its official forestation targets are achieved, this could offset approximately 43% of its anticipated residual fossil CO_2 emissions by 2060. A recent reconciliation of China's national greenhouse gas inventory estimates that this corresponds to net removals of about 1.3 Gt CO_2 yr^{-1} from the land-use, land-use change, and forestry (LULUCF) sector under extended forestation targets (He et al., 2024).

standing timber volume per unit of land—become critical. In this paper, we provide an empirical analysis of the causal effect of highway expansion on forest quality and examine the mechanisms underlying this relationship.

Two empirical challenges arise. The first is data availability. Because forests are often located far from human settlements, representative data on forest quality are scarce. Although satellite imagery effectively monitors changes in forest area (Burgess et al., 2012), it fails to capture changes underneath the canopy (Baragwanath and Shinde, 2025). This includes essential metrics of forest productivity, such as standing timber volume, which is the proper indicator of forest capital and carbon storage. To address this gap, we use China's National Forest Inventory (NFI), which covers over 18,000 geo-located plots spaced at 3–5 km intervals across 11 major forestry provinces. Each plot, about 0.067 hectares in size, is surveyed every five years. We build a panel from the 1999–2003 and 2009–2013 waves, with ground-based measures of tree volume, height, diameter, and canopy. We then link these data to detailed maps of national highways constructed between 2000 and 2010, yielding a unique panel to study how transport access shapes forest quality.

The second challenge is causal identification. Expressways are not randomly routed: planners may steer corridors toward areas with high expected economic returns, while mandatory environmental impact assessments can divert routes away from the most protected or ecologically sensitive forests. As a result, simple comparisons may confound highway effects with endogenous route selection and can bias estimated impacts toward zero. To address these concerns, we combine the panel structure of the data with an instrumental-variables strategy. We first estimate long-difference specifications that compare changes in forest outcomes around newly opened expressways across plots with differential exposure to construction. We then instrument each plot's distance to the nearest highway using two predetermined predictors of highway placement that are standard in the literature.

The first instrument is distance to a simulated least-cost-path (LCP) expressway network constructed from terrain slope and other physical impediments. This IV is widely used in the literature (for examples, see Faber (2014); Donaldson (2018); Fenske, Kala and Wei (2021)). The network predicts where expressways would be routed if planners connected pre-specified endpoint cities in the national expressway blueprint while minimizing engineering costs. Conditional on rich geographic controls and proximity to planned nodes, the identifying variation therefore comes from how cost-minimizing corridors traverse the rural hinterland, rather than from local economic conditions or forest dynamics.

The second instrument exploits path dependence in infrastructure construction by using distance to China's national road network as of 1962, following Baum-Snow et al. (2020). This IV follows a long line of papers that use historical roads to instrument modern road net-

works (for examples, see Baum-Snow (2007); Michaels (2008); Duranton and Turner (2011); Jedwab and Moradi (2016)). Many modern expressways were built by upgrading or paralleling earlier roads, so proximity to the 1962 network lowers the marginal cost of subsequent highway placement and predicts later expressway access. Because the 1962 network predates the modern expressway program and was shaped by mid-century administrative and military considerations, its spatial distribution is unlikely to reflect contemporary economic or environmental trends affecting forest outcomes. Under the identifying assumption that these predetermined predictors affect forest outcomes only through their influence on modern highway placement—after conditioning on geography and pre-period characteristics—the two instruments provide complementary and plausibly exogenous variation in highway access.

Our estimates indicate that highway expansion raises forest biomass along the intensive margin. Moving 10 km closer to a newly constructed highway increases standing timber volume by approximately 2%. The effects are concentrated within 1–20 km of new roads, where forest volume grows about 15% faster than in more distant locations, relative to an average growth of 37% over the period. We exclude plots within 0–1 km to avoid mechanical land-conversion effects. Gains are larger for plots connected earlier, consistent with gradual biological growth, and for steeper terrain, where the opportunity cost of land is lower. These increases imply substantial carbon benefits: the implied gains in forest biomass between 2000 and 2010 sequestered 216.9–551.9 million tons of CO_2, depending on reasonable assumptions about the conversion from standing timber volume to biomass, comparable to the annual emissions of Spain (lower bound) and Canada (upper bound). Our results are robust to a wide range of alternative specifications. We show that the estimates are not driven by historical socioeconomic advantages along planned routes, correlated economic or demographic growth, or the calibration of trade elasticity in the market access measure. The findings are also insensitive to the treatment of zero or missing forest measurements and to alternative sample definitions addressing pre-existing highways and roadside plots.

What are the underlying mechanisms? They lie in the interaction between land-use institutions and enhanced market access brought about by new highways. Two institutional features are central. First, stringent forest land-use regulations sharply limit land conversion, channeling the benefits of improved access into forest management rather than forest loss. Second, China's collective forest-tenure reforms strengthened land rights, giving households and specialized operators incentives to respond to new market opportunities. Two pieces of evidence are consistent with this interpretation. (i) Satellite data show no systematic reduction in forest land area, and administrative records indicate that highway-connected counties experience higher afforestation rates. (ii) Forest quality gains are concentrated in initially collectively operated plots subject to tenure reform, whereas state and initially

privately operated plots exhibit muted responses.

With the institutional context as the enabler, market access brought by highway expansion to rural forest areas drives growth in forest quality. Following Herzog (2021), we empirically decompose the growth in forest quality into the market access effect of roads while holding income constant and the effect of income while holding roads constant, and further divide income into that generated by timber-related downstream industries versus other manufacturing industries. We find that both access to timber-related downstream industries due to road improvement and the growth of timber-related downstream industries drive improvements in forest volumes. The results suggest that market access to downstream industries plays a key role. Complementary evidence reinforces this mechanism. Firm registration records reveal a sharp increase in timber-cultivation firm entry in counties within 20 km of new highways, suggesting that better access attracts specialized forestry operators through improving incentives. Highway expansion also increases access to labor and enables a selection effect that promotes labor specialization in forest management. Household-level data on farmer activities show that forestry-experienced households double timber sales, with increases of RMB 1,113, while non-forestry households add 35.4 migration days.[3] The household specialization patterns suggest that highways reallocate workers toward sectors where they hold comparative advantage.

Motivated by these empirical findings, we develop a quantitative spatial equilibrium model of trade and migration across regions to rationalize the observed adjustments and quantify the relative importance of different channels. The model features forest and urban sectors and heterogeneous households differentiated by their comparative advantage in forest production. Forest area is held fixed, so changes in forest sales reflect intensive-margin improvements in forest quality and management rather than land expansion. We calibrate the model to the year-2000 Chinese economy, covering 287 cities that together account for the vast majority of China's population and economic activity. At the aggregate level, highway expansion raises forest sales by about 17 percent. An exact-hat decomposition highlights the relative strength of the underlying mechanisms: improved access to downstream markets is the dominant driver, amplifying forest sales by roughly 40 percent in proportional terms, while cheaper urban inputs and labor reallocation contribute more modest gains of about 3 and 0.2 percent, respectively. These gross gains are partially offset by higher land rents, which erode the aggregate effect by about 20 percent.

This paper makes three contributions to the literature. First, it offers a complementary perspective on the role of transport infrastructure in nature conservation. A large

[3]Highways may also encourage specialized forestry workers to cover a larger area of forestry land, including neighboring and even distant counties.

body of work shows that roads often accelerate deforestation by lowering access costs and facilitating land conversion at the extensive margin, particularly in frontier forest settings (Souza-Rodrigues, 2019; Asher, Garg and Novosad, 2020; Baehr, BenYishay and Parks, 2023; Engert et al., 2024; Gollin and Wolfersberger, 2025; Lundberg et al., 2025). We show, however, that this mechanism need not apply to managed forests—which account for roughly half of global forest area—where land-use institutions may credibly constrain conversion. In such settings, improved connectivity can raise forest quality on the intensive margin. The limited attention to this channel in the literature likely reflects the scarcity of high-quality data on tree-level outcomes. Using a unique panel from China's National Forest Inventory, we measure standing timber volume and canopy structure—the relevant measures of the intensive-margin response—and document a quality-upgrading response that would be largely missed in analyses focused on forest cover alone (Baragwanath and Shinde, 2025).

Second, we show that this quality response operates primarily through changes in market access. By strengthening links to downstream timber-processing markets, improved connectivity raises the returns to intensive forest management and quality upgrading. A quantitative spatial equilibrium model helps organize these patterns and underscores downstream market access as the dominant channel, with smaller roles for input access and labor reallocation. More broadly, these results contribute to the literature on market integration and environmental outcomes by clarifying how connectivity can reshape incentives for investment in natural capital (Araujo et al., 2009; Xie, Berck and Xu, 2016; Baragwanath and Bayi, 2020; Blackman et al., 2017; Walker et al., 2025). Particularly improved connectivity increases incentives for intensive forest management only when producers hold secure and transferable claims and can reorganize production in response to market opportunities.

Finally, we extend the carbon accounting of transport infrastructure beyond loss-based metrics that focus exclusively on deforestation. Existing work largely emphasizes the carbon costs of roads in frontier settings (Souza-Rodrigues, 2019; Vilela et al., 2020). By incorporating intensive-margin changes in forest biomass, we show that infrastructure can generate net carbon benefits when improved connectivity strengthens incentives for higher forest quality. This reframes the climate implications of transport infrastructure by emphasizing that their carbon consequences depend not only on changes in forest area, but also on how connectivity reshapes incentives to accumulate carbon-dense forest capital on land that remains forested.

Section 2 describes the background and data. Section 3 presents the identification strategy and estimation results. Section 4 sets up the quantitative spatial model. Section 5 conducts counterfactual analysis. Section 6 concludes.

2 Background and Data

This section provides a brief background on China's forestry institutions and introduces the key datasets used in the analysis. We then describe our core plot-level panel from the National Forest Inventory and how we link it to the national highway rollout to study changes in forest quality over the decade.

2.1 Background

Forestry, forest quality and carbon sequestration Forests are central to nature-based solutions for mitigating climate change because carbon sequestration scales with standing timber volume and growth. Traditional attention has focused on preventing deforestation, especially in tropical regions such as the Amazon, which sometimes leads to the misconception that cutting trees is inherently harmful. In fact, sustainable harvesting can be essential for forests to continue growing and storing carbon, as harvested wood retains carbon in long-lived products when properly processed, while cutting stimulates regrowth that converts more atmospheric CO_2 into sequestered biomass (Jin, Yi and Xu, 2020). Therefore, increasing the lifetime value of forest volume is crucial for long-term climate mitigation. Given a fixed horizon, such as China's 2060 carbon neutrality goal, higher forest quality implies greater timber volume and more CO_2 absorbed. Both increasing forest cover at the extensive margin and improving forest quality at the intensive margin matter. This paper focuses on the latter.

China's forest ownership and management, incentives and specialization In China, plantations are the primary engine of timber growth. According to the 6th National Forest Inventories (NFIs), total forest area is about 169 Mha in 1999–2003, including 115.8 Mha of natural forests and 53.3 Mha of plantation forests. By the 8th NFI (2009–2013), total forest area further rose to about 191 Mha, including 121.8 Mha of natural forests and 69.3 Mha of plantations, reflecting a decade of intensive afforestation and large-scale plantation establishment across China.[4] Forest expansion has been extensive, but forest productivity and quality hinge on plantation management and the labor that sustains silvicultural work. While natural forests under conservation mandates are typically managed by state-owned enterprises (SOEs), which carry out thinning, tending, and pest control to maintain and improve stands under binding protection rules (USDA Foreign Agricultural Service, 2017;

[4]Summary data for the 6th and 8th NFIs are sourced from the National Forest Resources Reports (State Forestry Administration, 2010, 2014), respectively. To ensure temporal consistency, forest area follows the definition of "Forested Land" (*You Lin Di* in Chinese), which requires (i) canopy cover $\geq$ 20% and (ii) a minimum contiguous area of 0.067 ha.

Standing Committee of the National People's Congress, 2019; Farooq et al., 2021), plantations are largely collectively owned, with timber income a key revenue source for rural households.

The Collective Forest Tenure Reform (referred to as CFTR hereafter), launched in the early 2000s, marked a major shift from collective to household-based management. Before the reform, collective ownership entailed weak and uncertain rights, discouraging incentives to make sustained investments to enhance forest quality. The CFTR clarified long-term use rights and allowed the transfer, leasing, and mortgaging of forest-use contracts by farmer households (Food and Agriculture Organization of the United Nations, 2012; He et al., 2020). This institutional change is reflected in tenure patterns observed in the NFIs: the share of collectively managed forests declined from 38% in the 6th NFI to 18% in the 8th NFI, while privately managed holdings more than doubled from 21% to 44% (He et al., 2020). By strengthening residual claims over returns, clearer property rights increased individuals' incentives to make silvicultural investments—such as thinning, replanting, and species upgrading—and encouraged the consolidation of fragmented plots into more efficient management units. Consistently, empirical studies document significant increases in timber output and replanting following tenure clarification, indicating a shift from short-term extraction toward long-term forest investment (Xie, Berck and Xu, 2016; He et al., 2020).

Beyond incentives alone, effective forest management also requires appropriate skills and operational scale. Individual farmers often lack the silvicultural expertise and financial capacity needed to maximize forest volume and quality. Forest tenure individualization has partly relaxed this constraint by fostering an active forestland rental market (Siikamäki, Xu and Ji, 2015; Yi et al., 2022), which facilitates the transfer and consolidation of management rights to specialized forestry operators—such as experienced farmers, former state forest farm employees, or other agents with the interest and capacity to manage larger forest holdings. By reallocating land to managers with a comparative advantage in forestry, the reform has promoted specialization in forest management, which we view as an important channel contributing to improvements in forest quality.

Strict laws of forest protection Following the 1998 mega-floods, forest conservation moved to the top of the policy agenda. National programs, including the NFPP and the Grain-for-Green (Sloping Land Conversion) Program, tightened land-use controls, strictly policed illegal logging, and expanded ecological (public-benefit) forests (World Bank Group, 2002; Bennett, 2008; Liu et al., 2008; Standing Committee of the National People's Congress, 2019; Wang et al., 2025). Quotas, cutting permits, and transport checkpoints were rigorously enforced, raising the cost of unlawful extraction and large-scale conversion (Standing

Committee of the National People's Congress, 2019; Preferred by Nature, 2021; China Perspectives, 2001). Under this regime, improved highway access is less likely to induce deforestation and more likely to channel effort toward lawful harvesting, better management, and increased forest volume at the intensive margin. Our evidence indicates that the expansion of highways did not result in a decrease in China's forest land area.

Highway expansion China's National Expressway Network Plan (NENP) launched in 2004 was designed to connect major cities and industrial hubs to facilitate trade and labor mobility, rather than to target forests per se (Faber, 2014; World Bank, 2007; Xu, 2015). The official plan aims to connect 245 cities, nearly all of which are principal cities of prefecture-level administrative units. These cities are typically distant from major forest areas and are often surrounded by cultivated land or built-up surfaces. An official map of the planned roads and targeted cities is shown in Section A.1.

During the period from 2000 to 2010, the total length of highways in China increased by about fivefold (China Transport Yearbook Press, 2000, 2011). By sharply reducing travel times and logistics costs, expressways act as a market integration shock for rural and forested regions. On the demand side, improved connectivity links plantations more closely to urban downstream markets, increasing the effective price received for timber and forest products. On the supply side, better access lowers the cost of obtaining forestry inputs and services, enabling rural producers to source seedlings, fertilizers, machinery, and professional silvicultural support more easily, which facilitates faster replanting and higher-quality forest production (USDA Foreign Agricultural Service, 2017; Farooq et al., 2021). Improved connectivity also expands opportunities for temporary urban employment, indirectly promoting labor specialization and household income diversification.

At the same time, strong environmental scrutiny during planning and environmental impact assessments typically routed new corridors away from the most strictly protected forest areas. As a result, new highways often traverse less critical or slower-growing forests, generating a selection pattern that poses an empirical challenge for identifying causal impacts on forest quality (Standing Committee of the National People's Congress, 2019; Preferred by Nature, 2021). Although some forest land was converted for rights-of-way, in our setting we find that the increase in standing timber volume and quality near improved access exceeds the localized forest loss from construction.

2.2 Data

The analysis relies on a rich set of data sources to evaluate how new highways affect forest quality and to identify the mechanisms behind these effects. The core dataset is the National

Forest Inventory (NFI), complemented by detailed information on highway networks and rasterized measures of travel speed and travel time. We further draw on auxiliary sources, including rural household surveys, enterprise registration records, and satellite-based data on land cover.

2.2.1 Forest Quality from the National Forest Inventory

We measure forest quality using data from the National Forest Inventory (NFI), a comprehensive ground-based monitoring system administered by the National Forestry and Grassland Administration. To examine the long-term impact of infrastructure, we construct a panel dataset linking plots from the 6th (1999–2003) survey wave to the 8th (2009–2013) wave. This period broadly coincides with China's rapid highway expansion.

The unit of observation is the sample plot. The NFI utilizes a systematic sampling design, with approximately 415,000 plots situated at the intersections of a national 1 km $\times$ 1 km grid. Grid selection is random and independent of land use. On each 0.0667-hectare plot, professional enumerators conduct detailed field measurements whenever measurable trees are present.[5]

Due to data availability, our analysis focuses on 11 major forestry provinces.[6] These provinces are highly representative, covering 53% of national forest land and 43% of standing timber volume, and spanning the diverse ecological conditions of China's four major forest regions. Within these regions, our dataset comprises the full universe of permanent sample plots (approximately 18,000) that were classified as arbor forest in both the 6th and 8th waves. We restrict our attention to arbor forests because they constitute the primary stock of forest capital, accounting for 81.3% of China's forest land and 97.2% of its standing timber volume.[7] This focus on the intensive margin (stable arbor forests) comes at the cost of excluding land conversion, but it enables us to analyze internal forest quality dynamics with a precision typically absent from remote-sensing-based studies.[8] The spatial distribution of our sample plots is shown in Panel (a) in Figure 1.

Our primary dependent variable is *timber volume* on each survey plot, calculated from

[5]To prevent strategic behavior, plot markers are hidden. Analysis of MODIS NDVI data confirms that vegetation trends near plots do not differ from surrounding areas (see Table A.4), supporting the sample's representativeness.

[6]The 11 major forestry provinces include Jilin, Inner Mongolia, Yunnan, Guangxi, Guangdong, Fujian, Jiangxi, Hubei, Guizhou, Shanxi, and Gansu.

[7]Data from the 6th NFI. Arbor forest land is defined as woodland primarily composed of arbor trees with canopy density above 20%, or stands that have grown for a specific number of years even if they do not yet meet the density threshold (State Forestry Administration of the People's Republic of China, 2004).

[8]We also use satellite-based land-use data to examine changes at the extensive margin; see the data description on land classifications.

tree-level measurements of height and diameter at breast height (DBH).[9] This metric captures the physical stock of forest capital and is the direct determinant of carbon storage and commercial value. We also analyze average DBH, tree height, and canopy coverage as supplementary indicators of forest quality. To mitigate the influence of outliers, all continuous variables are winsorized at the 1% level.

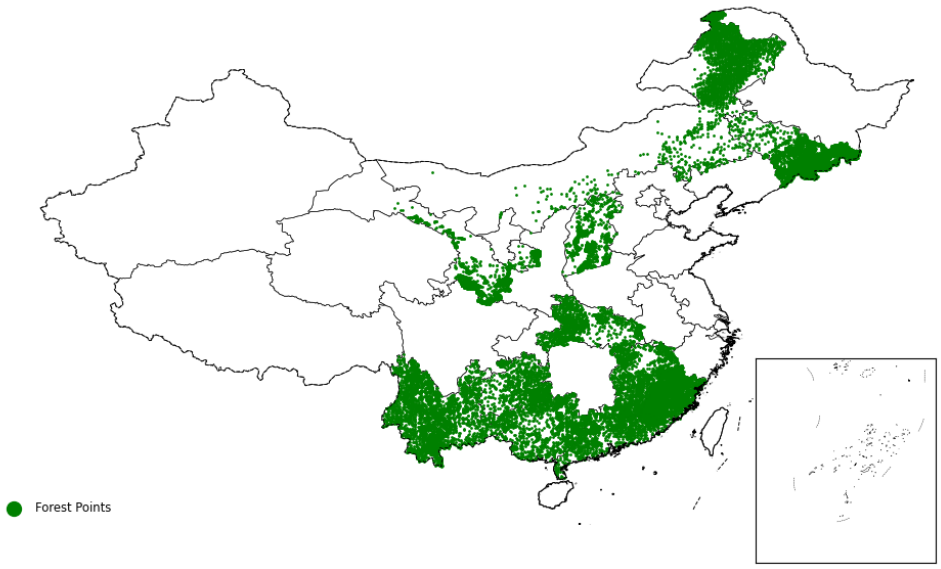
The richness of these ground-based measures offers a fundamental advantage over satellite-derived indicators, such as canopy cover or vegetation indices (e.g., NDVI, VFI). As illustrated in Panel (b) and Panel (c) in Figure 1, remote sensing primarily detects forest presence—the extensive margin—but often fails to capture structural changes underneath the canopy, the intensive margin (Baragwanath and Shinde, 2025). For instance, a stand of young, thin trees can exhibit the same canopy closure as a stand of mature, massive timber, yet the latter holds vastly more carbon and economic value. Satellite measures can also misclassify shrublands as loose forests or miss selective logging that thins the forest without breaking the canopy. By contrast, the NFI directly measures the volume and size of individual trees, enabling us to detect quality upgrading—such as tree growth, density management, and species shifts—that remote sensing would miss. This distinction is crucial in our context, where strict land-use regulations limit deforestation but market incentives drive intensive management.

Table A.1 reports summary statistics for the key forest variables. Forest quality improves substantially over the sample period. Average timber volume per plot increases from 6.04 to 7.67 m^3, a rise of 27%, corresponding to an increase from 90.6 to 115.0 m^3/ha. Canopy coverage increases from 56.2% to 62.2%, average tree height rises from 96.7 to 112.2 dm, and mean DBH remains of similar magnitude across survey waves. There is considerable heterogeneity across plots. The standard deviation of timber volume is comparable to its mean in both survey waves, and changes over time vary widely: while average log volume growth is positive (0.38), some plots experience sharp declines whereas others exhibit rapid growth. Similar dispersion characterizes changes in DBH, height, and coverage. This rich within-plot variation motivates the use of panel and long-difference specifications to identify how improved connectivity affects forest quality along the intensive margin.

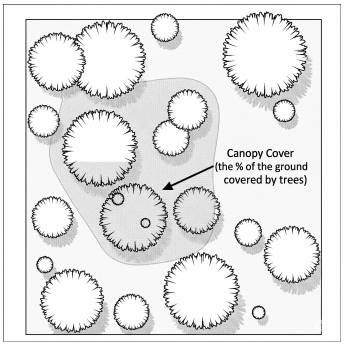
2.2.2 Highways and Inter-County Travel Times

We use the Transportation Networks Dataset from Ma and Tang (2024) to construct our primary measure of infrastructure access: the distance from each forest inventory plot to the

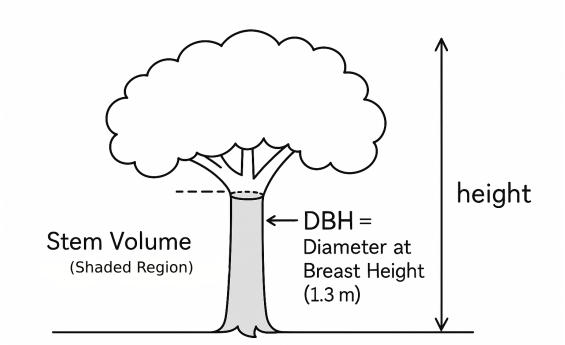
[9]The NFI protocol records volume only for trees exceeding specific thresholds, such as a DBH $\geq$ 5 cm. Plots with no trees meeting these criteria—for instance, newly planted areas or land that has recently undergone clear-cutting—are assigned a volume of zero.



(a) Forest Survey Points



(b) Covers



(c) Ground-based Measures

Figure 1: Key Metrics Used to Assess Forest Quality

Note: Panel (a) shows the spatial distribution of the survey points from NFI. Panel (b) depicts canopy cover, defined as the percentage of ground area shaded by tree crowns when viewed from above. Commonly derived from satellite imagery, canopy cover is widely used to assess whether a given area is forested. Panel (c) illustrates standing timber volume, a ground-based measure of forest quality derived from direct observations of individual trees. Timber volume is calculated using tree height and diameter at breast height (DBH, measured at 1.3 meters above ground). This metric reflects the physical structure and biomass of forests.

nearest highway. This dataset documents China's dramatic expressway expansion at a 1 km resolution. During our study period, the national highway network grew more than fivefold, expanding from approximately 11,600 km in 2000 to 65,000 km by the end of 2010. This massive infrastructure rollout significantly improved proximity to high-speed transport for remote forestry areas. As shown in Table A.1, the average distance from a forest plot to the nearest highway fell sharply from 174.6 km in 2000 to 61.4 km in 2010.

To examine how highways shape local markets and labor allocation, we construct a ma-

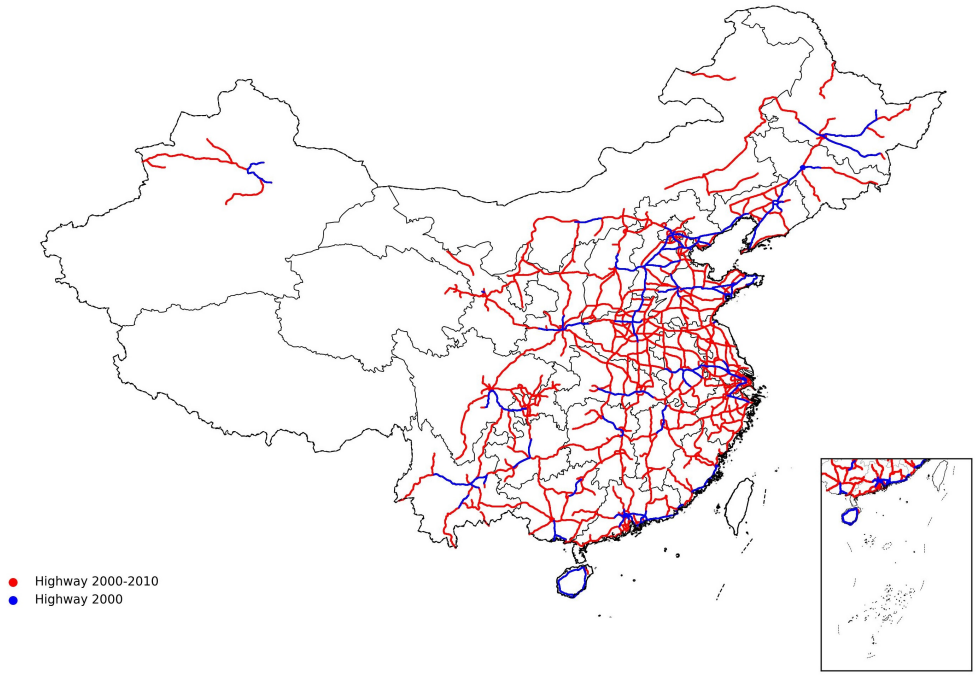


Figure 2: Highway Expansion Between 2000 and 2010

Note: Blue lines represent highways existing in early 2000; red lines represent newly constructed highways between 2000 and 2010.

trix of shortest travel times between approximately 2,800 county seats. We build a national 1 km speed raster that integrates highways, first-rate roads, and other national and provincial roads, with road-specific speeds assigned based on engineering standards and terrain.[10] Shortest travel times are computed over this unified network. Highway expansion substantially improves regional connectivity. Average inter-county travel time falls by about 17%, declining from 21.7 hours in 2000 to 18.0 hours in 2010. We use these travel-time measures both to construct market-access variables in the reduced-form analysis and to calibrate trade and migration costs in the quantitative spatial model.

2.2.3 Other Auxiliary Data

To shed light on the mechanisms underlying our main results, we supplement the core datasets with several auxiliary sources drawn from satellite imagery, administrative records, and household- and firm-level surveys.

[10] Appendix A.3 describes the construction of the speed raster, including data sources and assumed travel speeds for each road class, as well as the least-cost path algorithm.

Forest planting data We study the effect of highways on active afforestation using county-level forest planting data from 2002 to 2010 from the *China Forestry Statistical Yearbook*. The data report annual planting area in hectares, including both artificial afforestation and aerial seeding, for most counties and forestry management units nationwide. We harmonize these statistics to 2000 county boundaries and exclude special forestry units, yielding an unbalanced panel of 817 counties. Among them, 340 counties have complete annual records over the sample period and form a balanced panel. Average annual planting across all sample counties amounts to 1.87 million hectares, equivalent to 2.17% of their total forest area in 2000. The annual frequency of the data allows us to trace year-to-year variation in planting intensity and to study the dynamic response of forestry investment to improved highway access.

Land classification To track changes in forest land, we use the Global Land 30 dataset for 2000 and 2010, developed by the National Geomatics Center of China. With a spatial resolution of 30 meters and classifications derived primarily from Landsat and HJ-1 satellite imagery, the dataset distinguishes ten major land-cover types, including forest and cultivated land. The fine resolution enables us to detect relatively small-scale land-cover changes that are not observable in coarser global products such as MODIS. By comparing the two waves, we track shifts in forest boundaries at the county level and distinguish forest loss associated with highway construction from conversion to agricultural or urban uses. The use of a consistent classification algorithm across years further ensures comparability over time.

Enterprise registration data We incorporate enterprise registration records from 2000 to 2010 compiled by the State Administration for Industry and Commerce (SAIC). These data cover all formally registered enterprises in China and include registration dates, addresses, registered capital, and detailed business scope. Using keyword-based searches of business descriptions, we identify firms engaged in timber-related activities (e.g., logging, wood processing, and timber trade) as well as non-forestry agricultural enterprises. This information allows us to construct county-level measures of timber-related business activity at the beginning and end of the decade and to distinguish new entrants from continuing firms. We further separate upstream logging enterprises from downstream processors, giving a clearer view of how improved connectivity reshapes local forestry industries.

Rural fixed observation points To capture household-level mechanisms, we use the Rural Fixed Observation Point (RFOP) surveys for 2003 and 2010. This nationally representative panel, administered by the Ministry of Agriculture and Rural Affairs, covers sev-

eral hundred fixed villages across China and has been conducted continuously since the early 1980s. The RFOP data provide detailed information on household land use, agricultural and forestry production, timber sales, labor allocation, income sources, and demographics. We focus on two sets of outcomes: (i) household participation in forestry production and timber sales, and (ii) labor allocation between forestry, agriculture, and off-farm employment. Because the same villages and households are repeatedly surveyed, we can compare outcomes before and after highway expansion while controlling for time-invariant local characteristics.

Control variables Finally, we construct a comprehensive set of controls at both the county and plot levels. Socioeconomic controls include population, GDP per capita, urbanization rates, and sectoral employment shares in the year 2000. Plot-level geographic controls account for topography (slope) and proximity to administrative centers, measured by distance to county seats and urban areas. We also control for changes in key climate variables—temperature, precipitation, and sunshine—that affect forest growth. Table A.2 provides summary statistics for county level variables. Table A.3 summarizes the definitions and data sources for all control variables.

3 Empirical Analysis

We study how China's nationwide highway expansion between 2000 and 2010 affected nearby forest quality, exploiting fine-grained spatial variation in forest plots' distance to newly constructed highways. A key challenge is that highway placement may be endogenous to local economic and environmental conditions. To address this concern, we instrument realized highway access using simulated least-cost routes derived from the National Expressway Network Plan and the historical 1962 national road network.

3.1 Empirical Specification

Our empirical strategy relies on a long-difference design that compares changes in forest quality between 1999–2003 and 2009–2013. This specification is well suited to forest outcomes, which evolve slowly and reflect cumulative changes in growth, management, and labor allocation rather than short-run shocks. Using decade-scale differences also mitigates the influence of transitory weather fluctuations or temporary policy interventions that may confound short-horizon estimates.

Formally, we estimate the following baseline specification:

$$\Delta y_{icp} = \alpha + \beta\, Dist_{icp}^{HW} + \mathbf{x}_{c0}'\boldsymbol{\gamma} + \mathbf{m}_i'\boldsymbol{\eta} + \Delta\mathbf{z}_c'\boldsymbol{\delta} + \lambda_p + \varepsilon_{icp}, \tag{1}$$

where Δy_{icp} denotes the decadal change in forest quality outcomes for plot i in county c and province p.[11] The key regressor, $Dist_{icp}^{HW}$, measures the distance (in units of 10 km) from the plot to the nearest highway segment that opened between 2000 and 2010. The vector $\mathbf{x}_{c0}$ controls for baseline county characteristics measured in 2000, including log population, GDP per capita, and urbanization rates. Plot-level geographic characteristics are captured by $\mathbf{m}_i$, which includes indicators for urban proximity and steep terrain (slope above 15°), while $\Delta \mathbf{z}_c$ accounts for changes in local climatic conditions over the period. These controls are important because more developed or flatter regions tend to follow systematically different forest trajectories and were simultaneously more likely to receive highway investments. In particular, areas with higher population density and favorable topography often experience stronger land-use pressure. Province fixed effects, λ_p, absorb time-invariant institutional and policy differences across provinces. Standard errors are clustered at the county level.

To isolate the broader economic effects of improved connectivity, we exclude forest plots located within 1 km of highways. This restriction removes direct local impacts of road construction, such as physical disturbance (Forman and Alexander, 1998) or mechanical increases in vegetation driven by mandated roadside greening.[12] We additionally exclude plots located within 20 km of pre-2000 highways[13], as older transport corridors may exert long-lasting ecological effects unrelated to the expansion under study. Reintroducing these plots leaves our estimates virtually unchanged, as shown in Section 3.5.

In addition to the continuous distance measure, we estimate a specification that focuses on plots located within 1–20 km of newly constructed highways. The 20 km cutoff is motivated by binned regressions, which indicate that the largest effects of highway access occur within this range (see Section A.5 for details). The estimating equation is:

$$\Delta y_{icp} = \alpha + \beta \mathbf{I}(Dist_{icp}^{HW} < 20 \text{ km}) + \mathbf{x}_{c0}'\gamma + \mathbf{m}_i'\eta + \Delta \mathbf{z}_c'\delta + \lambda_p + \epsilon_{icp}, \tag{2}$$

where the indicator $\mathbf{I}(Dist_{icp}^{HW} < 20 \text{ km})$ equals one if a plot is located within 1–20 km of a newly constructed highway. All controls are the same as in equation 1. The binary specification facilitates the assessment of the implied total forest gain and carbon implications.

Identification We address the endogeneity in highway placement using two widely used instrumental variables from the literature. Highway placement is potentially endogenous

[11] For outcomes expressed as shares, we use percentage changes rather than logs.

[12] Although these plots account for only about 1% of the forest sample in both the NFI and GlobeLand30 datasets, excluding them ensures that the estimated effects reflect market access and reallocation mechanisms rather than mechanical deforestation from road construction land conversion.

[13] The number of plots located within 20 km of pre-2000 highways accounts for approximately 3% of the total number of observations.

because roads are sited in response to both economic and environmental conditions. New corridors may be routed toward areas with greater development potential or shaped by policy priorities that directly affect land use and forest outcomes (Asher and Novosad, 2020; Ghani, Goswami and Kerr, 2015; Faber, 2014; Baum-Snow et al., 2020; Fretz, Parchet and Robert-Nicoud, 2021; Banerjee, Duflo and Qian, 2020). For example, highway projects require approval from environmental authorities, which may steer routes away from ecologically sensitive areas and toward locations with different underlying forest growth potential, potentially attenuating ordinary least squares (OLS) estimates. At the same time, highways are more likely to be built in locations with stronger growth trends, which may themselves be correlated with forest growth trends. Because the relationship between economic development and forest outcomes may be non-monotonic (e.g., in the spirit of an environmental Kuznets curve), the direction of OLS bias is ambiguous ex ante.

Specifically, we instrument realized distance to highways ($Dist_{icp}^{HW}$) using two predetermined predictors of highway placement that are standard in the literature: distance to a simulated Least-Cost Path (LCP) network ($Dist_{icp}^{IV_{LCP}}$) and distance to the historical 1962 road network ($Dist_{icp}^{IV_{1962}}$). In the binary specification, we similarly instrument the indicator for whether an NFI plot lies within 20 km of a highway, $\mathbf{I}(Dist_{icp}^{HW} < 20\,\text{km})$, using corresponding indicators for proximity to the LCP network and the 1962 road network. These instruments are used jointly throughout the analysis, as they capture distinct but complementary sources of variation in predicted highway access. Figure A.3 shows that the binary proximity indicators for the two instruments (within 20 km) are only weakly correlated, with an R^2 of 0.15, indicating that they capture largely distinct sources of local exposure, while the continuous measures of the two instruments are correlated with an R^2 of 0.62.

The simulated LCP IV follows Faber (2014) and exploits variation from the National Expressway Network Plan (NENP). The NENP specifies a target expressway network connecting 245 major Chinese cities through 384 designated city pairs (see Section A.1). For each city pair, we compute a least-cost path that minimizes predicted construction costs as a function of terrain slope, hydrological barriers, and pre-existing built-up areas. This procedure is intended to approximate engineering feasibility rather than local economic conditions. The LCP instrument is widely used in the literature (Donaldson, 2018; Fajgelbaum and Redding, 2022; Bogart et al., 2022; Voigtländer and Voth, 2026) and builds on the straight-line instrument that connects nodal cities (Ghani, Goswami and Kerr, 2015; Asher and Novosad, 2020), while further incorporating geographic characteristics to predict where roads would be constructed based purely on construction costs, without considering economic trends.

By construction, the identifying variation in this instrument arises from how planned corridors traverse rural hinterlands between major urban endpoints, rather than from the

selection of those endpoints themselves. Because the NENP connects only major cities, the simulated paths are not designed to target intermediate towns or local administrative units whose development trajectories may affect forest outcomes. To address the concern that geographic features may affect forest quality, we include a rich set of controls—including slope, climate variables, and proximity to 333 urban centers and county seats,[14] which include the 254 target nodal cities. With these controls, a plot's distance to the simulated network is therefore primarily determined by its relative location within the national network. To further address the concern that locations lying along least-cost connections between major economic centers may differ systematically from other locations due to sorting, we follow Faber (2014) and include predetermined county-level controls measured prior to highway expansion. These include population, GDP per capita, and urbanization rates in the initial period, as well as province fixed effects.

The 1962 roads instrument follows Baum-Snow et al. (2020) and a broader literature using historical road networks as instruments for modern transport infrastructure (Baum-Snow, 2007; Michaels, 2008; Duranton and Turner, 2011; Jedwab and Moradi, 2016; Fajgelbaum and Redding, 2022). Roads constructed during this period primarily served agricultural transport and administrative or military objectives, rather than modern travel or freight demand. Historical accounts emphasize that these roads were typically low quality, locally initiated, and often unsuitable for motor traffic, with limited capacity to support interregional trade or industrial activity. As a result, the spatial distribution of the 1962 road network is unlikely to reflect the economic or environmental conditions that shape contemporary forest outcomes.

The relevance of the 1962 network arises from path dependence in infrastructure construction. Even low-quality historical roads established rights-of-way and alignments that could be upgraded to modern highways at substantially lower cost than building entirely new corridors. Consequently, locations closer to the 1962 road network were more likely, all else equal, to be incorporated into subsequent highway expansions. Our identifying assumption is therefore that, conditional on geographic characteristics and predetermined county attributes, proximity to the 1962 road network affects contemporary forest quality only through its influence on the placement of modern highways.

We assess the validity of our identifying assumptions in two ways. First, we exclude samples in the vicinity of nodal cities that the NENP was explicitly designed to connect. Although forest outcomes were not among the selection criteria, differential economic trends

[14]The 333 urban cores are identified as the contiguous built-up areas surrounding municipal government seats. These contiguous built-up areas are derived from the 1-km GLCNMO 2003 dataset Tateishi et al. (2011).

in these hubs could indirectly affect our results. While our baseline specifications control for distance to urban cores, we further restrict our analysis—for both forest outcomes and firm entry—to the rural hinterland by removing areas proximate to these nodal cities to test whether the estimates remain stable. Second, we experiment with controlling for economic and demographic trends over the study period. If the instrumental-variable networks were to affect forest outcomes through channels other than realized highway construction, the inclusion of such controls would attenuate the estimated effects. We do not include these variables in our main specifications because they are potentially endogenous outcomes of highway expansion and thus constitute "bad controls." We therefore use them only as robustness checks. Section 3.5 reports the corresponding results.

3.2 Empirical Results

Baseline results Table 1 examines the impact of highway expansion on forest biomass, measured as the change in log standing timber volume at the plot level. Columns (1)–(3) use the continuous distance to the nearest newly constructed highway (in 10-km units), while Columns (4)–(6) use a binary indicator for whether a plot lies within 20 km of a new highway. All specifications include province fixed effects and a rich set of county- and plot-level controls, with standard errors clustered at the county level to account for spatially correlated trends over time.

The results indicate a sizable and spatially concentrated increase in forest biomass following highway expansion. As shown in Column (2), moving 10 km closer to a highway increases timber volume growth by approximately 2.1%. Column (4) shows that plots located within 1–20 km of a new highway experience about 15% higher volume growth relative to more distant plots. Given that average timber volume rises by roughly 37% over the sample period, these effects are economically meaningful and concentrated in areas close to new highways.

Across both the continuous and binary specifications, IV estimates are larger in magnitude than their OLS counterparts (Columns (1) vs. (2) and Columns (4) vs. (5)). This pattern suggests that highways tend to be placed away from locations with higher underlying forest growth potential, consistent with our prior concern that OLS estimates are biased toward zero.

The first-stage results confirm the strength of the instruments. Columns (3) and (6) show that proximity to the simulated LCP network and the 1962 road network strongly predicts both realized highway distance and the probability of being within 20 km of a highway, with first-stage F-statistics exceeding 45.

Table 1: Highway Proximity and Changes in Forest Volume

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	$\Delta \ln$(Volume)		$Dist_{icp}^{HW}$	$\Delta \ln$(Volume)		$\mathbf{I}(Dist \leq 20)$
Variable	Distance (Continuous)			Connected (Binary)		
Method	OLS	IV	1st Stage	OLS	IV	1st Stage
$Dist_{icp}^{HW}$	-0.002	-0.021***				
	(0.003)	(0.008)				
$\mathbf{I}(Dist_{icp}^{HW} < 20$ km)				0.042**	0.150*	
				(0.020)	(0.086)	
LCP IV			0.261***			
			(0.037)			
1962 Road			0.076*			
			(0.041)			
LCP IV (< 20km)						0.179***
						(0.024)
1962 Road (< 20km)						0.081***
						(0.016)
Observations	16226	16226	16226	16226	16226	16226
Mean of Dep. Var.	0.37	0.37	6.28	0.37	0.37	0.20
F-statistic in 1st stage		45.5			49.5	
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
County-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports plot-level estimates for $\Delta \ln$(Volume). Columns (1)–(3) use the continuous distance to the nearest highway as the explanatory variable, while Columns (4)–(6) use a binary indicator for distance ≤ 20 km. Columns (1) and (4) report OLS estimates, (2) and (5) report IV estimates, and (3) and (6) report the first stage. Instruments are based on the LCP and 1962 Road Network. Plots within 20 km of a highway in 2000 or within 1 km in 2010 are excluded. All specifications include province fixed effects and controls. Standard errors are clustered at the county level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Other forest outcomes Appendix Table A.5 shows that the effect mostly comes from changes in coverage and DBH, with limited effects on average height. Moving 10 km closer to highways increases $\Delta \ln(\text{DBH})$ by about 0.010 (Column (3)), while the effects on $\Delta \ln(\text{Avg. Height})$ are small and statistically insignificant (Column (4)). Given standard allometric relationships in which tree volume scales approximately with $\text{DBH}^2 \times \text{Height}$, a 1% increase in DBH alone implies roughly a 2% increase in per-tree volume even with no change in height, closely matching the estimated volume effect.[15] With a 10 km change in highway access, canopy coverage rises by about 0.36 percentage points (Column (2)), indicating greater canopy closure (e.g., larger crowns, denser foliage, and fewer gaps). The coverage response is likely to be naturally growing canopy structure that accompanies the DBH-driven biomass gains.

Extensive margin Beyond the intensive margin, we examine whether highway access affects the extensive margin of forest area. Using 30-meter GlobeLand30 land-cover maps for 2000 and 2010, we compute the change in each county's forest share, excluding pixels within 1 km of the 2010 highway network to net out the mechanical footprint of road construction. We measure highway exposure as the median distance from baseline forest plots within a county to the nearest highway and instrument this distance using predicted routes from the NENP and the historical 1962 road network. Appendix Table A.12 shows that highway expansion is not associated with deforestation over the study period. If anything, proximity to highways is associated with increases in forest share. The IV estimate using the continuous distance measure is negative and statistically significant (Column (2)), implying larger forest-share gains in counties closer to highways. Similarly, when we use a binary connectivity measure (median distance < 20 km), the estimated effects are positive and statistically indistinguishable from zero. Detailed estimates are reported in Appendix A.6. These results support the claim that forest land use is highly regulated during the period, and land conversion is largely excluded from the choice set of forest farmers.[16]

Carbon sequestration benefits The gains in standing timber volume documented above imply substantial carbon sequestration benefits. Aggregating across affected areas, we estimate a gross increase of approximately 435.8 million m^3 in timber volume. This estimate

[15] In forestry, tree volume follows standard allometric relationships and is well approximated by basal area times height. Because basal area is proportional to DBH^2, volume scales approximately with $\text{DBH}^2 \times \text{Height}$. In logs, this implies $\ln(\text{Volume}) \approx 2\ln(\text{DBH}) + \ln(\text{Height})$, so holding height fixed, a 1% increase in DBH corresponds to about a 2% increase in volume.

[16] Given that the strength of our data lies in measuring tree-level outcomes on the intensive margin, and the lack of reliable biomass measures for the extensive margin, the remainder of the paper focuses on the intensive margin.

assumes that forests located within 20 km of highways—our primary effect zone—experience an average volume increase of 15% (Table 1).[17] To derive a conservative estimate, we subtract the direct loss of forest stock from highway construction, assuming a 60-meter rights-of-way width that fully removes existing vegetation. Even after accounting for these mechanical losses, we estimate a net increase of 430.2 million m^3 in forest stock. This increase corresponds to an expansion of the carbon sink by 216.9–551.9 Mt CO_2 over the decade—comparable to the annual emissions of Spain at the lower bound and Canada at the upper bound.[18] Taken together, these calculations indicate that, under strong land-use governance, infrastructure investment can generate net climate co-benefits by intensifying management on existing forest land. Appendix A.7 provides the detailed back-of-the-envelope calculation.

3.3 Mechanisms

Highway expansion improves forest quality primarily by increasing the profitability of forestry rather than by reducing land conversion or logging pressure. We identify three mechanisms through which this occurs: improved market access to timber-intensive downstream industries, the entry of specialized timber firms, and increased household specialization in forestry.[19]

3.3.1 Market Access

Improved access to timber processors lowers transportation costs and increases the marginal return to forestry. Road-induced improvements in access to timber-intensive downstream industries—particularly paper manufacturing and wood products—lead to significant increases in forest volume.[20] We arrive at this result by separating the effect of highway-induced con-

[17] We conservatively set extensive-margin contributions to zero because we lack sufficient data to quantify biomass gains from newly planted forests.

[18] These estimates depend on Biomass Conversion and Expansion Factors (BCEFs), which convert timber volume into total ecosystem biomass and vary by tree species and stand characteristics. We use a range of 0.55 to 1.4, reflecting typical values for forests in China, to derive the reported lower and upper bounds.

[19] Alternative channels related to labor migration and reduced agricultural pressure are unlikely to drive these results(see Appendix A.8). First, we find no causal effect of aggregate labor supply shocks on forest volume , suggesting that rural depopulation does not passively lead to forest recovery. Second, there is no significant reduction in cultivated land near highways, providing no support for the farmland abandonment hypothesis.

[20] In the 2007 input–output table, 98 percent of forestry output is absorbed as intermediate inputs, so demand is overwhelmingly downstream-industry driven. The largest users are wood products (C20), paper manufacturing (C22), and furniture production (C211); the table does not separately report wooden furniture, so the furniture share is an upper bound. These downstream industries value different timber attributes and therefore reward quality in different ways. Pulp and paper producers are sensitive to fiber characteristics that affect pulp yield and processing efficiency; many wood products (including sawnwood

nectivity improvements from demand shocks due to broader industrial expansion. Specifically, we decompose changes in market access into a road-improvement component and an income-growth component. The road improvement effect (RE) captures changes in market access driven solely by reductions in travel time, holding downstream output fixed:

$$RE_{kc} = \ln\left(\sum_{j\neq c} Y_{kjt-1} \cdot time_{cjt}^{-\theta}\right) - \ln\left(\sum_{j\neq c} Y_{kjt-1} \cdot time_{cjt-1}^{-\theta}\right), \tag{3}$$

where Y_{kjt-1} denotes baseline output of industry k in county j, $time_{cjt}$ is the travel time between counties c and j, and θ is the trade elasticity parameter (Ma and Tang, 2024). By construction, this measure isolates changes in market access attributable to highway construction and speed upgrades.

In parallel, the income effect (IE) holds travel times fixed while allowing downstream output to change:

$$IE_{kc} = \ln\left(\sum_{j\neq c} Y_{kjt} \cdot time_{cjt}^{-\theta}\right) - \ln\left(\sum_{j\neq c} Y_{kjt-1} \cdot time_{cjt}^{-\theta}\right). \tag{4}$$

This component absorbs variation arising from geographic advantages or historical development patterns that may simultaneously influence downstream output growth and forest outcomes.[21]

We estimate forest responses separately for timber-intensive downstream industries and for non-timber manufacturing sectors. Specifically, we construct RE and IE for timber-related industries and for the remainder of manufacturing. Controlling for access to non-timber industries is crucial for isolating the forestry-specific market-access channel. The same highway segments that improve access to timber processors also improve access to other manufacturing sectors, which can increase outside employment opportunities and draw labor away from forestry management. We estimate a county-level specification that replaces plot-level exposure in equation 1 with county-level market-access measures:

$$\begin{aligned}\Delta y_{icp} = \alpha_i &+ \beta_{1k} RE_c^{\text{Specific}} + \beta_{2k} RE_c^{\text{Other}} + \beta_{3k} IE_c^{\text{Specific}} + \beta_{4k} IE_c^{\text{Other}} \\ &+ \mathbf{x}_{c0}'\boldsymbol{\gamma} + \mathbf{m}_i'\boldsymbol{\eta} + \Delta\mathbf{z}_c'\boldsymbol{\delta} + \lambda_p + \varepsilon_{icp}.\end{aligned} \tag{5}$$

and engineered wood) require strength, density, and straightness; and furniture production typically pays a premium for appearance-grade lumber (uniform grain and fewer defects). Improved access to these buyers thus steepens the local price gradient with respect to quality, strengthening incentives for forest-quality upgrading. Wood products, paper manufacturing, and furniture account for 41.20%, 6.09%, and 7.78% of total intermediate forestry-product use, respectively.

[21] We use 2007 output data as the baseline because the 2008–2010 firm-level data are widely considered less reliable due to severe sample truncation and measurement errors in key inputs (Brandt, Van Biesebroeck and Zhang, 2014).

Road-induced improvements in access to timber-intensive downstream industries significantly increase forest volume, while improved access to other manufacturing sectors reduces it, as shown by the first two rows in Column (1) of Table 2. In contrast, the income effect terms for timber-intensive industries are small and statistically insignificant, indicating that the observed forest responses are driven primarily by reductions in travel time rather than by contemporaneous growth in downstream output. The large and positive coefficient on the income effect for other manufacturing reflects a demand shock due to income growth resulting from broader industrial expansion that spills over to demand for forestry products. Disaggregating by downstream industry, the positive effects are concentrated in paper manufacturing and wood products, while the coefficient for wood furniture is small and statistically insignificant.[22]

Table 2: Disentangling Road Expansion Effects on Forest Performance

Dep. Var.	(1)	(2)	(3)	(4)
		Specific Industry		
Δ ln(Volume)	Downstream Industries	Paper Manufacturing	Wood Furniture	Wood Products
RE_c^{Specific}	0.691**	0.551**	0.073	0.476*
	(0.279)	(0.238)	(0.380)	(0.246)
RE_c^{Other}	-0.802**	-0.634*	-0.034	-0.495
	(0.371)	(0.337)	(0.366)	(0.308)
IE_c^{Specific}	-0.090	0.122	0.217*	-0.137*
	(0.137)	(0.138)	(0.119)	(0.078)
IE_c^{Other}	0.946***	0.739***	0.531*	0.919***
	(0.245)	(0.234)	(0.271)	(0.233)
Observations	16226	16226	16226	16226
Province fixed effects	Yes	Yes	Yes	Yes
Mean of IE_c^{Specific}	1.491	1.333	1.745	1.709
Mean of RE_c^{Specific}	0.249	0.253	0.241	0.243
County-level controls	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes

Note: Estimates correspond to the specification in Equation (5), with the trade elasticity parameter $\theta = 1.5$ following Herzog (2021). Column (1) pools the three timber-intensive downstream industries. Columns (2)–(4) report results separately for paper manufacturing, wood furniture, and wood products. RE_c^{Specific} measures road-induced changes in market access to the specified downstream industry, while RE_c^{Other} captures changes in access to all other manufacturing industries. IE_c^{Specific} and IE_c^{Other} measure changes due to industry output growth (income effect). All specifications include province fixed effects, county-level controls, and plot-level controls. Robust standard errors are clustered at the county level and reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

[22]This pattern reflects the structure of China's forestry sector: domestic forests are dominated by fast-growing, lower-quality species such as eucalyptus and poplar, which are mainly used for pulp and low-value wood products, whereas higher-quality timber used in furniture production is largely imported.

3.3.2 Highways Attract Specialized Forestry Firms

Highway connectivity induces substantial entry of specialized timber-cultivation firms, which increases access to urban inputs. Our IV estimates indicate a sharp increase in timber-cultivation enterprises following highway connection, pointing to the arrival of larger-scale and more technically capable operators that complement household forestry. To quantify firm entry, we assemble county-level enterprise registration records from 2000–2010 for counties containing National Forest Inventory (NFI) plots.[23] We estimate the following specification:

$$\Delta y_{cp} = \alpha + \beta \cdot \mathbf{I}(Dist_{cp}^{HW} < 20 \text{ km}) + \mathbf{x}_{c0}'\boldsymbol{\gamma} + \lambda_p + \varepsilon_{cp}, \tag{6}$$

where Δy_{cp} denotes the change in the number of registered timber-cultivation firms in county c between 2000 and 2010. The treatment indicator equals one if the county lies within 20 km of a highway completed by 2010.[24] We control for baseline county characteristics in 2000, including log county area, forest-land share, log population, and log GDP per capita, as well as province fixed effects.[25]

Table 3 reports OLS and IV estimates of the effect of highway proximity on firm entry. Counties connected to newly built highways experience a significant increase in registered timber-cultivation firms. The IV estimates imply a 58.6 percent increase in the number of timber firms (Column 2), substantially larger than the 9.7 percent increase obtained under OLS (Column 1). This divergence suggests that OLS estimates are attenuated by endogenous highway placement. Areas prioritized for highway construction tend to experience rapid growth in industrial and service activities, which can compete with forestry for land, labor, and capital. Such structural transformation may discourage forestry investment in otherwise well-connected locations, biasing OLS estimates downward.

As a placebo test, Columns (3)–(4) examine non-timber, crop-based agricultural firms. We find no statistically significant effects under either specification. This null result is consistent with the predominantly mountainous nature of our sample counties, which are poorly suited for crop agriculture but well suited for forestry. Together, these findings indicate that highway connectivity selectively attracts specialized forestry firms rather than generating a

[23] Firm addresses are typically registered in county seats and do not reliably reflect operating locations in forested areas; aggregating to the county level mitigates this misclassification. Timber-cultivation enterprises are identified by parsing business-scope descriptions. Firms are classified as timber-related if their scope includes specific tree species (e.g., *Phoebe zhennan*, *Eucalyptus*, *Cunninghamia lanceolata*, or fast-growing forests) and forestry activities such as planting, harvesting, cultivating, tending, or afforestation. Enterprises explicitly referencing forestry, silviculture, or the timber industry are also directly coded as timber firms.

[24] Distance is measured as the median distance from all forest plots within a county to the nearest highway.

[25] We exclude 86 counties that were already within 20 km of highways in 2000 and one county with no timber-cultivation firms, leaving 817 counties in the estimation sample.

broad expansion of rural enterprises, reinforcing the interpretation that improved market access reallocates productive activity toward forestry-specific investments.

We further probe the nature of these new entrants to assess whether improved connectivity also strengthens access to upstream, urban-supplied inputs and services used in forestry. Using firms' registered business scopes, we construct overlapping, input-oriented categories by searching for keyword strings linked to key production factors. In particular, machinery identifies firms referencing forestry machinery, logging and cultivation equipment, and related mechanical services, while seedlings captures firms engaged in nursery stocks, tree seedlings, and seed provision for timber plantations.[26] Given the prevalence of zero entrants in these narrow categories, we employ IV-PPML to accommodate the zero-valued observations while maintaining the standard percentage-change interpretation of the coefficients. Columns (5)–(6) of Table 3 show that highway connectivity significantly increases the entry of timber firms whose scopes involve machinery and seedling inputs. Together, these patterns support an upstream-input mechanism: by lowering transport and search costs, highways facilitate the local availability of urban intermediates and specialized services required for intensive forest management.

Table 3: Highway Connection and County-Level Entry of Specialized Forestry Firms

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	log(Num. Firms)				Num. Firms (Poisson)	
Industry/Input	Timber		Non-Timber Agric.		Machinery	Seedlings
Method	OLS	IV	OLS	IV	IV-PPML	IV-PPML
$\mathbf{I}(Dist_{cp}^{HW} < 20\text{km})$	0.097*	0.586***	-0.007	0.074	0.789**	0.736*
	(0.052)	(0.196)	(0.038)	(0.132)	(0.311)	(0.379)
Observations	817	817	817	817	817	817
F-statistic in 1st stage		27.76		27.76	27.76	27.76
Mean of Num. Firm Entry	70.25	70.25	632.97	632.97	8.81	38.06
Province Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Control	Yes	Yes	Yes	Yes	Yes	Yes

Note: Estimates are based on the specification in Equation (6). The highway connection indicator is instrumented using both the simulated least-cost path network and the 1962 national road network. Columns (1)–(4) report results for log(number of firms), where columns (1)–(2) focus on specialized forestry (timber) firms and (3)–(4) focus on non-timber agricultural firms. Columns (5)–(6) report IV-PPML estimates for the count of timber-related firms categorized by specific keywords in their registered business scopes. All specifications include province fixed effects and baseline county characteristics in Equation 1. Robust standard errors are reported in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

[26]To exclude agricultural activities unrelated to forestry, we restrict the scope of these input-oriented categories to the timber-cultivation firms analyzed in columns (1) and (2) of Table 3.

3.3.3 Household Specialization in Forestry Work

Highway expansion reshapes rural labor allocation by inducing both migration and household specialization. Households with prior forestry experience nearly double their timber sales following highway connection, while households without such experience respond primarily by increasing migration for off-farm work.

We use household- and individual-level data from the Rural Fixed Observation Point (RFOP) and implement a two-period difference-in-differences design comparing outcomes in 2003 and 2010.[27] The treatment group consists of villages located within 20 km of highways completed by 2010.[28] To be consistent with the NFI data-based sample, we further restrict the sample to villages located in counties containing NFI survey plots in 2003. We also exclude villages with no forestry activity in 2003 to ensure comparability across both villages and households while focusing on areas with forestry potential. We aggregate individual records to the household level, defining household migration duration as the average days worked outside the township among all eligible labor force members. The sample is restricted to households with at least one active labor force member.[29] The final sample includes 1,291 households across 26 villages, approximately half of which had members engaged in forestry work in 2003, which we define as households with forestry experience.[30] We estimate the following specification:

$$y_{vht} = \alpha + \beta \mathbf{I}(Dist^{HW}_{v,2010} < 20 \text{ km}) \times Post_t + \mathbf{x}'_{c0}\boldsymbol{\gamma} + \lambda_h + \varepsilon_{vht}, \tag{7}$$

where v, h, and t index villages, households, and years. The dependent variable y_{vht} captures migration duration, timber sales, or forestry income. The treatment indicator $\mathbf{I}(Dist^{HW}_{v,2010} < 20 \text{ km})$ equals one if village v lies within 20 km of a highway in 2010. The control vector $\mathbf{x}_{c0}$ includes county-level log GDP per capita and log population in 2000, as well as village-level indicators for average slope greater than 15° and distance to county seats.

Table 4 presents the results. Households with prior forestry experience increase timber sales revenue by 1,113 RMB—approximately 100% of the 2003 mean—and timber quantity

[27] Data from 2000 are excluded because key variables, including migration duration and timber sales, are unavailable.

[28] To isolate the effect of new connections, we exclude villages already within this distance in 2003.

[29] The RFOP survey shifted its definition of migrant workers from "outside the village" (pre-2008) to "outside the township" (post-2008). To ensure longitudinal consistency, we utilize granular individual records on migration destinations and durations to unify the measure as "working outside the township" across all years. We exclude households with missing or ambiguous migration records. The eligible labor force is restricted to rural *hukou* holders, specifically males aged 16–65 and females aged 16–60. Household-level migration days are then calculated as the average of individual migration days for all such eligible members.

[30] Forestry experience is defined as participation in family-based forestry management and production, excluding employment in state-owned forest farms or large private forestry enterprises.

by 0.53 m^3, nearly 200% of the baseline mean though insignificant, following highway connection. In contrast, households without forestry experience respond by increasing migration duration by 35.4 days, while experienced households show no significant change in migration. These patterns suggest that prior forestry experience confers a comparative advantage in exploiting improved market access: experienced households deepen specialization in timber production, while inexperienced households shift labor toward non-farm work. This result is consistent with Ma et al. (2025), who show that improved connectivity promotes agricultural specialization and selective rural out-migration. Highways therefore operate as a selection device, magnifying pre-existing differences across households and reinforcing specialization along lines of comparative advantage.

Taken together with the firm-level evidence, these results shed light on the mechanisms of market access and labor specialization. Highway connectivity deepens the engagement of experienced households while simultaneously attracting specialized forestry-cultivation firms. By supplying services and intermediate inputs, these firms lower production costs. Households with forestry experience are better positioned to benefit from this complementary infrastructure, leading to higher productivity and more intensive forest management in connected areas.

Table 4: Selection Effects of Highway Access on Households

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)
	Migrant Days		Forestry Revenue(yuan)		Wood Sale(m^3)	
Forest Household 2003	Yes	No	Yes	No	Yes	No
$\mathbf{I}(Dist^{HW}_{v,2010} < 20 \text{ km}) \times Post_t$	-7.33	35.43**	1112.69**	96.37	0.53	0.01
	(12.98)	(14.51)	(492.74)	(126.33)	(0.33)	(0.04)
Household Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Control	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1474	1108	1474	1108	1474	1108
Mean of Num. Firms 2003	52.13	65.96	810.95	0.00	0.25	0.00
Share of HHs treated	0.66	0.31	0.66	0.31	0.66	0.31

Note: The regression is based on the specification in Equation (7). We restrict the sample to villages located in counties with at least one forest plot to ensure consistency. We further limit the sample to villages that had at least one household engaged in forestry management in 2003 (i.e., ≥ 1 household reporting positive days in forestry management). Controls include county-level ln(GDP per capita), ln(County Population), village-level slope $> 15°$, and distance to county seat, all interacted with year. Standard errors in parentheses. Robust standard errors are clustered at the village level. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

3.4 Heterogeneity

Property rights Highway effects are markedly larger in forests that were collectively owned prior to the reform, as shown in Column (1) of Table 5. For non-collective forests, a

10 km reduction in distance to the nearest highway increases forest volume by about 1.7 percent. In contrast, for forest plots that were initially collectively owned and subsequently subject to tenure reform, the corresponding effect rises to approximately 3.9 percent—more than twice the baseline response.

This pattern is consistent with the view that improvements in market access translate into real changes in resource management only when operators hold secure, long-term, and residual claims over returns. As described in the institutional background, during our study period China's collective forest tenure reform reassigned forestland management rights and timber ownership from village collectives to individual households, while legally permitting households to transfer or lease these rights to enterprises. This reform substantially strengthened incentives for long-term investment and active forest management.

In this institutional environment, improved connectivity raises expected returns to forestry by lowering transport costs and expanding access to downstream timber markets, while secure and transferable tenure ensures that these gains accrue to forest managers rather than being dissipated at the collective level (Besley, 1995; Banerjee and Iyer, 2005; Acemoglu, Johnson and Robinson, 2005; Goldstein and Udry, 2008; Hornbeck, 2010). By contrast, where forests remain state-managed and managerial incentives are weak, similar improvements in market access generate much smaller responses, consistent with prior evidence that secure property rights are central to sustainable forest management (Walker et al., 2025; Xie, Berck and Xu, 2016; Santika et al., 2021). In line with this interpretation, the interaction term for privately owned forests is close to zero and statistically insignificant, reflecting the fact that these forests were not directly affected by the collective tenure reform. Finally, we note that our land-tenure classification is observed only in the initial period, which prevents us from directly tracking subsequent tenure transitions at the plot level.

Space and time The heterogeneous spatial patterns help explain the average effects documented in the baseline specifications. As shown in Column (2) in Table 5, the impact of highway proximity on forest volume varies with terrain. On steeper slopes, defined as being greater than 15°, moving closer to a highway is associated with a forest volume gain of approximately 4.7 percent, more than twice the baseline average effect of 1.9 percent.[31] This pattern is consistent with differences in land-use opportunity costs: forestry investment responds more strongly in locations where agricultural alternatives are limited, as opposed to flatter areas where agricultural opportunity costs are higher. Time since connection further accounts for the magnitude of the average treatment effect. Columns (3)–(4) show that

[31]The similarity between the plain-terrain estimate and the average effect is likely to reflect the fact that plots on flatter land constitute a large share of the sample and therefore receive substantial weight in the estimator.

forest volume gains accrue primarily in plots exposed to highways for longer periods.[32] Plots connected between 2000 and 2005 experience statistically significant increases in forest volume (Column (3)), while plots connected after 2005 display negligible effects over the sample window. Adding back plots connected before the sample period shows that the effect is even larger for those connected before 2000 (Column (4)).

Table 5: Heterogeneity: Highway Impact on Forest Volume

Dep. var	(1)	(2)	(3)	(4)
	$\Delta \ln$(Volume)			
	Interactions (Double IV)		Duration Effects	
Sample	Post-1999	Post-1999	Post-1999	All
$Dist_{icp}^{HW}$	-0.017**	-0.019***		
	(0.007)	(0.007)		
Collective $\times$ $Dist_{icp}^{HW}$	-0.022***			
	(0.008)			
Collective	0.211***			
	(0.041)			
Slope$\geq$ 15 $\times$ $Dist_{icp}^{HW}$		-0.028***		
		(0.007)		
Slope$\geq$ 15		0.038		
		(0.038)		
$\mathbf{I}(Dist_{icp}^{HW,bf99} < 20\text{km})$				0.118**
				(0.048)
$\mathbf{I}(Dist_{icp}^{HW,9905} < 20\text{km})$			0.064**	0.061**
			(0.029)	(0.029)
$\mathbf{I}(Dist_{icp}^{HW,0510} < 20\text{km})$			0.030	0.027
			(0.026)	(0.026)
Observations	16138	16226	16226	16726
Mean of Dep. Var.	0.367	0.368	0.368	0.373
F-statistic in 1st stage	19.6	11.7		
Pre-1999 Share				0.030
1999-2005 Share			0.085	0.082
2005-2010 Share			0.119	0.115
Province FE	Yes	Yes	Yes	Yes
County Controls	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes

Notes: This table presents heterogeneity analysis and period effects based on the specifications in Equation (1) and Equation (2). Columns (1)–(2) use the LCP and the 1962 road network jointly as instruments. Columns 3–4 report OLS estimates because first-stage F-statistics are insufficient when disaggregating treatment by connection periods. Column 1 reports estimates with tenure interaction, where the dummy indicates Collective forests. Column 2 reports estimates with geography interaction (steep slopes $\geq 15°$). Column 3 estimates separate effects for connections during 2000–2005 and 2005–2010 periods (excluding pre-2000). Column 4 includes all periods with separate indicators. Controls are as specified in Table 1. Robust standard errors are clustered at the county level and reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

[32] We report OLS estimates for these specifications because the first-stage F-statistics are insufficient when disaggregating the treatment by connection periods.

3.5 Robustness Checks

Removing nodal locations A potential identification concern is that the highway network was explicitly designed to connect nodal cities, which often possess historical geographic and economic advantages. If forest outcomes near these cities partly reflect urban agglomeration forces rather than highway access itself, the estimated effects could be confounded. To address this concern, we restrict the sample to forest plots located more than 20 km and 46 km away from nodal city centers.[33] As shown in Table A.6 (Columns 3–6), the resulting estimates are similar in magnitude to the baseline results, suggesting that the main findings are unlikely driven by proximity to nodal cities. The firm entry results are also robust to removing nodal locations. We replicate the county-level firm entry analysis after excluding counties in which the median forest plot lies within 20 km or 46 km of a nodal city center. As reported in Table A.7, the IV estimates for timber-cultivation firm entry remain economically meaningful and statistically significant even under the more restrictive 46 km exclusion. This pattern is consistent with highways influencing firm entry in rural areas beyond the immediate influence of nodal cities.

Controlling for economic and population growth A potential concern is that highway expansion may be correlated with broader regional economic and demographic trends that could independently affect forest outcomes, for example through changes in local demand, governance capacity, or environmental enforcement. Our identification strategy is already designed to address this concern by isolating plausibly exogenous variation in highway placement. As an additional robustness check, we further augment the baseline specification with controls for local economic and population growth. Because these variables are potentially affected by highway expansion and therefore constitute "bad controls", they are included only for robustness purposes and are not part of the baseline specification. Specifically, we include county-level population growth between 2000 and 2010, as well as growth in nighttime light intensity, a widely used proxy for local economic activity. Nighttime light growth is measured using inter-calibrated DMSP-OLS stable lights data from 2000 and 2010, following Hsu et al. (2015).

As reported in Table A.8, the estimated effect of highway proximity on forest volume changes little in magnitude and remains statistically significant after controlling for both population and economic growth. This stability suggests that our results are not driven by correlated long-run development trends, but instead reflect a distinct impact of improved transport connectivity on forest management and quality.

[33] The 20 km and 46 km thresholds correspond to the 20th and 50th percentiles of the distance from forest plots to highways. City centers are identified based on the locations of government seats.

Varying trade elasticity parameters We assess the robustness of the market access mechanism underlying our main results by varying the trade elasticity parameter θ in equation 5. As shown in Table A.9, the qualitative pattern of the estimates remains stable across a range of values commonly used in the literature, including $\theta = 1$, $\theta = 1.5$ (following Herzog (2021)), and $\theta = 3.6$ (following Donaldson and Hornbeck (2016)).

Across all specifications, improved market access to timber-intensive downstream industries is associated with statistically significant increases in forest product volumes, while market access to other manufacturing sectors exhibits weaker or opposite effects. Although the magnitudes of the coefficients scale mechanically with θ, the relative importance of timber-specific demand and the sign of the estimated effects remain unchanged. This stability indicates that our results are not sensitive to the particular calibration of trade elasticity.

Alternative sample definitions Our results are robust to alternative sample selection criteria. In Table A.10, Column (1) reports the baseline, which excludes plots within 20 km of pre-2000 highways. Column (2) reintroduces plots within 20 km of pre-2000 highways, while Column (3) applies a stricter 50 km exclusion around pre-2000 highways. Estimates are stable across these specifications, indicating that the results are not driven by lasting impact of earlier infrastructure. Column (4) includes roadside plots within 1 km, which are excluded in the baseline to avoid mechanical effects. Finally, Column (5) includes plots near pre-2000 highways and defines proximity using the full 2010 network. The similarity of these estimates to the baseline supports that our findings are robust to alternative definitions of highway exposure.

4 Quantitative Framework

The empirical results are most consistent with a market-access mechanism. Highways raise standing timber volume and shift households toward more intensive forest management, with effects concentrated near new roads and strongest where tenure security and geography make scale-up feasible. The data also suggest a smaller composition response: as access improves and outside options strengthen, some low-productivity households exit forestry, concentrating production among higher-productivity managers. We then quantify these margins in a spatial trade model with mobility, so the reduced-form effects can be translated into counterfactuals that speak to policy, including where roads deliver the largest forest gains and where complementary reforms are needed to amplify them.

4.1 Model Environment

We consider an economy with N regions. Each region $i \in \{1, \ldots, N\}$ is divided into an urban area ($a = u$) and a rural area ($a = r$), so a *geographic unit* is given by the region–area pair (i, a), yielding a total of $2N$ units. There are two sectors: forestry (f), located in rural areas, and urban production (u), located in urban areas. Forestry production requires both land and intermediate inputs sourced from urban goods. Urban goods are produced with labor as the sole input, consumed in all locations, and also used as inputs in forestry.[34] Households are heterogeneous in their comparative advantage in forest production. Specifically, there are two types indexed by $z = 1, 2$, with productivity levels normalized to 1 and $\psi > 1$, respectively. This heterogeneity captures the fact that some households are inherently more suited to forest production than others. Households choose their residence and sectoral engagement based on their productivity type, location of birth, and prevailing economic conditions, subject to migration and trade costs across regions.

Consumption Household utility in area a of region n is Cobb–Douglas in final goods from the forest sector f and the urban sector u:

$$u_{n,a} = \phi_{n,a}(Q^f_{n,a})^{\eta}(Q^u_{n,a})^{1-\eta}, \tag{8}$$

where η and $1-\eta$ are the expenditure shares on forest and urban goods. The amenity term $\phi_{n,a} = \bar{\phi}_{n,a}L^{\beta}_{n,a}$ combines a location-specific component $\bar{\phi}_{n,a}$ with a congestion effect, where $\beta < 0$ captures the disutility from crowding. Forest goods produced in rural areas, Q^f_n, and urban goods produced in urban areas, Q^u_n, are both aggregated across differentiated varieties using a CES function with elasticity of substitution ϵ:

$$Q^s_n = \left[\sum_i \int_{\omega \in \Omega^s_{ni}} \left(q_{ni}(\omega)\right)^{\frac{\epsilon-1}{\epsilon}} d\omega\right]^{\frac{\epsilon}{\epsilon-1}}, \qquad s \in \{f, u\}, \tag{9}$$

where $q_{ni}(\omega)$ is the quantity of variety ω sourced from region i. The CES structure ensures that varieties from multiple regions substitute for one another, so lower trade costs expand the range of accessible goods and reduce their effective prices.[35]

[34] We abstract from agricultural production since land-use data show little systematic change in cultivated area before and after highway construction. The main margin of adjustment occurs within forestry itself and through migration out of rural areas, rather than through shifts between agriculture and forestry.

[35] Rural and urban subareas within region n operate in a common market for forest goods produced in rural areas and urban goods produced in urban areas; accordingly, they share the same sectoral price indices, P^f_n and P^u_n.

Forest production The forest sector is modeled under monopolistic competition. Each rural household produces a distinct variety of forest goods, ω^f, using a technology that exhibits increasing returns to scale. Producers compete against each other while taking the prices of other varieties as given. Each household operates its own farm, combining land and urban intermediate inputs to produce output. Labor is not modeled separately, as the household acts simultaneously as owner and worker; farm income is therefore fully captured by profits. Production requires a fixed cost $\bar{f}$, measured in units of the numeraire, which reflects expenditures such as land use, machinery, and overhead. The presence of fixed costs ensures increasing returns, making variety-level production consistent with monopolistic competition. Because each household produces a single variety, producers can be indexed either by type z or by variety ω^f. The production function for a household of type z in region i is given by

$$q_i^f(\omega^f) = \left(\bar{A}_i \psi_z\right) \left(\frac{\mathbb{m}_i}{\upsilon}\right)^{\upsilon} \left(\frac{\mathbb{Q}_i^u}{1-\upsilon}\right)^{1-\upsilon}, \tag{10}$$

where υ and $1-\upsilon$ are the cost shares of land and urban inputs, respectively. Total factor productivity has two components: a location-specific term $\bar{A}_i$, and a household-specific term ψ_z that captures heterogeneity in forestry skills. The variable $\mathbb{m}_i$ denotes land use by the household, while $\mathbb{Q}_i^u$ is the bundle of urban intermediates employed in production. Each rural region i is endowed with a fixed stock $\bar{N}_i$ of land, owned by immobile landlords. Landlords consume their income and share the same utility function as households, as specified in equation 8. Given the CES demand structure, the forest-goods price index in region n is: $P_n^f = \left[\sum_i \left(\frac{\epsilon}{\epsilon-1}\frac{\chi_i}{\bar{A}_i\bar{\psi}_i}\tau_{ni}^f\right)^{1-\epsilon} L_{i,r}\right]^{\frac{1}{1-\epsilon}}$.

Notably, the model does not incorporate endogenous quality investment, since our data lack precise measures of activities such as replanting, thinning, or the use of higher-quality inputs. Instead, heterogeneity in forestry performance is captured by ψ_z, which reflects differences in farmers' inherent skills or comparative advantage. We assume that effective forest productivity depends linearly on ψ_z, consistent with evidence that more skilled households remain active and intensify management when market access improves. A more general mapping between farmer productivity and effective forest productivity would require accurate measures of farmer-level TFP, which are rarely available. Treating forestry skills as an exogenous driver of productivity therefore provides a tractable and empirically grounded way to capture how household heterogeneity shapes forest outcomes. A central object in the model is region-i effective forestry productivity, which is pinned down by the composition of households that engage in production. Formally, effective forest productivity in region i is:

$$\bar{\psi}_i = \left(\frac{L_{1i,r} + \psi^{\epsilon-1} L_{2i,r}}{L_{i,r}}\right)^{\frac{1}{\epsilon-1}}, \tag{11}$$

where $L_{1i,r}$ and $L_{2i,r}$ denote the numbers of low- and high-productivity farmers in the rural area of i, and $L_{i,r}$ denotes their sum. This expression shows that regional effective forest productivity rises not only with the number of active farmers but, more importantly, with the share of higher-productivity households. Market access influences $\bar{\psi}_i$ directly: if less productive farmers exit while more skilled ones expand, average productivity improves. Road expansion therefore matters not only for aggregate output but also for the productivity of forestry, by altering the composition of producers who remain in the sector. Detailed derivations of the forest production framework are provided in Appendix B.1.

Urban production Each urban variety (ω^u) in region i is produced using labor as the sole input:

$$q_i^u(\omega^u) = \xi_i(\omega^u) B_i \ell, \tag{12}$$

where $B_i = \bar{B}_i L_{i,u}^{\alpha}$ captures location-specific productivity, determined by a fundamental term $\bar{B}_i$ and agglomeration effects, with $\alpha > 0$ measuring the elasticity of productivity with respect to the urban workforce. The urban sector operates under perfect competition. Productivity shocks $\xi_i(\omega^u)$ are i.i.d. across varieties and locations, drawn from a unit Fréchet distribution with shape parameter θ. Following Eaton and Kortum (2002), consumers in each location purchase from the lowest-cost supplier, yielding the sourcing probability:

$$\pi_{ni}^u = \frac{\left(w_i \tau_{ni}^u\right)^{-\theta} \left(B_i\right)^{\theta}}{\sum_{i'} \left(w_{i'} \tau_{ni'}^u\right)^{-\theta} \left(B_{i'}\right)^{\theta}}. \tag{13}$$

The corresponding price index for urban goods in region n is $P_n^u = \gamma \left[\sum_i \left(w_i \tau_{ni}^u\right)^{-\theta} \left(B_i\right)^{\theta}\right]^{-\frac{1}{\theta}}$ where $\gamma = \left[\Gamma\left(\frac{\theta+1-\epsilon}{\theta}\right)\right]^{1/(1-\epsilon)}$, and $\Gamma(\cdot)$ is the Gamma function. Highways matter here because lower trade costs τ_{ni}^u increase the competitiveness of nearby urban centers, raising demand for urban labor and pulling less-productive rural households into cities. This migration channel reshapes who remains in forestry: as weaker farmers exit, those with stronger comparative advantage in forest production stay, intensify management, and drive up effective forest productivity. Thus, the urban sector not only supplies intermediates to forestry but also acts as the destination sector that underpins the selective mechanism of road expansion.

4.2 Sorting of Individuals by Comparative Advantage.

Migration Households choose their location by weighing real income, amenities, and migration costs. The indirect utility of a type-z household ω, born in (i, a_0) and choosing to live in (n, a), is $v_{zin,a_0a}(\omega) = \frac{\phi_{n,a} y_{zn,a}}{P_n \mu_{in,a_0a}} \epsilon_{zin,a_0a}(\omega)$, where P_n is the CES price index of forest

and urban goods, $P_n = (P_n^f)^{\eta}(P_n^u)^{1-\eta}$, and $y_{zn,a}$ denotes household income in location (n, a):

$$y_{zn,a} = \begin{cases} \epsilon^{-1} R_n^f (L_{1n,r} + \psi^{\epsilon-1} L_{2n,r})^{-1} & \text{if } z = 1 \text{ and } a = r \\ \epsilon^{-1} R_n^f (L_{1n,r} + \psi^{\epsilon-1} L_{2n,r})^{-1} \psi^{\epsilon-1} & \text{if } z = 2 \text{ and } a = r \\ w_n & \text{if } a = u \end{cases} \tag{14}$$

Here, R_n^f denotes forest revenues in region n, and w_n the urban wage. Following Ahlfeldt et al. (2015), idiosyncratic tastes ϵ_{zin,a_0a} are i.i.d. across households and drawn from a Fréchet distribution with shape parameter κ. This structure yields the migration probability that a type-z household born in (i, a_0) chooses (n, a):

$$\lambda_{zin,a_0a} = \frac{\left(\frac{\phi_{n,a} y_{zn,a}}{P_n \mu_{in,a_0a}}\right)^{\kappa}}{\sum_{n'} \sum_{a'} \left(\frac{\phi_{n',a'} y_{zn',a'}}{P_{n'} \mu_{in',a_0a'}}\right)^{\kappa}}. \tag{15}$$

Sorting of individuals by comparative advantage The migration probability in equation 15 embodies how households sort across space according to their comparative advantage. To build intuition, consider the relative odds that a type-z individual currently in rural region i chooses to move to the urban area of j rather than stay in i,

$$\tilde{\lambda}_{zij,ru} = \frac{\lambda_{zij,ru}}{\lambda_{zii,rr}} = \left(\frac{\phi_{j,u} y_{zj,u} / P_j}{\phi_{i,r} y_{zi,r} / P_i}\right)^{\kappa} \mu_{ij,ru}^{-\kappa}. \tag{16}$$

The ratio captures the utility gain from migration relative to staying, which depends directly on the household's productivity across the two sectors. Since rural incomes vary with forestry productivity, the migration decision reflects where a household's skills are more highly rewarded. Dividing the relative out-migration probability of type 1 by that of type 2 and substituting income terms yields

$$\tilde{\lambda}_{1ij,ru} / \tilde{\lambda}_{2ij,ru} = \psi^{(\epsilon-1)\kappa}, \tag{17}$$

where $\psi > 1$ captures the productivity edge of type-2 households in forestry. Lower-productivity farmers (type 1) are more likely to leave rural forestry for urban work, whereas higher-productivity farmers (type 2) are more likely to stay and scale up their forestry activities. For ease of reference, we label type 1 as Urban CA (comparative advantage in urban work) and type 2 as Forest CA. Migration costs $\mu_{in,ru}$ govern the strength of this sorting. When frictions are high, even Urban-CA households may not find it worthwhile to move. As costs fall, the out-migration response increases, and it increases disproportionately for type

1: $\Delta\tilde{\lambda}_{1ij,ru}/\Delta\tilde{\lambda}_{2ij,ru} = \psi^{(\epsilon-1)\kappa}$. Reductions in migration frictions therefore widen the compositional shift in rural forestry: Urban-CA households exit, while Forest-CA households remain and expand their share. In equilibrium, highway expansion raises effective forest productivity through the sorting channel formalized in equation 11.

Finally, in equilibrium, rural forest revenues balance against expenditures on urban goods, while urban wages balance against expenditures on forest goods. In the rural sector, revenues are divided among farmers, landowners, and input suppliers, with land rents adjusting to clear local land markets. In the urban sector, wages adjust so that the value of production matches demand from both rural and urban consumers. These conditions jointly determine the vector of land rents and wages that clear all markets. Highways enter by reducing trade and migration frictions, shifting sourcing probabilities and reallocating revenues, incomes, and labor across space. Appendix B.2 provides the full derivations. Appendix B.3 develops an exact hat-algebra version of the model, which maps changes in transportation and migration costs from highway expansion into equilibrium changes in wages, rents, and price indices without re-solving the model in levels.[36]

5 Quantitative Results

This section outlines the calibration of key model parameters and reports the quantitative results. We first describe how the model is disciplined by initial conditions, and then present simulations that quantify the aggregate effects of road connections on effective forest productivity.

5.1 Model Calibration

We calibrate the model to China in 2000. The baseline sample includes $N = 287$ cities with valid migration and trade data. We split each city into rural and urban parts, yielding $2N = 574$ locations. Starting from this baseline and using our estimated highway-induced reductions in trade and migration costs from 2000 to 2010, we solve for the counterfactual equilibrium and simulate the implied migration and spatial reallocation.

[36]The exact hat algebra method computes relative changes in equilibrium variables without re-solving the system in levels. Let w_i, r_i, P_i^f, P_i^u denote the baseline equilibrium under policy variables $\tau_{ni}^f, \tau_{ni}^u, \mu_{in,a_0a}$, and $w_i', r_i', P_i^{f'}, P_i^{u'}$ the new equilibrium under the updated policy. The hat variables then measure the ratios of new to old values, providing a tractable way to link changes in trade and migration frictions to adjustments in wages, land rents, and price indices.

Trade flows by sector We construct bilateral trade flows from the 42-sector, 30-province input–output table of Chen et al. (2023). Because the IO data are reported at the provincial level, we infer city-to-city flows by distributing each province-pair flow across the corresponding city pairs using observable proxies. For urban goods, we define the sector as manufacturing and mining, and allocate each interprovincial flow to city pairs in proportion to the GDP of the origin and destination cities. For forest goods, we start from the provincial aggregate for agriculture, forestry, animal husbandry, and fishery (AFHF). To isolate forestry within AFHF, we compute forestry's share in AFHF using the 135-sector national IO table for 2007 (National Bureau of Statistics of China, 2010) and apply this share to the AFHF flows. We then allocate provincial forestry flows across city pairs using the origin city's forest land area (from GlobeLand30) as a proxy for supply capacity and the destination city's forestry-related demand (from the 2004 Economic Census). Appendix B.4 provides full details.

Migration flows by type We construct bilateral migration flows from the 2000 Population Census, which reports each respondent's hukou location, current residence, education, and occupation at the county level. We map the hukou county to the origin city (rural vs. urban by hukou type) and the current county to the destination city. We restrict the sample to working-age individuals (16–64). We split migrants into two types by schooling: type 1 (Urban CA) has at least high school education, and type 2 (Forest CA) has middle school education or less. This split follows a standard selection logic, consistent with Lagakos and Waugh (2013): lower-education workers are more likely to sort into forestry-related activities, while more educated workers are more likely to find higher returns in the urban sector. Occupation codes identify forestry workers as residing in rural areas. When forestry cannot be cleanly separated from the broader AFHF workforce in migration records, we apportion rural hukou migration using the provincial forestry-to-AFHF labor ratio. Aggregating by origin–destination pairs yields a $2N \times 2N$ migration matrix for each type, L_{1in,a_0a}.

Travel-time reductions and implied cost changes We measure transport improvements using inter-county travel times from the Transportation Networks Dataset in Ma and Tang (2024) (see Section 2.2.2). We aggregate county-to-county times to the city level using weights proportional to the product of 2000 populations in the origin and destination counties. For each city pair (i, n), the change in travel time between 2000 and 2010 is $\widehat{times}_{in} = times_{i,n}^{2010}/times_{i,n}^{2000}$, which captures the relative decline in travel time induced by highway expansion. We map these travel-time changes into migration costs by specifying

the bilateral moving cost from (i, a_0) to (n, a) as

$$\ln \mu_{in,a_0a} = \beta_0 + \beta_1 \ln times_{in} + \beta_2 I_{in} \ln times_{in} + \beta_3 I_{in} + \nu_{a_0a}. \tag{18}$$

where I_{in} indicates within-province pairs and ν_{a_0a} are move-type fixed effects (rural–rural, rural–urban, urban–rural, urban–urban). The implied change in migration costs is therefore

$$\ln \hat{\mu}_{in,a_0a} = \beta_1 \ln \widehat{times}_{in} + \beta_2 I_{in} \ln \widehat{times}_{in}, \tag{19}$$

with parameter estimates reported in Appendix B.4. For trade costs, bilateral flows are observed only at the provincial level, so we do not estimate city-level gravity directly. Instead, we discipline the trade-cost response using the internal-trade gravity estimates in Fan (2019).

Other parameters We pin down the productivity advantage of type-2 over type-1 farmers, ψ, from the sorting condition in equation 16. Using calibrated migration flows, we compute the mean ratio $\tilde{\lambda}_{1ij,ru}/\tilde{\lambda}_{2ij,ru}$ and interpret it as $\psi^{(\epsilon-1)\kappa}$. Given externally chosen values for κ and ϵ, this mapping implies $\psi = 2.5$. The land cost share in forest production, υ, is calibrated from the Rural Fixed Observation Point survey. For each household, we impute land rent as managed forest area times the national average rent of about 910 yuan per hectare (Zhang et al., 2018), and take the ratio of imputed rent to forestry income. Averaging across households yields $\upsilon = 0.68$. The consumption share of forest products, η, is taken from the 2007 national IO table: agriculture and forestry goods account for about 10% of household final consumption, so we set $\eta = 0.10$. We set $\kappa = 1.5$ (Tombe and Zhu, 2019), $\epsilon = 3$ (Hsieh and Klenow, 2009), and the trade elasticity $\theta = 4$ (Simonovska and Waugh, 2014). Agglomeration and congestion elasticities follow Allen and Arkolakis (2022), with $\alpha = 0.1$ and $\beta = -0.3$. Appendix B.4 details the calibration, and Appendix Table B1 summarizes all parameter values.

5.2 Model Performance

We begin by assessing whether the model reproduces rural-to-urban migration, the key margin behind the urbanization patterns in the data. In the data, the counterpart is the share of a city's migrants whose destination is urban. Figure B1 compares this share in the 2010 Census to the model's predicted probabilities under highway expansion. Each point is a city. For Urban CA migrants (Panel (a)), the fit is tight: the slope is 0.65 with an R^2 of 0.49. Panel (b) shows a comparable fit for Forest CA migrants, indicating that the model captures the broad urbanization dynamics from 2000 to 2010. Across cities, Forest CA

migrants are consistently less likely to move to urban destinations than Urban CA migrants, as the model implies. Table 6 summarizes these differences using geometric averages. In the 2000 baseline (Column (1)), rural-to-urban migration probabilities are 34.6% for Urban CA and 16.8% for Forest CA. Under highway expansion (Column (2)), they rise to 36.9% and 18.3%, respectively. The implied increases (Column (3)) are 2.3 and 1.5 percentage points, showing that highways promote urbanization for both groups, with a larger response among Urban CA households.

Table 6: Highway Expansion and Rural-to-Urban Migration Probabilities

	(1) Initial Economy, 2000	(2) After Highway Expansion, 2010	(3) Highway-Induced Urbanization: (2)-(1)
Urban CA	34.61%	36.86%	2.25%
Forest CA	16.84%	18.32%	1.48%

Notes: This table reports rural-to-urban migration probabilities for two types of individuals: Urban CA (comparative advantage in urban production) and Forest CA (comparative advantage in forest production). Columns (1) and (2) present geometric averages across municipalities in the initial economy and after highway expansion, respectively. Column (3) shows the highway-induced change in migration probabilities.

This asymmetry implies that highways encourage selective migration: individuals with higher urban comparative advantage are more likely to relocate, while those with forest-specific skills remain disproportionately in rural areas. The resulting composition effect strengthens the concentration of labor with a comparative advantage in forestry, providing a potential channel through which transport improvements contribute to observed gains in effective forest productivity. This mechanism is consistent with the model's implications in Section 4.2.[37] Putting these migration results together with China's forest-tenure constraints and the stable forest cover in the data, we hold forest area fixed in the model. Any change in forest sales or quality therefore reflects intensive-margin adjustment, through productivity, management intensity, and market access—rather than an expansion of forest land.

5.3 Market Access, Selective Migration, and Forest Productivity

Better transportation strengthens sorting by comparative advantage and, through that channel, raises effective forest productivity. From equation 11, the proportional change in average forestry productivity in city i can be approximated by

$$\ln \hat{\bar{\psi}}_i \approx \Theta \ln(\hat{L}_{i2,r}/\hat{L}_{i1,r}), \tag{20}$$

[37] Figure B2 further compares the model-predicted rural workforce in the forest sector of each city—shaped by urbanization and selective migration—with the corresponding 2010 Census observations, both expressed in logarithms. The results yield slopes of $\beta \approx 1$ for both types, which further reinforces the reliability of our model.

where $\Theta > 0$ since $\psi > 1$ and $\epsilon > 1$ (See details in Appendix B.6). The equation links productivity gains to composition: when type-1 households, with stronger urban comparative advantage, leave rural areas more than type 2, the rural workforce tilts toward the more forestry-productive type, raising $\bar{\psi}_i$. Table 7 quantifies this margin. On average across cities, highway expansion increases effective forest productivity by 0.0012 log points, reflecting a 0.01 log-point rise in the rural type-2-to-type-1 ratio. Both types decline in rural areas, consistent with the higher rural-to-urban out-migration probabilities in Table 6. The larger outflow of type 1 nevertheless tilts the rural workforce toward type 2 and pushes up productivity. However, the small magnitude reflects offsetting reallocation. Highways improve matching by drawing skilled workers and managers toward high-return, well-connected places, raising productivity there but draining remote forest areas. These gains and losses largely offset in the aggregate. If the stock of forestry labor or skilled managers grows over time through entry, training, or improved practices, the productivity effect would be larger under the same mechanism.

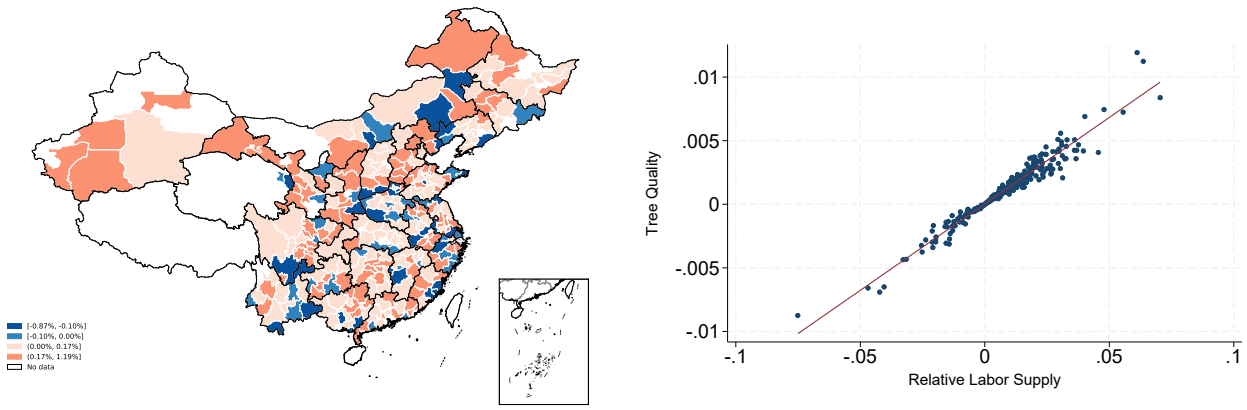
Table 7: Effective Forest Productivity and the Composition of Rural Labor

(1)	(2)	(3)	(4)
Effective Forest Productivity, $\ln \hat{\bar{\psi}}_i$	Rural Labor Composition, $\ln(\hat{L}_{i2,r}/\hat{L}_{i1,r})$	Urban CA, $\ln \hat{L}_{i1,r}$	Forest CA, $\ln \hat{L}_{i2,r}$
0.0012	0.0095	-0.0557	-0.0462

Notes: This table reports average log changes in effective forest productivity and rural labor composition relative to the year 2000 baseline. Values are simple averages across cities, consistent with the additive decomposition in Equation (20). Effective forest productivity improves despite a decline in the absolute number of both Urban CA and Forest CA farmers, reflecting the disproportionate out-migration of Urban CA individuals and the resulting increase in the relative share of Forest CA labor in rural areas.

The gains in effective forest productivity are not evenly distributed across regions. Panel (a) of Figure 3 illustrates the geographic variation, with warmer colors indicating improvements and cooler colors indicating declines. Most cities experience productivity gains following the highway expansion, but the magnitude differs across space. Panel (b) shows that these gains are larger in areas with a relatively higher concentration of farmers holding a rural comparative advantage, consistent with the mechanism in equation 20.

Decomposing the sources of forest sales To pinpoint the channels through which highways raise rural forest sales, we use exact-hat algebra to decompose the model-implied change in sales. Guided by China's forest-tenure institutions and policy constraints, and supported by the data, we treat forest area (and hence aggregate forest cover) as essentially fixed over our sample period. With land held constant, changes in forest sales can only reflect intensive-margin adjustment: higher effective forest productivity, stronger management incentives and input use, and improved access to downstream markets, rather than an



(a) Percentage Change in $\bar{\psi}_i$ (b) Percentage Change in $\bar{\psi}_i$ and $L_{i2,r}/L_{i1,r}$

Figure 3: Model-Predicted Changes in Effective Forest Productivity across Cities

Notes: Panel (a) maps the percentage change in effective forest productivity, $\bar{\psi}_i$, across cities predicted by the model. Warmer (cooler) colors indicate larger (smaller) improvements. Panel (b) plots the predicted change in effective forest productivity against the corresponding change in the relative supply of Forest CA to Urban CA farmers ($L_{i2,r}/L_{i1,r}$).

expansion of land under forest cover. The resulting decomposition is:

$$\hat{R}_i^f = \underbrace{(\hat{r}_i)^{\upsilon(1-\epsilon)}}_{\text{Land Rent}} \times \underbrace{\left(\hat{P}_i^u\right)^{(1-\upsilon)(1-\epsilon)}}_{\text{Urban Input}} \times \underbrace{\left(\hat{\bar{\psi}}_i\right)^{\epsilon-1}}_{\text{Effective Forest Productivity}} \times \underbrace{\hat{M}_i^f}_{\text{Market Access}} . \tag{21}$$

Appendix B.7 derives equation 21 and shows how improved transportation maps into four general-equilibrium forces: land rents, the cost of urban intermediates, effective forestry productivity (captured by $\bar{\psi}_i$, which embeds changes in worker composition and management intensity), and market access. With $\epsilon > 1$ and $0 < \upsilon < 1$, higher rents and higher input costs work against expansion along the intensive margin, whereas stronger market access and higher effective productivity increase forest sales.

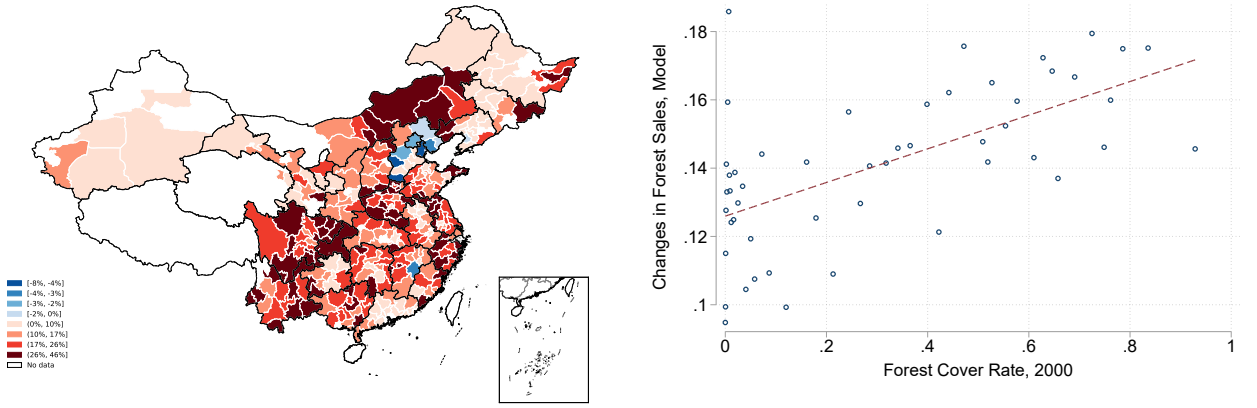
Table 8: Decomposition of Changes in Forest Sales

(1) Sale, $\hat{R}_i^f$	(2) Land Rent, $(\hat{r}_i)^{\upsilon(1-\epsilon)}$	(3) Urban Input, $\left(\hat{P}_i^u\right)^{(1-\upsilon)(1-\epsilon)}$	(4) Effective Forest Productivity, $\left(\hat{\bar{\psi}}_i\right)^{\epsilon-1}$	(5) Market Access, $\hat{M}_i^f$
1.169	0.808	1.027	1.002	1.405

Note: This table decomposes the change in forest sales into its contributing components, expressed as hat changes relative to the year 2000 baseline. Reported values are geometric averages across municipalities, consistent with the multiplicative structure of Equation 21. Increases in sales are driven primarily by cheaper access to urban inputs, improved market access, and higher effective forest productivity. The latter reflects selective migration that leaves rural areas increasingly concentrated in Forest CA farmers, while the input channel is consistent with the entry of specialized forestry firms. By contrast, higher land rents partially offset these gains.

Table 8 reports the decomposition of the model-implied change in forest sales. On net, highway expansion raises forest sales by about 17%. The main force is improved market

access, which accounts for 40.5% of the gross increase. Cheaper urban inputs add a further 2.7%, consistent with the entry and scaling of specialized forestry activities that rely on purchased intermediates and services. The effective-productivity contribution is small (0.2%) because, in general equilibrium, sorting is largely a reallocation rather than a net gain. Better connectivity makes nearby locations more attractive, so skilled workers and managers concentrate where returns are highest. That raises productivity in connected areas, but it can reduce productivity in more remote forest areas that lose talent. When the decomposition aggregates across locations, these gains and losses mostly offset, leaving a small net term.[38] Offsetting these gains, land rents rise with improved connectivity and reduce forest sales by about 19.2%.



(a) Percentage Change in R_i^f (b) Percentage Change in R_i^f and Forest Coverage

Figure 4: Model-Predicted Changes in Forest Sales across Cities

Notes: Panel (a) maps the predicted percentage change in forest sales across cities. Warmer colors indicate larger increases. Panel (b) plots changes in forest sales against initial forest coverage in 2000 using a binned scatterplot.

As cities differ in both highway connectivity and baseline forestry endowments, the gains from highway expansion are uneven. Figure 4b maps these distributional effects. Panel (a) shows that the geography of forest-sales growth closely tracks the spatial pattern of forest-productivity gains in Figure 3. Panel (b) links changes in forest sales to initial forest coverage in 2000 using a binned scatterplot.[39] The relationship is strongly positive: sales

[38] In addition to this possibility, it is helpful to view $\hat{\tilde{\psi}}_i$ as a residual term in the exact-hat decomposition: it captures productivity changes not already accounted for by prices and market access. Since market access is the primitive shock and the force that triggers sorting, part of the quality upgrading induced by reallocation is absorbed by $\hat{M}_i^f$ rather than appearing separately in $\hat{\tilde{\psi}}_i$. Moreover, the model holds the stock of forestry-relevant skills fixed, so sorting mainly reallocates existing talent; any expansion of skilled forestry labor through entry, training, or improved practices would mechanically amplify the productivity contribution even under the same mechanism.

[39] To reduce noise from cities with near-zero initial forest coverage, we bin cities by baseline coverage.

gains are concentrated in initially more forested cities. This pattern matches the model's core prediction that improved market access reinforces comparative advantage, by strengthening selective migration and facilitating the entry and expansion of specialized forestry firms.

6 Conclusions

China's highway expansion between 2000 and 2010 improved forest *quality* rather than causing deforestation in our setting. A 10 km reduction in distance to a highway raises standing timber volume by approximately 2%, with effects concentrated within 1–20 km of new roads. The implied gains correspond to roughly 216.9–551.9 million tons of CO_2 sequestered, illustrating the potential environmental benefits of infrastructure expansion when intensive-margin responses dominate.

These patterns point to a mechanism in which improved connectivity raises forest quality by strengthening incentives for managing existing forests. During our study period, forest land-use regulations limited large-scale land conversion, so new highways primarily affected the returns to forest management by improving access to downstream processors and urban input markets. Collective forest-tenure reforms further reinforced this response by clarifying use rights and facilitating more intensive forest management. Consistent with this interpretation, forest quality gains are largest in collectively owned plots subject to tenure reform, and household data indicate increased specialization: producers with forestry experience expand timber sales, while others reallocate labor toward off-farm employment.

A quantitative spatial model, calibrated to migration, trade, and forest data, rationalizes these behavioral adjustments and clarifies their broader economic implications. Holding forest area fixed, the model implies that highway expansion increased forest sales by about 17% over the study period, reflecting intensive-margin improvements in effective forest productivity and management. An exact-hat decomposition shows that improved downstream market access is the primary driver of these gains, complemented by better input access and selective reallocation of labor.

The results illustrate that infrastructure need not erode natural capital. When transport investment strengthens incentives for intensive forest management, it can support higher forest quality and increased carbon storage. This perspective highlights an important implication for infrastructure evaluation: market integration can shift adjustment toward productivity and specialization in resource management rather than extensive land-use change. Admittedly, our analysis focuses on forest biomass and carbon storage and does not capture other environmentally valuable outcomes, such as biodiversity. Nor do we rule out potential adverse effects of highways on relatively undisturbed natural ecosystems. Our results

are best interpreted as applying to managed forest landscapes where human presence and economic activity are already substantial and largely unavoidable. Future work may explore to a broader range of forest-related environmental outcomes and the long-run climate consequences of transport investment.

References

Acemoglu, Daron, Simon Johnson, and James A. Robinson. 2005. "The rise of Europe: Atlantic trade, institutional change, and economic growth." *American Economic Review*, 95(3): 546–579.

Ahlfeldt, Gabriel M., Stephen J. Redding, Daniel M. Sturm, and Nikolaus Wolf. 2015. "The economics of density: Evidence from the Berlin Wall." *Econometrica*, 83(6): 2127–2189.

Allen, Treb, and Costas Arkolakis. 2022. "The welfare effects of transportation infrastructure improvements." *The Review of Economic Studies*.

Amatulli, G., S. Domisch, M.-N. Tuanmu, B. Parmentier, A. Ranipeta, J. Malczyk, and W. Jetz. 2018. "A suite of global, cross-scale topographic variables for environmental and biodiversity modeling." *Scientific Data*, 5: 180040.

Araujo, Claudio, Catherine Araujo Bonjean, Jean-Louis Combes, Pascale Combes Motel, and Eustaquio J. Reis. 2009. "Property rights and deforestation in the Brazilian Amazon." *Ecological Economics*, 68(8): 2461–2468.

Asher, Sam, and Paul Novosad. 2020. "Rural roads and local economic development." *American Economic Review*, 110(3): 797–823.

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

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.

Banerjee, Abhijit, and Lakshmi Iyer. 2005. "History, institutions, and economic performance: The legacy of colonial land tenure systems in India." *American Economic Review*, 95(4): 1190–1213.

Banerjee, Abhijit, Esther Duflo, and Nancy Qian. 2020. "On the road: Access to transportation infrastructure and economic growth in China." *Journal of Development Economics*, 145: 102442.

Baragwanath, Kathryn, and Ella Bayi. 2020. "Collective property rights reduce deforestation in the Brazilian Amazon." *Proceedings of the National Academy of Sciences*, 117(34): 20495–20502.

Baragwanath, Kathryn, and Nilesh Shinde. 2025. "Beyond the canopy: How satellite data detection thresholds influence policy evaluation and deforestation behavior." *Journal of Environmental Economics and Management*, 134: 103219.

Baum-Snow, Nathaniel. 2007. "Did Highways Cause Suburbanization?" *Quarterly Journal of Economics*, 122(2): 775–805.

Baum-Snow, Nathaniel, J. Vernon Henderson, Matthew A. Turner, Qinghua Zhang, and Loren Brandt. 2020. "Does investment in national highways help or hurt hinterland city growth?" *Journal of Urban Economics*, 115: 103124.

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

Bennett, Michael T. 2008. "China's sloping land conversion program: Institutional innovation or business as usual?" *Ecological Economics*, 65(4): 699–711.

Besley, Timothy. 1995. "Property rights and investment incentives: Theory and evidence from Ghana." *Journal of Political Economy*, 103(5): 903–937.

Blackman, Allen, Leonardo Corral, Eirivelthon Santos Lima, and Gregory P. Asner. 2017. "Titling indigenous communities protects forests in the Peruvian Amazon." *Proceedings of the National Academy of Sciences*, 114(16): 4123–4128.

Bogart, Dan, Xuesheng You, Eduard J. Alvarez-Palau, Max Satchell, and Leigh Shaw-Taylor. 2022. "Railways, divergence, and structural change in 19th century England and Wales." *Journal of Urban Economics*, 128: 103390.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting event study designs: Robust and efficient estimation." *The Review of Economic Studies*, 91(1): 3–35.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2014. "Challenges of working with the Chinese NBS firm-level data." *China Economic Review*, 30: 339–352.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. "The political economy of deforestation in the tropics." *The Quarterly Journal of Economics*, 127(4): 1707–1754.

Burgess, Robin, Remi Jedwab, Edward Miguel, Ameet Morjaria, and Gerard Padró i Miquel. 2015. "The Value of Democracy: Evidence from Road Building in Kenya." *American Economic Review*, 105(6): 1817–1851.

Chen, Quanrun, Yuning Gao, Chen Pan, Dingyi Xu, Kun Cai, Dabo Guan, Qi He, Shantong Li, Wanqi Liu, Bo Meng, Zhi Wang, Yang Wang, Xianchun Xu, Peihao Yang, Meichen Zhang, and Yuanqi Zhou. 2023. "An interprovincial input–output database distinguishing firm ownership in China from 1997 to 2017." *Scientific Data*, 10(1): 293.

Chen, Zuoqi, Bailang Yu, Yuyu Zhou, Hongxing Liu, Chengshu Yang, Kaifang Shi, and Jianping Wu. 2019. "Mapping Global Urban Areas From 2000 to 2012 Using Time-Series Nighttime Light Data and MODIS Products." *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(4): 1143–1153.

China Perspectives. 2001. "The natural forest protection program: China's logging ban policy." *China Perspectives*.

China Transport Yearbook Press. 2000. *China Transport Yearbook 2001.* Beijing, China:China Transport Yearbook Press.

China Transport Yearbook Press. 2011. *China Transport Yearbook 2011.* Beijing, China:China Transport Yearbook Press.

Donaldson, Dave. 2018. "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure." *American Economic Review*, 108(4-5): 899–934.

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American economic growth: A "market access" approach." *The Quarterly Journal of Economics*, 131(2): 799–858.

Duranton, Gilles, and Matthew A. Turner. 2011. "The Fundamental Law of Road Congestion: Evidence from US Cities." *American Economic Review*, 101(6): 2616–2652.

Eaton, Jonathan, and Samuel Kortum. 2002. "Technology, geography, and trade." *Econometrica*, 70(5): 1741–1779.

Engert, Jayden E., Mason J. Campbell, Joshua E. Cinner, Yoko Ishida, Sean Sloan, Jatna Supriatna, Mohammed Alamgir, Jaime Cislowski, and William F. Laurance. 2024. "Ghost roads and the destruction of Asia-Pacific tropical forests." *Nature*, 629(8011): 370–375.

Esfahani, Hadi Salehi, and María Teresa Ramírez. 2003. "Institutions, infrastructure, and economic growth." *Journal of Development Economics*, 70(2): 443–477.

Faber, Benjamin. 2014. "Trade integration, market size, and industrialization: Evidence from China's national trunk highway system." *The Review of Economic Studies*, 81(3): 1046–1070.

Fajgelbaum, Pablo, and Stephen J. Redding. 2022. "Trade, Structural Transformation, and Development: Evidence from Argentina 1869–1914." *Journal of Political Economy*, 130(5): 1246–1318.

Fan, Jingting. 2019. "Internal geography, labor mobility, and the distributional impacts of trade." *American Economic Journal: Macroeconomics*, 11(3): 252–288.

Farooq, T. H., H. Q. Hu, J. Xu, et al. 2021. "Perspectives of plantation forests in the sustainable forest development of China." *iForest*, 14: 225–235.

Fenske, James, Namrata Kala, and Jinlin Wei. 2021. "Railways and Cities in India." University of Warwick, Department of Economics 1349.

Food and Agriculture Organization of the United Nations. 2012. "Forest tenure reform in China: Impacts and challenges."

Food and Agriculture Organization of the United Nations. 2020. "Global Forest Resources Assessment 2020: Key findings." Food and Agriculture Organization of the United Nations, Rome.

Forman, Richard T. T., and Lauren E. Alexander. 1998. "Roads and their major ecological effects." *Annual Review of Ecology and Systematics*, 29(1): 207–231.

Fretz, Stephan, Raphaël Parchet, and Frédéric Robert-Nicoud. 2021. "Highways, Market Access, and Spatial Sorting." *The Economic Journal*. ueab070.

Ghani, Ejaz, Arti Grover Goswami, and William R. Kerr. 2015. "Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing." *The Economic Journal*, 126(591): 317–357.

Goldman, Elizabeth. 2024. "Forest Extent Indicator." *Global Forest Review, World Resources Institute*, Last updated April 4, 2024. Accessed 2026-02-02.

Goldstein, Markus, and Christopher Udry. 2008. "The profits of power: Land rights and agricultural investment in Ghana." *Journal of Political Economy*, 116(6): 981–1022.

Gollin, Douglas, and Julien Wolfersberger. 2025. "Agricultural trade and deforestation: The role of new roads." *The Economic Journal*, ueaf125.

He, Guojun, Yuhang Xie, and Bing Zhang. 2020. "Expressways, GDP, and the environment: The case of China." *Journal of Development Economics*, 145: 102485.

He, Jun, Bereket Kebede, Adrian Martin, and Nicole Gross-Camp. 2020. "Privatization or communalization: A multi-level analysis of changes in forest property regimes in China." *Ecological Economics*, 174: 106629.

Hersbach, H., B. Bell, P. Berrisford, et al. 2020. "The ERA5 global reanalysis." *Quarterly Journal of the Royal Meteorological Society*.

Herzog, Ian. 2021. "National transportation networks, market access, and regional economic growth." *Journal of Urban Economics*, 122: 103316.

He, Yue, Shilong Piao, Philippe Ciais, Hao Xu, and Thomas Gasser. 2024. "Future land carbon removals in China consistent with national inventory." *Nature Communications*, 15: 10426.

Hornbeck, Richard. 2010. "Barbed wire: Property rights and agricultural development." *The Quarterly Journal of Economics*, 125(2): 767–810.

Hsieh, Chang-Tai, and Peter J. Klenow. 2009. "Misallocation and manufacturing TFP in China and India." *The Quarterly Journal of Economics*, 124(4): 1403–1448.

Hsu, Feng-Chi, Kimberly E. Baugh, Tilottama Ghosh, Mikhail Zhizhin, and Christopher D. Elvidge. 2015. "DMSP-OLS Radiance Calibrated Nighttime Lights Time Series with Intercalibration." *Remote Sensing*, 7(2): 1855–1876.

Interagency Working Group on Social Cost of Greenhouse Gases. 2016. "Technical support document: Technical update of the social cost of carbon for regulatory impact analysis under Executive Order 12866." United States Government.

Jedwab, Remi, and Adam Storeygard. 2021. "The Average and Heterogeneous Effects of Transportation Investments: Evidence from Sub-Saharan Africa 1960-2010." *Journal of the European Economic Association*. jvab027.

Jedwab, Remi, and Alexander Moradi. 2016. "The Permanent Effects of Transportation Revolutions in Poor Countries: Evidence from Africa." *Review of Economics and Statistics*, 98(2): 268–284.

Jin, Lu, Yuanyuan Yi, and Jintao Xu. 2020. "Forest carbon sequestration and China's potential: The rise of a nature-based solution for climate change mitigation." *China Economic Journal*, 13(2): 200–222.

Kaczan, David J. 2020. "Can Roads Contribute to Forest Transitions?" *World Development*, 129: 104898.

Lagakos, David, and Michael E. Waugh. 2013. "Selection, agriculture, and cross-country productivity differences." *American Economic Review*, 103(2): 948–980.

Lesiv, Myroslava, Dmitry Shchepashchenko, Marcel Buchhorn, Linda See, Martina Dürauer, Iva Georgieva, Martin Jung, Florian Hofhansl, et al. 2022. "Global forest management data for 2015 at a 100 m resolution." *Scientific Data*, 9(1): 199.

Liu, Jianguo, Shuxin Li, Zhiyun Ouyang, Chi-yung Tam, and Xiaodong Chen. 2008. "Ecological and socioeconomic effects of China's policies for ecosystem services." *Proceedings of the National Academy of Sciences*, 105(28): 9477–9482.

Lundberg, Matthew, Matthew Clark, Ian Herzog, and Yue Yu. 2025. "Transportation infrastructure, market access, and deforestation." SSRN Working Paper No. 5355483.

Ma, Lin, and Yang Tang. 2024. "The distributional impacts of transportation networks in China." *Journal of International Economics*, 103873.

Ma, Lin, Yuan Mei, Qunfeng Wu, and Mingzhi Xu. 2025. "Connectivity and Selective Rural Migration for Agricultural Productivity Gains." *Available at SSRN 5381231*.

Michaels, Guy. 2008. "The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System." *Review of Economics and Statistics*, 90(4): 683–701.

Ministry of Transport of the People's Republic of China. 2004. "Technical standard of highway engineering." Ministry of Transport of the People's Republic of China JTG B01-2003. Issued January 29, 2004.

National Bureau of Statistics of China. 2010. "Input–output tables of China 2007." Data release.

Peng, Cong, and Yao Wang. 2026. "Roads to Development? Urbanization Without Growth in Zambia." Available at SSRN.

Pfeifroth, U., et al. 2024. "Satellite-based climate data records of surface solar radiation (SARAH-3)." *Earth System Science Data*.

Preferred by Nature. 2021. "China timber legality risk assessment, version 2.0."

Santika, Truly, Kerrie A. Wilson, Sugeng Budiharta, Elizabeth A. Law, Tun-Min Poh, Marc Ancrenaz, Matthew J. Struebig, and Erik Meijaard. 2021. "No aggregate deforestation reductions from rollout of community land titles in Indonesia yet." *Proceedings of the National Academy of Sciences*, 118(43): e2100741118.

Siikamäki, Juha, Jintao Xu, and Yanjun Ji. 2015. "Post-reform forest land markets in China." *China Economic Review*, 32: 220–237.

Simonovska, Ina, and Michael E. Waugh. 2014. "The elasticity of trade: Estimates and evidence." *Journal of International Economics*, 92(1): 34–50.

Souza-Rodrigues, Eduardo. 2019. "Deforestation in the Amazon: A unified framework for estimation and policy analysis." *The Review of Economic Studies*, 86: 2713–2744.

Standing Committee of the National People's Congress. 2019. "Forest Law of the People's Republic of China (2019 revision)." English translation hosted by the Ministry of Ecology and Environment.

State Forestry Administration. 2010. *Forest resources statistics of China: The 7th national forest inventory (2004–2008).* Beijing:State Forestry Administration.

State Forestry Administration. 2014. *China forest resources report (2009–2013).* China Forestry Publishing House. National forest resource inventory report.

State Forestry Administration of the People's Republic of China. 2004. "Technical regulations for national forest continuous inventory." Beijing, China, (In Chinese).

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Tateishi, R., U. Bayaer, H. Al-Bilbisi, M. Aboel Ghar, J. Tsend-Ayush, T. Kobayashi, A. Kasimuf, N. Thanh Hoan, A. Shalaby, B. Alsaaideh, T. Enkhzaya, Gegentana, and H. P. Sato. 2011. "Production of global land cover data – GLCNMO." *International Journal of Digital Earth*, 4(1): 22–49.

Tombe, Trevor, and Xiaodong Zhu. 2019. "Trade, migration, and productivity: A quantitative analysis of China." *American Economic Review*, 109(5): 1843–1872.

USDA Foreign Agricultural Service. 2017. "Strong demand and declining production encourage wood imports." Summarizes China's 8th National Forest Resource Inventory (2009–2013).

U.S. Environmental Protection Agency. 2022. "Report on the social cost of greenhouse gases: Estimates incorporating recent scientific advances." U.S. Environmental Protection Agency. Draft report.

Vilela, Thais, Alfonso Malky Harb, Aaron Bruner, Vera Laísa da Silva Arruda, Vivian Ribeiro, Ane Auxiliadora Costa Alencar, Annie Julissa Escobedo Grandez, Adriana Rojas, Alejandra Laina, and Rodrigo Botero. 2020. "A better Amazon road network for people and the environment." *Proceedings of the National Academy of Sciences*, 117(13): 7095–7102.

Voigtländer, Nico, and Hans-Joachim Voth. 2026. "Highway to Hitler." *American Economic Journal: Applied Economics*, 18(1): 120–59.

Walker, Sarah, Jennifer Alix-García, Anne Bartlett, and Alice Calder. 2025. "Land tenure security and deforestation: Evidence from a framed field experiment in Uganda." *Journal of Environmental Economics and Management*, 131: 103137.

Wang, Zhi, Mingwei Xu, Qinghua Zhang, and Li-An Zhou. 2025. "Greening along the slope and resource reallocation: Evidence from China's Grain for Green program." SSRN Working Paper No. 5111936.

World Bank. 2007. "Domestic trade impacts of the expansion of the national expressway network in China." World Bank, East Asia and Pacific Region EASTR Working Paper 14.

World Bank Group. 2002. "China—Sustainable forestry development project." World Bank Project Appraisal Document 22271, Washington, D.C.

Xie, Lunyu, Peter Berck, and Jintao Xu. 2016. "The effect on forestation of the collective forest tenure reform in China." *China Economic Review*, 38: 116–129.

Xu, X. 2015. "Economic regulation and the national trunk highway system in China." *Utilities Policy*, 35: 30–41.

Yi, Yuanyuan, Fredrik Carlsson, Gunnar Köhlin, and Jintao Xu. 2022. "Allocative efficiency or misallocation of resources? The emergence of forestland rental markets and the forest devolution reform in China." *European Review of Agricultural Economics*.

Zhang, Ying, Cheng Sun, Yuhui Sun, and Yu Liu. 2018. "Forest resources accounting in China." *Open Journal of Forestry*, 8(1): 66–80.

A Empirical Appendix

A.1 Appendix Figures

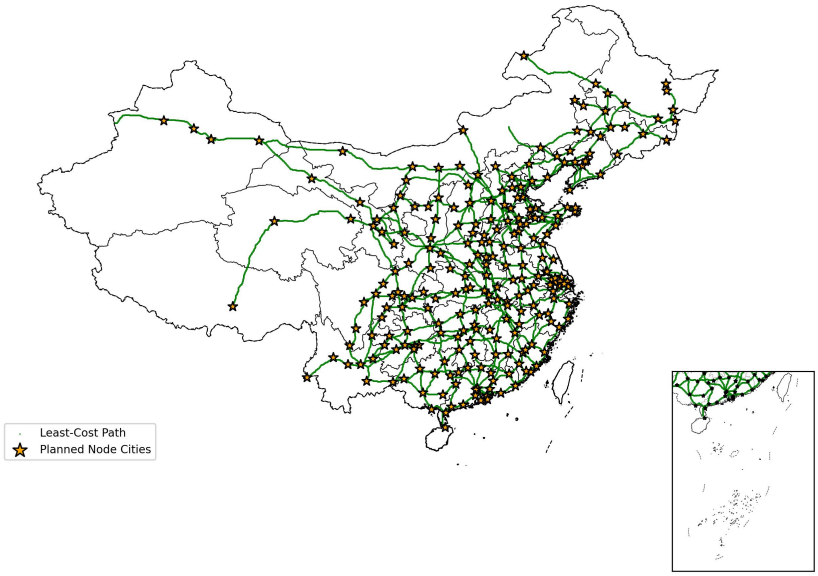
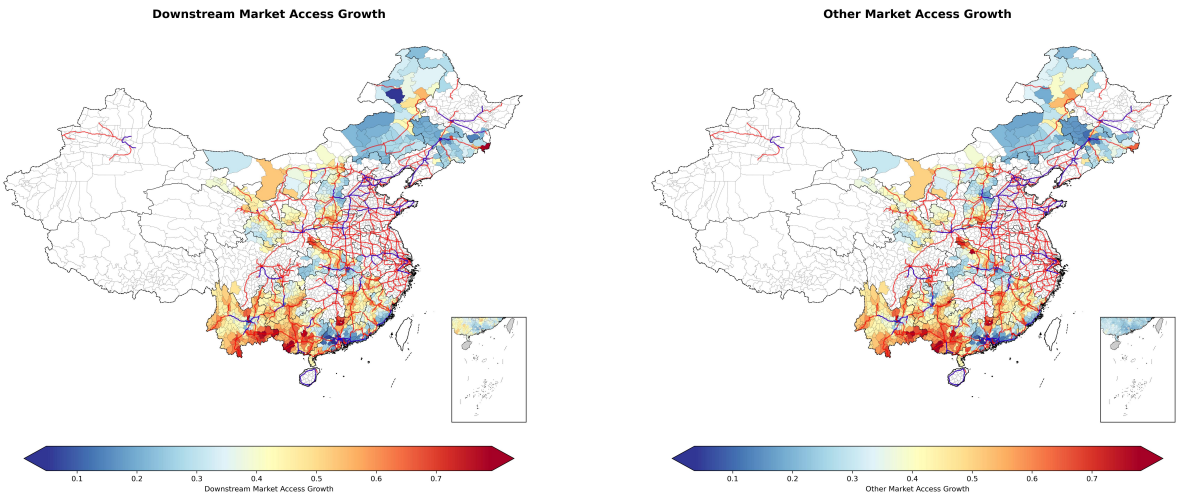


Figure A.1: "7–9–18" Trunk-network Plan

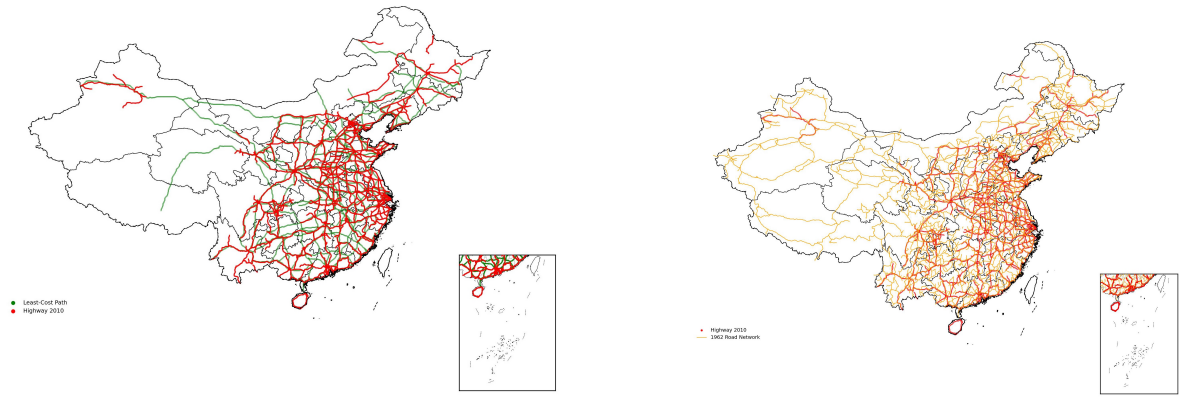
Note: The map shows the National Expressway Network Plan (NENP, the "7-9-18" grid, formerly NTHS), approved by the State Council in 2004. This plan aimed to establish an 85,000 km network connecting all provincial capitals and most cities with an urban population over 200,000. We construct the least-cost path route between each pair of planned target cities, detailed in Section A.4. Source: China's Ministry of Transport (2005) / National Development and Reform Commission (NDRC).



(a) RE_c^{Specific} (b) RE_c^{Other}

Figure A.2: Impact of Highway Expansion on Market Access

Note: This figure maps the relative change in county-level market access (RE_c) resulting from highway expansion between 2000 and 2010. The decomposition of market access is based on Equation (5). Panel (a) shows the change in market access to the pooled timber-intensive downstream industries (paper, wood furniture, and wood products), which is the Road-Induced Specific Market Access ($\text{RE}_c^{\text{Specific}}$). Panel (b) shows the change in market access to all other manufacturing sectors, measuring the Road-Induced Other Market Access ($\text{RE}_c^{\text{Other}}$).



(a) Least-Cost Path and 2010 Highways (b) 1962 Road and 2010 Highways

Figure A.3: Planned Highway Network and Historical Road Infrastructure

Note: This figure displays the 2010 National Expressway Network (red lines) alongside the two instrumental variables. Panel (a) depicts the Least-Cost Path (LCP) network (green lines), which simulates cost-minimized routes connecting major nodal cities designated in the NENP. Panel (b) shows the 1962 historical road network (orange lines).

A.2 Appendix Table

Table A.1: Summary Statistics of Plot-Level Variables

Panel A: NFI Survey Data	N	Mean	Sd	Min	Max
Forest Quality (2003 Wave)					
Volume (m^3)	18111	6.04	5.86	0.00	31.91
Coverage (%)	18111	56.25	22.91	0.00	100.00
Average Height (dm)	18111	96.70	50.03	0.00	235.00
DBH (mm)	18111	126.73	69.01	0.00	375.00
Forest Quality (2013 Wave)					
Volume (m^3)	18111	7.67	5.94	0.00	32.38
Coverage (%)	18111	62.26	18.97	20.00	99.00
Average Height (dm)	18111	112.26	44.67	20.00	234.00
DBH (mm)	18111	117.92	72.57	9.00	358.00
Forest Quality (Diff)					
Δ ln(DBH)	17337	-0.27	1.00	-3.73	4.71
Δ ln(Avg. Height)	17643	0.17	0.46	-2.38	4.05
Δ Coverage	18111	6.00	21.80	-80.00	92.00
Δ ln(Volume)	16945	0.38	0.72	-4.06	5.13
Forest Tenure					
State-Operated	18111	0.46	0.50	0.00	1.00
Collective-Operated	18111	0.35	0.48	0.00	1.00
Private-Operated	18111	0.19	0.39	0.00	1.00
Panel B: Other Plot-Level Data	N	Mean	Sd	Min	Max
Highway Distance					
Dis. to Highway (2000, km)	18111	174.55	128.26	0.00	796.70
Dis. to Highway (2010, km)	18111	61.33	63.38	0.00	445.28
Share of Dis. to highway < 20 (2000)	18111	0.03	0.18	0.00	1.00
Share of Dis. to highway < 20 (2010)	18111	0.24	0.42	0.00	1.00
Dis. to least-cost IV(km)	18111	67.31	88.04	0.08	548.61
Dis. to 1962 Road(km)	18111	40.09	36.17	0.00	255.51
Topography					
Slope	18111	18.12	7.85	0.78	44.63
Slope $> 15°$	18111	0.34	0.47	0.00	1.00
Proximity					
Dis. to County Seat(km)	18111	25.05	22.68	0.43	165.37
Dis. to Urban Area(km)	18111	87.28	65.65	0.21	461.46
Land Cover Shares					
Cultivated (%,1500m radii,2000)	18111	13.66	18.93	0.00	100.00
Forest (%,1500m radii,2000)	18111	74.71	26.95	0.00	100.00
Diff. Cultivated (%,1500m radii)	18111	0.17	3.28	-77.18	71.80
Diff. Forest (%,1500m radii)	18111	-0.18	8.33	-92.79	100.00

Note: Panel A reports plot-level variables from the National Forest Inventory (NFI). Missing values are coded as zero, following the survey definition for plots with "no measurable above-scale trees". Panel B reports additional variables constructed from external data sources. Land-cover measures are expressed as percentages within 1,500 m buffers, while means of indicator variables (e.g., slope $> 15°$) can be interpreted as shares. Log-transformed variables may have fewer observations due to zeros in the underlying values.

Table A.2: Summary Statistics of County-Level Variables

	N	mean	sd	min	max
Demographics					
Population (10k people)	904	41.09	29.85	2.23	273.51
GDP per capita (10k yuan)	904	0.50	0.60	0.06	10.23
Urbanization Rate	904	0.29	0.22	0.00	1.00
Primary Employment Share	904	70.92	20.43	0.21	96.19
Secondary Employment Share	904	12.05	11.58	0.49	81.18
Forest and Area					
Forest Area/Total Area	904	0.44	0.26	0.00	0.93
County Area (km^2)	904	3093.94	5373.33	53.21	87306.89
Forest Growth (Log Diff, GEE)	896	0.05	0.51	-2.25	7.13
Annual Forest Planting (ha)	340	3272.97	2854.73	88.44	15903.33
Highway Connectivity					
Connected 2000	904	0.10	0.29	0.00	1.00
Connected 2010	904	0.38	0.49	0.00	1.00
Median Dist. Hwy 2000	904	122.23	90.72	0.00	658.65
Median Dist. Hwy 2010	904	37.38	35.70	0.00	339.96
Firms					
Timber Firms Entry	904	77.55	88.25	0.00	975.00
Non-Timber Agriculture Firms Entry	904	665.63	879.65	24.00	16230.00
Non-Agriculture Firms Entry	904	14852.67	19159.99	615.00	300848.00

Note: All variables are aggregated at the county level. Population and GDP per capita are expressed in 10,000 persons and 10,000 yuan, respectively. Using data from the Global Land 30 dataset, forest area shares are calculated as the ratio of forested land to total county area, and forest growth is measured as the log difference in forest area between 2000 and 2010. Median highway distance represents the median distance from forest plots within the county to the nearest highway segment; accordingly, a county is defined as connected to the highway if this distance is less than 20 km in the respective year.

Table A.3: Variable Definitions and Data Sources

Variable Category	**Variables**	**Data Source**
Panel A: Forest Quality (Plot Level)		
Plot Forest Characteristics	Timber Volume (m^3), Canopy Coverage (%), Average Tree Height (dm), Diameter at Breast Height (DBH, mm)	National Forest Inventory (NFI), 6th (1999–2003) and 8th (2009–2013) waves
Panel B: Highway and Infrastructure (Plot and County Level)		
Highway	Distance to nearest highway (2000, 2010, 2000–2010 Newly Constructed); Share of area/plots within 20 km of highway; Connected dummy	Transportation Networks Dataset (Ma and Tang, 2024)
Historical Roads (IV)	Distance to 1962 Road	Historical Road Maps / Transportation Networks Dataset (Baum-Snow et al., 2017)
Market Access	County-pair travel times	Ma and Tang (2024) and National Fundamental Geographic Information Dataset (2000)
Panel C: Topography and Geography (Plot Level)		
Topography	Slope	SRTM 30 m Digital Elevation Model (DEM)
Urban Proximity	Distance to nearest urban area and county seat	Chen et al. (2019) and National Fundamental Geographic Information Dataset
Panel D: Land Cover and Forestry Activities (County/Local)		
Land Use	County-level forest area; Share of Forest/Cultivated Land	GlobeLand 30 (2000, 2010)
Forest Planting	Annual area of approved artificial afforestation	China Forestry Statistical Yearbook (2002–2010)
Panel E: Socioeconomic Controls (County Level)		
Demographics	Population, Urbanization Rate, Employment Shares (Primary/Secondary)	Fifth National Population Census (2000)
Economic Development	GDP per capita	Prefecture and County Fiscal Statistics (2000)
Panel F: Firm Dynamics (County Level)		

Continued on next page

Table A.3 – *Continued from previous page*

Variable Category	**Variables**	**Data Source**
Firm Entry	Number of new Timber, Non-Timber Agriculture, and Non-Agriculture firms	SAIC Enterprise Registration Data (2000–2010)
Panel G: Climate Controls		
Weather	Mean annual temperature, precipitation, sunshine hours	ERA5 (Hersbach et al., 2020), SARAH-3 (Pfeifroth et al., 2024)
Panel H: Economic Flows and Structural Linkages		
Migration Matrix	Inter-prefectural and inter-sectoral population migration matrix (2000, 2010)	Microdata from Fifth and Sixth National Population Census
Trade Flows	Inter-prefectural forestry and urban-goods trade flows	2007 National and Inter-provincial IO Tables (Chen et al., 2023) ; National Economic Census (2004)

Note: This table summarizes the data sources for the key variables used in the descriptive statistics and regression analysis.

Table A.4: NDVI Changes within Buffers around Forest Survey Plots, 2000–2010

Buffer Size	250m	1000m	5000m
mean	-0.0018	-0.0019	-0.0019
sd	(0.0528)	(0.0474)	(0.0404)

Note: The table reports mean changes in NDVI from 2000 to 2010, with standard deviations in parentheses, calculated within three buffer sizes (250 m, 1,000 m, and 5,000 m) around forest survey plots. NDVI changes are similar across buffer sizes, suggesting that the survey itself does not affect surrounding forest dynamics.

Table A.5: Effects of Highway Access on Other Forest Outcomes

	(1) Δ ln(Volume)	(2) Δ Coverage	(3) Δ ln(DBH)	(4) Δ ln(Avg. Height)
$Dist_{icp}^{HW}$	-0.021***	-0.360**	-0.010**	-0.007
	(0.008)	(0.153)	(0.004)	(0.005)
Observations	16226	17350	16602	16896
Mean of Dep. Var.	0.37	6.01	-0.28	0.17
F-statistic in 1st stage	45.5	41.8	39.8	40.2
Province fixed effects	Yes	Yes	Yes	Yes
County-level controls	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes

Note: Outcomes are differences between two waves of surveys: Δ ln(Volume), Δ Coverage (percentage points), Δ ln(DBH), and Δ ln(Avg. Height). The continuous distance to the nearest highway is instrumented by the distance to the LCP and the 1962 Road Network. The sample excludes plots within 20 km of any highway in 2000 and within 1 km in 2010. All specifications include province fixed effects, county-level controls, and plot-level geographic controls. Robust standard errors, clustered at the county level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Robustness: Removing nodal locations for forest outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	Δ ln(Volume)					
Sample	Baseline		Drop < 20km		Drop < 46km	
IV Type	Continuous	Dummy	Continuous	Dummy	Continuous	Dummy
$Dist_{icp}^{HW}$	-0.021***		-0.021***		-0.021**	
	(0.008)		(0.008)		(0.008)	
$\mathbf{I}(Dist_{icp}^{HW} < 20$ km)		0.150*		0.175*		0.193*
		(0.086)		(0.092)		(0.109)
Observations	16226	16226	15910	15910	14120	14120
Mean of depvar	0.37	0.37	0.37	0.37	0.36	0.36
F-statistic in 1st stage	45.5	49.5	45.8	44.2	40.9	30.6
Province fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
County-level controls	Yes	Yes	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: Results are based on the specification in Table 1. All estimates use instrumental variables derived from the LCP and the 1962 Road Network. Columns (1)–(2) use the baseline sample. The sample progressively excludes plots located within 20 km (Columns 3–4) and 46 km (Columns 5–6) of district seats and county seats. Robust standard errors are clustered at the county level and reported in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table A.7: Robustness: Removing nodal locations for firm entry

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)
	log(Num. Firms Registered)					
Sample	Baseline		Drop < 20km		Drop < 46km	
Method	OLS	IV	OLS	IV	OLS	IV
$\mathbf{I}(Dist_{cp}^{HW} < 20\text{km})$	0.097*	0.586***	0.095*	0.606***	0.065	0.486**
	(0.052)	(0.196)	(0.053)	(0.207)	(0.056)	(0.214)
Observations	817	817	796	796	671	671
F-statistic in 1st stage		27.757		24.744		22.331
Province Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Control	Yes	Yes	Yes	Yes	Yes	Yes
Mean of Dep. Var.	70.25	70.25	69.92	69.92	63.97	63.97

Note: This table presents robustness checks for the baseline firm entry results reported in Table 3. Columns (1)–(2) replicate the baseline OLS and IV estimates. Columns (3)–(6) progressively exclude counties located near nodal centers: Columns (3)–(4) exclude counties within 20 km of nodal cities, and Columns (5)–(6) exclude counties within 46 km. The dependent variable is the logarithm of the number of newly registered timber-related firms. Highway connection is instrumented using the planned least-cost path and 1962 road. Robust standard errors are reported in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table A.8: Robustness: Controlling for economic and population growth

Dep. Var	(1)	(2)	(3)	(4)
	$\Delta \ln$(Volume)			
$Dist_{icp}^{HW}$	-0.021***	-0.022***	-0.021***	-0.022***
	(0.008)	(0.008)	(0.008)	(0.008)
$\Delta \ln$(Pop)		-0.043		-0.041
		(0.043)		(0.043)
$\Delta \ln$(NTL)			-0.029	-0.026
			(0.044)	(0.044)
Observations	16226	16226	16226	16226
Mean of depvar	0.37	0.37	0.37	0.37
F-statistic in 1st stage	45.48	43.91	45.61	43.97
Province fixed effects	Yes	Yes	Yes	Yes
County-level controls	Yes	Yes	Yes	Yes
Weather controls	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes

Note: This table checks the robustness of the baseline IV estimates (Column 1) by adding controls for county-level population growth (Column 2), nighttime light growth (Column 3), and both simultaneously (Column 4). All specifications include the full set of baseline controls. Population growth is measured as the log difference in county population between 2000 and 2010. Nighttime light growth is measured using NOAA DMSP-OLS stable lights data (F15 in 2000 and F18 in 2010), inter-calibrated following Hsu et al. (2015) to account for sensor differences. Robust standard errors are clustered at the county level and reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.9: Robustness: Varying trade elasticity parameters

Dep. Var.	(1)	(2) $\Delta \ln$(Volume)	(3)
Trade elasticity θ	1	1.5 (Herzog21)	3.6 (DNS16)
RE_c^{Specific}	1.932*	0.691**	0.053*
	(1.166)	(0.279)	(0.030)
RE_c^{Other}	-2.097	-0.802**	-0.056
	(1.276)	(0.371)	(0.049)
IE_c^{Specific}	-0.170	-0.090	-0.016
	(0.441)	(0.137)	(0.012)
IE_c^{Other}	3.423***	0.946***	0.080***
	(0.833)	(0.245)	(0.024)
Observations	16226	16226	16226
Mean of Income Effect	1.506	1.491	1.367
Mean of Road Effect	0.181	0.249	0.296
Province Fixed Effect	Yes	Yes	Yes
County Controls	Yes	Yes	Yes
Raster Controls	Yes	Yes	Yes

Note: Results are based on the specification in Equation (5). This table presents a sensitivity analysis of the market access mechanism by varying the trade elasticity parameter θ. Columns (1), (2), and (3) report results using $\theta = 1$, $\theta = 1.5$ (the baseline, following Herzog (2021)), and $\theta = 3.6$ (following Donaldson and Hornbeck (2016)), respectively. Controls are as specified in Table 2. Robust standard errors are clustered at the county level and reported in parentheses. Significance levels: * p<0.1, ** p<0.05, *** p<0.01.

Table A.10: Robustness: Alternative sample definitions

Dep. var.	(1)	(2)	(3) $\Delta \ln$(Volume)	(4)	(5)
Sample	Baseline	Incl. Pre-1999	Excl. 50km Pre-1999	Incl. 2010 Roadside (<1km)	Incl. Pre-1999
$Dist_{icp}^{HW}$	-0.021***	-0.024***	-0.017**	-0.021***	
	(0.008)	(0.008)	(0.007)	(0.008)	
$Dist_{icp}^{HW}$ (2010)					-0.022***
					(0.008)
Observations	16226	16726	15092	16401	16726
Mean of dep. var.	0.37	0.37	0.36	0.37	0.37
F-statistic in 1st stage	45.5	43.3	42.8	46.0	45.3
Province fixed effects	Yes	Yes	Yes	Yes	Yes
County-level controls	Yes	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes	Yes

Note: This table assesses robustness by varying sample selection and distance definitions using the joint IV estimation. In Columns (1)–(4), highway proximity is measured using only highways constructed between 2000 and 2010, whereas Column (5) uses the full 2010 highway network (including both pre-2000 and post-2000 segments). Within this framework: Column (1) is the baseline, which excludes plots within 20 km of pre-2000 highways. Subsequent columns modify this baseline as follows: Column (2) includes the plots within 20 km of pre-2000 highways; Column (3) replaces the 20 km exclusion with a stricter 50 km exclusion; and Column (4) includes roadside plots within 1 km of 2010 highways, which are omitted in the baseline. Finally, Column (5) includes the pre-2000 plots (as in Column 2) but applies the expanded network definition. Controls are as specified in previous tables. Robust standard errors clustered at the county level are reported in parentheses. Significance levels: * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

A.3 Methodology for Shortest Path Calculation and Travel Time Construction

To measure spatial connectivity across China, we implement a least-cost path (LCP) approach using a high-resolution raster representation of the national transportation network in 2000 and 2010. The methodology proceeds in three steps: construction of pixel-level speed surfaces, graph-based computation of least-cost paths, and aggregation to county-level travel times.

We begin by constructing national speed surfaces at a 1 km resolution. Following the main text, travel speeds are assigned based on road type and terrain. For highways and first-rate roads, we directly use pixel-level design speed information from the Transportation Networks Dataset (Ma and Tang, 2024). For other national and provincial roads, speeds are assigned as a function of slope: 60 km/h for slopes below 10° and 40 km/h otherwise.[40] To capture connectivity across locations not directly linked by observed roads, we impose a minimum travel-speed threshold of 40 km/h for all traversable pixels. Finally, to integrate administrative centers into the network, we connect each county centroid to the nearest road segment using a nearest-road search combined with the Bresenham line algorithm.

The resulting speed raster is then converted into a directed, 8-directional weighted graph, where edge weights correspond to travel times between adjacent pixels. Because computing least-cost paths between all county pairs is computationally costly given more than 2,800 counties, we adopt a two-stage procedure. In the first stage, we compute pixel-level least-cost paths only between neighboring county pairs.[41] In the second stage, these local connections are aggregated into a county-level graph that summarizes feasible inter-county travel links.

A key feature of our approach is the explicit modeling of the transportation system as a two-layer network, which captures the costs of accessing high-speed corridors. The traditional road network is treated as an open-access layer, while the highway network constitutes a closed-access layer. Transfers between layers are modeled by a transit cost, defined as the travel time from a county centroid to the nearest highway interchange via local roads. This structure allows counties without direct highway access to reach the high-speed network by first traveling to neighboring counties with access, while the ubiquitous traditional road layer guarantees nationwide connectivity. Final inter-county travel times are obtained by applying

[40] These assumptions follow the Technical Standard of Highway Engineering (Ministry of Transport of the People's Republic of China, 2004). For second-rate arterial roads, the preferred design speed is 80 km/h, but 60 km/h is common where traffic is mixed or intersections are dense. In mountainous terrain, 40 km/h is standard. For roads below second-rate, the maximum design speed is 40 km/h.

[41] Two counties are defined as neighbors if they share a common administrative boundary or if the Euclidean distance between their centroids is less than twice the sum of their equivalent radii, calculated as $\sqrt{\text{Area}/\pi}$.

Dijkstra's algorithm to this integrated, multi-layer network.

A.4 Least-Cost Path IV Construction

This section describes the construction of least-cost paths between target cities specified in China's National Expressway Network Plan. The plan identifies 245 major cities to be connected by the national expressway system, yielding 384 pairs of adjacent target cities. Following Faber (2014) and He, Xie and Zhang (2020), we construct a least-cost path for each city pair to proxy the planned expressway layout.

We begin by constructing a rasterized cost surface, where the traversal cost c_i for each land pixel i is defined as

$$c_i = 1 + \text{slope}_i + 25 \cdot \text{Developed}_i + 25 \cdot \text{Water}_i + 25 \cdot \text{Wetland}_i. \tag{A.1}$$

This cost function favors short and flat routes, while imposing large penalties on paths that cross developed areas, water bodies, or wetlands. Land cover information is obtained from GLCNMO (Tateishi et al., 2011), and slope is computed from SRTM v4.1 elevation data (Amatulli et al., 2018).

Using this cost raster, we compute least-cost paths for all 384 city pairs by applying Dijkstra's algorithm. The cost of moving between adjacent grid cells is defined as $c_{od1} = (c_o + c_{d1})/2$ for horizontal or vertical movements, and $c_{od2} = \sqrt{2}(c_o + c_{d2})/2$ for diagonal movements, reflecting differences in travel distance across grid directions.[42]

The resulting least-cost path network closely mirrors the spatial structure of the realized expressway system. For reference, Figure A.3 compares the constructed least-cost paths with the observed 2010 highway network.

A.5 Effects by distance to highways

The positive effects on forest quality are concentrated within 20 km of newly constructed highways. To characterize how impacts vary spatially, we estimate a banded OLS specification. Specifically, we partition distance to the nearest new highway into nine intervals based on deciles of the distance distribution: 1–11 km, 11–20 km, 20–28 km, 28–37 km, 37–46 km, 46–57 km, 57–71 km, 71–92 km, 92–134 km and beyond 134 km. We use the 92–134 km interval as the omitted category, since plots located very far from highways are more likely to lie in remote or environmentally distinct areas—such as deserts or high-altitude regions—that

[42] We deliberately avoid using Kruskal's minimum spanning tree algorithm, which tends to generate overly sparse and stylized networks that poorly approximate actual highway layouts.

may exhibit systematically different forest growth trends.

The estimating equation is:

$$\Delta y_{icp} = \alpha + \sum_{m} \beta_m \, \mathbf{1}\Big[Dist_{icp}^{HW} \in I_m\Big] + \mathbf{x}_{c0}'\boldsymbol{\gamma} + \mathbf{m}_i'\boldsymbol{\eta} + \Delta\mathbf{z}_c'\boldsymbol{\delta} + \lambda_p + \varepsilon_{icp}, \tag{A.2}$$

where $\mathbf{1}[Dist_{icp}^{HW} \in I_m]$ indicates whether plot i lies within distance band I_m, defined as concentric rings around highways opened between 2000 and 2010. All controls and fixed effects are defined as in the baseline specification.

Table A.11: Binned Regressions

Dep. var.	(1) Δ ln(Volume)	(2) Δ Coverage	(3) Δ ln(DBH)	(4) Δ ln(Avg. Height)
$\mathbf{I}(Dist_{icp}^{HW} < 11\text{km})$	0.035	1.993*	0.017	0.050*
	(0.038)	(1.123)	(0.022)	(0.026)
$\mathbf{I}(11 \leq Dist_{icp}^{HW} < 20\text{km})$	0.083**	3.933***	0.037*	0.059**
	(0.037)	(1.074)	(0.022)	(0.026)
$\mathbf{I}(20 \leq Dist_{icp}^{HW} < 28\text{km})$	0.005	1.767*	0.010	0.010
	(0.035)	(0.995)	(0.022)	(0.025)
$\mathbf{I}(28 \leq Dist_{icp}^{HW} < 37\text{km})$	0.009	1.718*	0.001	0.018
	(0.033)	(0.924)	(0.020)	(0.023)
$\mathbf{I}(37 \leq Dist_{icp}^{HW} < 46\text{km})$	0.014	0.675	0.020	0.028
	(0.035)	(1.015)	(0.020)	(0.023)
$\mathbf{I}(46 \leq Dist_{icp}^{HW} < 57\text{km})$	0.028	1.682*	0.028	0.036
	(0.033)	(0.926)	(0.020)	(0.024)
$\mathbf{I}(57 \leq Dist_{icp}^{HW} < 71\text{km})$	0.028	0.980	0.012	0.018
	(0.032)	(0.892)	(0.017)	(0.022)
$\mathbf{I}(71 \leq Dist_{icp}^{HW} < 92\text{km})$	0.024	1.215	0.012	-0.001
	(0.030)	(0.840)	(0.015)	(0.017)
$\mathbf{I}(Dist_{icp}^{HW} > 134\text{km})$	-0.026	-0.161	-0.001	0.038
	(0.037)	(0.983)	(0.021)	(0.026)
Observations	16226	17350	16602	16896
Mean of dep. var.	0.37	6.01	-0.28	0.17
Province fixed effects	Yes	Yes	Yes	Yes
County-level controls	Yes	Yes	Yes	Yes
Point-level controls	Yes	Yes	Yes	Yes

Note: Outcomes are differences between two waves of surveys: Δ ln(Volume), Δ Coverage (percentage points), Δ ln(DBH), and Δ ln(Avg. Height). Estimates are based on banded proximity indicators, with distance from 92 km to 134 km (80-90th percentiles) as the reference category. The sample excludes plots within 20 km of any highway in 2000 and within 1 km in 2010. All specifications include province fixed effects, county-level controls, and plot-level geographic controls. Robust standard errors, clustered at the county level, are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A.11 reports the banded OLS estimates. Consistent with the baseline results, the effects on forest quality are concentrated close to highways. While statistically significant increases in forest volume and DBH appear primarily in the 11–20 km band, positive effects begin to emerge even closer to the highway. In particular, plots within 11 km already exhibit significant gains in canopy coverage and average height. These effects intensify and become significant across all four forest metrics in the 11–20 km band, which represents the zone of peak impact. Guided by this spatial pattern, we define the area within 1–20 km of newly constructed highways as the primary effect zone and use it in the subsequent mechanism and

heterogeneity analyses.

A.6 Extensive Margin and Forest Planting Response

This section examines whether highway expansion affects forests along the extensive margin, focusing on changes in forest area and afforestation activity at the county level. We first describe the empirical strategy used to identify the effect of highway expansion on forest area and planting, and then present the corresponding results.

Methods We begin by estimating the impact of highway expansion on changes in county-level forest area between 2000 and 2010. The baseline specification relates the log change in forest area to highway proximity, measured by the median distance from forest plots within a county to the nearest highway:

$$\Delta \ln(\text{Forest Area}_c) = \alpha + \beta Dist_c^{HW} + \mathbf{x}_{c0}'\boldsymbol{\gamma} + \lambda_p + \varepsilon_c, \tag{A.3}$$

where $Dist_c^{HW}$ captures highway exposure and $\mathbf{x}_{c0}$ denotes county controls in the initial period. To address potential endogeneity in highway placement, we instrument $Dist_c^{HW}$ with the median distance to the least-cost path (LCP) network constructed from pre-determined geographic characteristics.

We also estimate an alternative specification that uses a binary measure of highway access following the specification below:

$$\Delta \ln(\text{Forest Area}_c) = \alpha + \beta \cdot \mathbf{I}(Dist_c^{HW} < 20 \text{ km}) + \mathbf{x}_{c0}'\boldsymbol{\gamma} + \lambda_p + \varepsilon_c. \tag{A.4}$$

where the indicator $\mathbf{I}(Dist_c^{HW} < 20 \text{ km})$ equals one if the median forest-plot distance to the nearest highway is below 20 km. This indicator is instrumented using a dummy variable indicating distance to the LCP network. In both specifications, we exclude counties that were already connected to highways in 2000 (median distance below 20 km) to isolate the effects of network expansion rather than pre-existing access.

To study afforestation, we exploit annual county-level data on forest planting areas from the *China Forestry Statistical Yearbook* (2002–2010). Using the panel structure of these data, we estimate a staggered difference-in-differences model:

$$\ln y_{ct} = \alpha + \beta \mathbf{I}(Dist_{ct}^{HW} < 20 \text{ km}) + \mathbf{x}_{c0}'\boldsymbol{\gamma} + \lambda_c + \lambda_t + \varepsilon_{ct}, \tag{A.5}$$

where $\ln y_{ct}$ denotes the logarithm of forest planting area in county c and year t. Counties are classified as treated once their median forest-plot distance to the nearest highway falls

below 20 km. We exclude counties connected prior to 2002 and include county and year fixed effects, along with interactions between baseline controls and year dummies, consistent with the baseline specification in equation A.2.

Table A.12: Extensive Margin Changes in Forest Share at County Level

	(1)	(2)	(3)	(4)	(5)	(6)
Dep. Var.	Δ Forest Share		$Dist_{cp}^{HW}$	Δ Forest Share		$\mathbf{I}(Dist < 20)$
Variable	Distance (Continuous)			Connected (Binary)		
Method	OLS	IV	1st Stage	OLS	IV	1st Stage
$Dist_{cp}^{HW}$	-0.010	-0.030***				
	(0.006)	(0.011)				
$\mathbf{I}(Dist_{cp}^{HW} < 20\text{km})$				0.007	0.044	
				(0.031)	(0.117)	
LCP IV			0.364***			
			(0.048)			
1962 Road			0.201***			
			(0.061)			
LCP IV (< 20km)						0.268***
						(0.035)
1962 Road (< 20km)						0.080**
						(0.033)
Mean. Dep. Var.	0.058	0.058	3.971	0.058	0.058	0.317
Mean Forest Share 2000	0.458	0.458	0.458	0.458	0.458	0.458
Observations	810	810	810	810	810	810
F-statistic in 1st stage		48.724			36.505	
Province Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes
Control	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports county-level estimates for changes in forest share. The explanatory variable is the distance to the nearest highway, defined as the median of the shortest distances from each forest plot within the county to the network (realized highways or IVs). Columns (1)–(3) use the continuous distance, while Columns (4)–(6) use a binary indicator for distance < 20 km. Columns (1) and (4) report OLS estimates, (2) and (5) report IV estimates, and (3) and (6) report the first stage. Instruments are based on the LCP and 1962 Road Network. All specifications control for province fixed effects and baseline county characteristics measured in 2000, as in Equation 1. The sample excludes counties connected in 2000. Standard errors are reported in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Results Table A.12 reports the estimates for changes in county-level forest area. Across both continuous and binary specifications, we find no evidence that highway expansion leads to forest loss. If anything, proximity to highways is associated with weakly positive changes in forest area, suggesting that improved connectivity does not induce extensive-margin deforestation once mechanical land conversion from highway construction is excluded.

Turning to afforestation activity, Table A.13 presents estimates from a balanced panel of 340 counties with continuous planting records. Columns (1) and (2) show that highway connection increases total forest planting area by approximately 25% relative to unconnected counties. Column (3) focuses on artificial planting, which captures more intensive and ac-

tively managed afforestation.[43] The estimated effect is larger, at approximately 35.3%, and statistically significant.

To interpret the magnitude, note that the average annual artificial planting area in these counties amounts to roughly 2.53% of their total forest area in 2000.[44] The estimated highway effect therefore corresponds to an annual increase of about 0.63% of the existing forest stock. Accumulated over a decade, this implies an expansion of forest area of approximately 6% attributable to improved highway access.

Finally, Figure A.4 examines the dynamic response of forest planting using an event-study design. We find no evidence of differential pre-trends prior to highway connection, supporting the validity of the identification strategy. Following connection, forest planting increases sharply in the first year and remains persistently higher thereafter, indicating a sustained afforestation response rather than a short-lived adjustment.

Table A.13: Forest Planting at County-Level from 2002 to 2010

Dep. Var.	(1)	(2) ln(Planting Area)	(3)
Type	All		Artificial
$\mathbf{I}(Dist_{ct}^{HW} < 20\text{km})$	0.256**	0.240**	0.353***
	(0.122)	(0.120)	(0.135)
Observations	3060	3060	3040
Controls	No	Yes	Yes
County FE	Yes	Yes	Yes
Year FE	Yes	Yes	Yes
Mean of Dep. Var.	7.44	7.44	7.19

Note: Estimates are based on the specification in Equation (A.5). The dependent variable is the logarithm of forest planting area. Columns (1)–(2) use total planting area, while Column (3) uses artificial planting area, which refers to forests established through direct human planting (excluding aerial seeding and natural regeneration). All columns use difference-in-differences (DID) with annual panel data 2002–2010 from the balanced sample. All specifications exclude counties connected before 2002 and include county and year fixed effects, and county-level controls. Robust standard errors, clustered at the county level, are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

[43] Artificial planting refers to forests established through direct human planting, excluding aerial seeding and natural regeneration.

[44] This ratio is calculated as the average annual artificial planting area (11,141 km^2) over 2002–2010 divided by total forest area in 2000 (439,902 km^2), based on GlobeLand30 data.

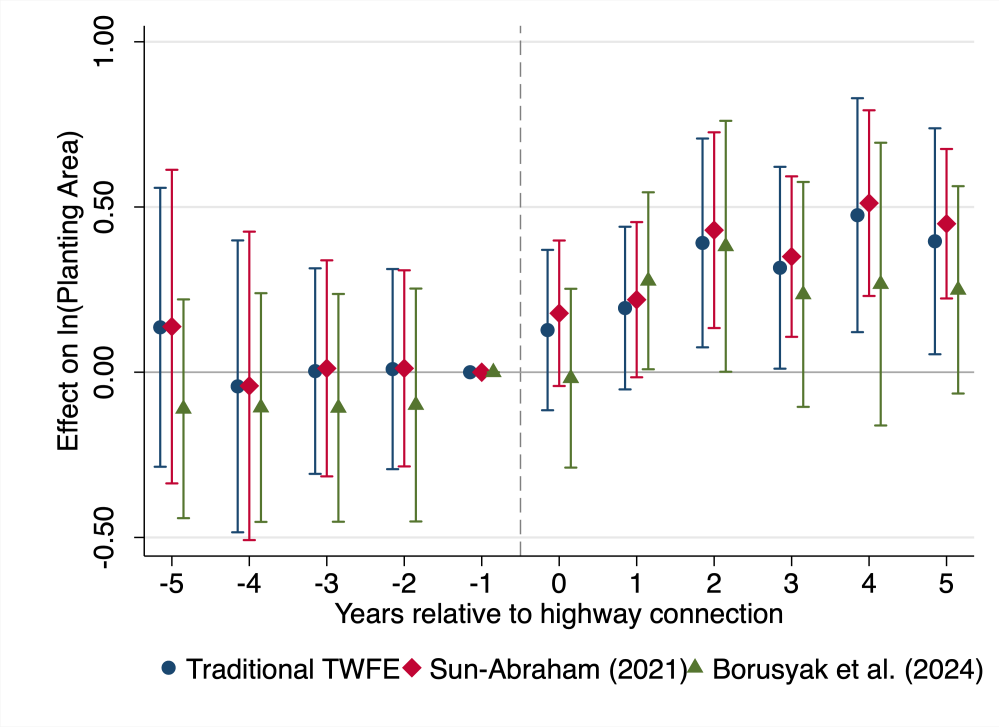


Figure A.4: Event Study (Balanced Panel)

Note: The dependent variable is log of forest planting area. All specifications exclude counties connected before 2002. The figure shows event study estimates using three methods: traditional TWFE, Sun and Abraham (2021), and Borusyak, Jaravel and Spiess (2024). For TWFE and Sun-Abraham, we bin periods ≤ -6 and ≥ 6 to ensure single-period effects in the displayed range [-5, 5]. For Borusyak, Jaravel and Spiess (2024), we restrict the estimation window to [-5, 5]. The reference period is $t = -1$. Error bars represent 95% confidence intervals with standard errors clustered at the county level.

A.7 Back-of-the-Envelope Calculations

Our baseline estimates imply that highway expansion increases forest carbon storage by raising standing timber volume. To assess the aggregate carbon consequences of these effects, we conduct a back-of-the-envelope calculation that translates the estimated intensive-margin response into national-scale carbon gains, while explicitly accounting for the offsetting carbon losses from the physical footprint of highway construction.

We first scale the intensive-margin effect implied by our estimates. Applying the updated 15% increase in standing timber volume to the share of China's forest area that experienced improved market access—23.43% of total forest land—yields a gross increase of approximately 435.8 million m^3 of standing timber. This calculation is anchored to the national forest stock of 1.24×10^{10} m^3 reported in the 6th National Forest Inventory (1999–2003) and captures the primary channel through which highways affect forest carbon stocks.

We then incorporate the direct carbon loss associated with highway rights-of-way as a corrective term. Between 2000 and 2010, China added roughly 53,400 km of expressways.

Assuming a conservative rights-of-way width of 60 meters and combining this footprint with the average forest share within a 1 km highway buffer (20.77%), we estimate that approximately 665.5 km^2 of forest land was directly converted. Valued at the national average standing timber density of 8,473 m^3/km^2, this corresponds to a gross loss of about 5.64 million m^3 of standing timber.

Subtracting this direct land-conversion loss from the intensive-margin gains yields a *net* increase in national standing timber volume of approximately 430.2 million m^3. Expressed in carbon terms, this calculation indicates that improvements in forest quality induced by highway expansion dominate the carbon losses from rights-of-way, even under conservative assumptions regarding land conversion.

As summarized in Table A.14, the implied net carbon sequestration ranges from 216.9 to 551.9 Mt CO_2. These estimates are obtained by converting net standing timber volume gains into carbon using a standard accounting framework: carbon gain equals net volume multiplied by the Biomass Conversion and Expansion Factor (BCEF), a carbon fraction of 0.5, the molecular weight ratio of CO_2 to carbon (44/12), and a permanence adjustment of 0.5. The BCEF converts standing timber volume into total aboveground biomass; we consider values of 0.55 and 1.40, corresponding to the lower and upper bounds of the typical range for forests in China. The resulting carbon gains are comparable in magnitude to the annual fossil-fuel emissions of countries such as Spain (lower bound) or Canada (upper bound). Valuing these carbon benefits using a Social Cost of Carbon (SCC) range of $51 per ton (Interagency Working Group on Social Cost of Greenhouse Gases, 2016) to $190 per ton (U.S. Environmental Protection Agency, 2022), the implied welfare gains span $11.1–41.2 billion under the conservative parameterization and increase to $20.1–74.9 billion under the baseline scenario.

Table A.14: Sensitivity Analysis of Net Forest Carbon Gains

Parameter Scenario	BCEF	Net Volume Gain (Million m^3)	Net Carbon Gain (Mt CO_2)	Social Value (Billion USD)
Conservative	0.55	430.2	216.9	11.1 – 41.2
Baseline	1.00	430.2	394.4	20.1 – 74.9
Optimistic	1.40	430.2	551.9	28.1 – 104.9

Note: Carbon gain = Net Volume × BCEF × Carbon Fraction (0.5) × (44/12) × Permanence (0.5). BCEF (Biomass Conversion and Expansion Factor) converts standing timber volume to total biomass; values of 0.55 and 1.40 correspond to the lower and upper bounds of the typical range for forests in China. Social Value is calculated using a Social Cost of Carbon (SCC) range of $51/ton (Interagency Working Group on Social Cost of Greenhouse Gases, 2016) to $190/ton (U.S. Environmental Protection Agency, 2022).

A.8 Alternative Channels

This section tests whether the observed forest recovery is driven by alternative channels unrelated to market access, specifically passive vegetation regrowth due to labor out-migration and associated changes in cultivated land.

Labor out-migration To explicitly test the labor drain mechanism, we estimate the elasticity of forest volume to local employment changes:

$$\Delta y_{icp} = \beta \Delta \ln(\text{Labor}_c) + \mathbf{x}'_{c0}\boldsymbol{\gamma} + \mathbf{m}'_i\boldsymbol{\eta} + \Delta\mathbf{z}'_c\boldsymbol{\delta} + \lambda_p + \varepsilon_{icp} \tag{A.6}$$

where Δy_{icp} denotes the change in forest volume at the plot level, and $\Delta \ln(\text{Labor}_c)$ represents the change in county-level agricultural employment or prefecture-level forestry employment. The vectors $\mathbf{x}_{c0}, \mathbf{m}_i, \Delta\mathbf{z}_c$ represent the standard set of controls defined in equation 1.

Table A.15 reports the OLS estimates of this labor-volume elasticity. We find no statistically significant relationship between changes in local labor supply (either agricultural or forestry) and forest volume growth. This null result, derived from plots located more than 1 km from highways to exclude direct conversion effects, supports our interpretation that forest recovery is an active response to market incentives rather than a passive byproduct of rural depopulation.

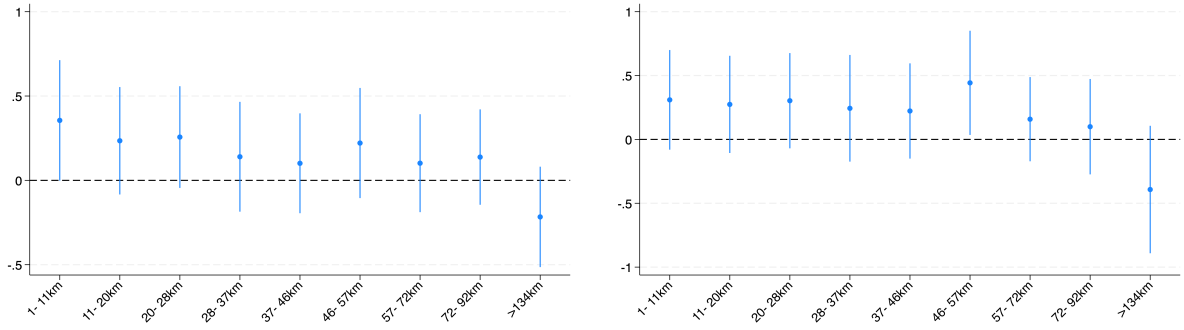
Table A.15: Labor Growth and Forest Volume Change, 2000–2010

Dep. Var.	(1)	(2)	(3)	(4)	(5)	(6)
	$\Delta \ln$(Volume)					
	County-level Agr. Emp.			Prefecture-level Forestry Emp.		
$\Delta \ln$(Agr. Emp.)	-0.021	0.008	-0.002			
	(0.020)	(0.023)	(0.022)			
$\Delta \ln$(Forestry Emp.)				-0.010	0.002	0.005
				(0.029)	(0.025)	(0.024)
Observations	16115	16115	16115	14730	14730	14730
Mean of Dep. Var.	0.37	0.37	0.37	0.37	0.37	0.37
Province FE	Yes	Yes	Yes	Yes	Yes	Yes
County controls	No	Yes	Yes	No	Yes	Yes
Point controls	No	No	Yes	No	No	Yes

Note: This table examines the relationship between labor migration and forest volume change based on the specification in Equation (A.6). Columns (1)–(3) use the change in county-level agricultural employment as the key explanatory variable, while Columns (4)–(6) use the change in prefecture-level forestry employment. Controls are as specified in Table 1. Robust standard errors clustered at the county level are reported in parentheses. * p<0.1, ** p<0.05, *** p<0.01.

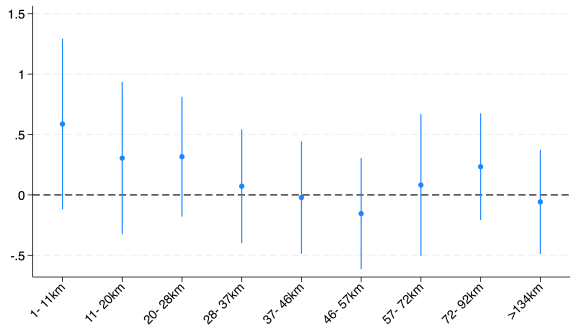
Changes in cultivated land We examine cultivated land trends around forest plots to test the possibility of passive land cover change. If the observed forest recovery is merely a

byproduct of labor outflow and cropland contraction, we should observe a concurrent decline in nearby cropland. Figure A.5 presents the results for this test, using the same method as equation A.2. We find no significant reduction in cultivated land share within 1500 m of forest plots near highways. The coefficients are small the pattern holds across both the full sample and subsamples split by slope. This suggests that mechanical cropland contraction is not a primary driver of the observed forest dynamics.



(a) Full Sample

(b) Slope $> 15°$



(c) Slope $< 15°$

Figure A.5: Change in Cultivated Land by Distance to Highway

Note: The dependent variable in this regression is the percentage change in cultivated land share within a 1500 m radius surrounding each forest plot. The regression specification is otherwise identical to the one presented in Equation (A.2).

B Model Appendix

B.1 Forest-Sector Prices, Trade Shares, and Effective Forest Productivity

This appendix provides additional details on the derivation of the cost structure, pricing rule, and trade equilibrium conditions for the forest sector.

Input cost and pricing under monopolistic competition The unit cost of producing one unit of a forest variety in region i depends on land and urban intermediate inputs. With cost shares υ and $1-\upsilon$, respectively, the unit cost of the input bundle is

$$\chi_i = (r_i)^{\upsilon}(P_i^u)^{1-\upsilon}, \tag{B1}$$

where r_i denotes the land rental rate and P_i^u is the local price index of urban intermediates.

Under monopolistic competition with CES demand, each farmer charges a constant markup over marginal cost. Specifically, with elasticity of substitution ϵ, the markup factor is $\epsilon/(\epsilon-1)$. The producer price of a forest variety ω^f from a type-z household in region i is therefore

$$p_{zi}^f = \frac{\epsilon}{\epsilon-1} \cdot \frac{\chi_i}{\bar{A}_i \psi_z}, \tag{B2}$$

where $\bar{A}_i$ is location-specific productivity and ψ_z represents the household's forestry skills.

Forest price index and trade costs Consumers in region n face CES preferences over all forest varieties sourced from all regions. The corresponding CES price index is

$$P_n^f = \left[\sum_i \sum_z \left(\frac{\epsilon}{\epsilon-1} \frac{\chi_i}{\bar{A}_i \psi_z} \tau_{ni}^f \right)^{1-\epsilon} L_{zi,r} \right]^{\frac{1}{1-\epsilon}}, \tag{B3}$$

where $\tau_{ni}^f \geq 1$ denotes iceberg trade costs from i to n.

This expression shows that the effective price index depends not only on input costs and trade frictions but also on the distribution of forestry skills across households.

Aggregation over household types To simplify, we aggregate across households by defining effective forest productivity in region i:

$$\bar{\psi}_i = \left(\frac{L_{1i,r} + \psi^{\epsilon-1} L_{2i,r}}{L_{1i,r} + L_{2i,r}} \right)^{\frac{1}{\epsilon-1}}, \tag{B4}$$

where $L_{1i,r}$ and $L_{2i,r}$ denote the numbers of type-1 and type-2 households in the rural area of region i. This CES aggregator gives greater weight to the more productive type when $\psi > 1$, reflecting their disproportionate influence on average productivity. Substituting $\bar{\psi}_i$ into the price index yields a more tractable expression:

$$P_n^f = \left[\sum_i \left(\frac{\epsilon}{\epsilon - 1} \frac{\chi_i}{\bar{A}_i \bar{\psi}_i} \tau_{ni}^f \right)^{1-\epsilon} L_{i,r} \right]^{\frac{1}{1-\epsilon}}. \tag{B5}$$

Import probabilities (trade shares) Finally, the probability that consumers in region n source a forest good from region i—equivalently, the expenditure share of n on goods from i—is given by:

$$\pi_{ni}^f = \left(\frac{\epsilon}{\epsilon - 1} \frac{\chi_i}{\bar{A}_i \bar{\psi}_i} \tau_{ni}^f \right)^{1-\epsilon} L_{i,r} \cdot (P_n^f)^{\epsilon - 1}. \tag{B6}$$

This trade-share expression highlights how relative input costs, local productivity, farmer skill composition, and trade costs jointly determine sourcing patterns in the forest sector. In particular, regions with higher average forestry skills $\bar{\psi}_i$ or stronger location-specific productivity $\bar{A}_i$ are more competitive suppliers, while greater land rents or trade frictions raise delivered prices and reduce market shares.

B.2 Market Clearing and Equilibrium

This appendix provides a fuller discussion of the market-clearing conditions in both rural (forest) and urban (manufacturing) locations. The goal is to show how revenues, incomes, and expenditures balance across space, ensuring that all goods markets and factor markets clear in equilibrium.

Rural locations (forestry) Let $X_{n,a}^f$ denote total expenditure on forest goods in area a of region n. The value of sales from region i's forest producers to region n is

$$R_{ni}^f = \pi_{ni}^f (X_{n,r}^f + X_{n,u}^f), \tag{B7}$$

where π_{ni}^f is the sourcing probability defined earlier. Total revenue of the forest sector in region i is then

$$R_i^f = \sum_n \pi_{ni}^f X_{n,r}^f + \sum_n \pi_{ni}^f X_{n,u}^f, \tag{B8}$$

which aggregates demand from both rural and urban consumers across all destinations. Revenues are distributed among farmers, landowners, and intermediate suppliers. Farmers earn a share $1/\epsilon$ (profits), landowners receive $\upsilon(\epsilon - 1)/\epsilon$, and the remainder covers intermediate

inputs. Hence total rural income in region i is

$$Y_{i,r} = \left(\frac{1}{\epsilon} + \upsilon\frac{\epsilon-1}{\epsilon}\right) R_i^f. \tag{B9}$$

The rural land market clears when payments to land equal landowners' rental income:

$$r_i \bar{N}_i = \upsilon\frac{\epsilon-1}{\epsilon} R_i^f. \tag{B10}$$

Rural expenditure on urban goods has two parts: household consumption and intermediate input demand for forestry production:

$$X_{i,r}^u = (1-\eta)Y_{i,r} + (1-\upsilon)\frac{\epsilon-1}{\epsilon}R_i^f = \Xi R_i^f, \qquad \Xi \equiv \frac{1-\eta+(\epsilon-1)(1-\eta\upsilon)}{\epsilon}. \tag{B11}$$

Trade balance requires that rural expenditures on urban goods equal forest sales to urban destinations:

$$\Xi R_i^f = \sum_n \pi_{ni}^f X_{n,u}^f. \tag{B12}$$

Using equation B8 and noting that rural consumption of forest goods is $X_{i,r}^f = \eta Y_{i,r}$, we obtain:

$$(1-\Xi)R_i^f = \sum_n \pi_{ni}^f \eta \left(\frac{1}{\epsilon} + \upsilon\frac{\epsilon-1}{\epsilon}\right) R_n^f. \tag{B13}$$

Rearranging yields:

$$R_i^f = \sum_n \pi_{ni}^f R_n^f, \tag{B14}$$

which implicitly defines the vector of land rents r_i that clears rural forest markets.

Urban locations (manufacturing) Urban regions operate under perfect competition, so total urban income equals labor income:

$$Y_{i,u} = w_i L_{i,u}, \tag{B15}$$

where $L_{i,u}$ is the urban labor force. Let $X_{n,a}^u$ denote total expenditure on urban goods in area a of region n. Sales from urban region i to region n are:

$$R_{ni}^u = \pi_{ni}^u (X_{n,r}^u + X_{n,u}^u). \tag{B16}$$

Total urban revenue in region i is therefore

$$R_i^u = \sum_n \pi_{ni}^u X_{n,r}^u + \sum_n \pi_{ni}^u X_{n,u}^u. \tag{B17}$$

Trade balance requires that urban expenditures on forest goods equal urban sales to rural areas. Since urban demand for forest goods is $X_{i,u}^f = \eta Y_{i,u}$, this condition is:

$$\sum_n \pi_{ni}^u X_{n,r}^u = \eta Y_{i,u}. \tag{B18}$$

Using equation B17 and noting $X_{i,u}^u = (1-\eta)Y_{i,u}$, we obtain:

$$(1-\eta)Y_{i,u} = \sum_n \pi_{ni}^u (1-\eta) Y_{n,u}. \tag{B19}$$

Equivalently,

$$w_i L_{i,u} = \sum_n \pi_{ni}^u w_n L_{n,u}, \tag{B20}$$

which implicitly defines the wage vector w_i that clears the urban goods markets.

Together, equation B14 and equation B20 characterize general equilibrium in this economy. Rural revenues depend not only on local forestry productivity but also on the composition of farmers who remain in the sector, while urban wages are tied to labor inflows driven by migration decisions. Highways enter the system by reducing trade and migration costs, which alters sourcing probabilities π_{ni}^f and π_{ni}^u, thereby reshaping both rural and urban market equilibria.

B.3 Exact Hat Algebra and Relative Change in Equilibrium

To quantify the impact of highway expansion, we employ the exact hat algebra method to solve for relative changes in equilibrium. Let $\{w_i, r_i, P_i^f, P_i^u\}$ denote the initial equilibrium under policy variables $\{\tau_{ni}^f, \tau_{ni}^u, \mu_{in,a_0a}\}$, and let $\{w_i', r_i', P_i^{f'}, P_i^{u'}\}$ denote the equilibrium under the new policy. Define the hat variables $\{\hat{w}_i, \hat{r}_i, \hat{P}_i^f, \hat{P}_i^u\}$ to represent the relative changes in equilibrium outcomes, corresponding to relative changes in the policy variables $\{\hat{\tau}_{ni}^f, \hat{\tau}_{ni}^u, \hat{\mu}_{ni,a_0a}\}$. The equilibrium conditions in terms of relative (hat) variables are then given by:

$$\hat{\chi}_i = (\hat{r}_i)^{\upsilon}(\hat{P}_i^u)^{1-\upsilon} \tag{B21}$$

$$\hat{\bar{\psi}}_i = \left(\frac{L_{1i,r}\hat{L}_{1i,r} + \psi^{\epsilon-1}L_{2i,r}\hat{L}_{2i,r}}{L_{1i,r} + \psi^{\epsilon-1}L_{2i,r}}\right)^{1/(\epsilon-1)} \left(\hat{L}_{i,r}\right)^{1/(1-\epsilon)} \tag{B22}$$

$$\hat{P}_n^f = \left[\sum_i \pi_{ni}^f \left(\hat{\chi}_i \hat{\tau}_{ni}^f / \hat{\bar{\psi}}_i\right)^{1-\epsilon} \hat{L}_{i,r}\right]^{1/(1-\epsilon)} \tag{B23}$$

$$\hat{\pi}_{ni}^f = \left(\hat{\chi}_i \hat{\tau}_{ni}^f / \hat{\bar{\psi}}_i\right)^{1-\epsilon} \hat{L}_{i,r} \left(\hat{P}_n^f\right)^{\epsilon-1} \tag{B24}$$

$$\hat{P}_n^u = \left[\sum_i \pi_{ni}^u (\hat{w}_i \hat{\tau}_{ni}^u)^{-\theta} (\hat{L}_{i,u})^{\alpha\theta}\right]^{-1/\theta} \tag{B25}$$

$$\hat{\pi}_{ni}^u = (\hat{w}_i \hat{\tau}_{ni}^u)^{-\theta} (\hat{L}_{i,u})^{\alpha\theta} (\hat{P}_n^u)^{\theta} \tag{B26}$$

$$\hat{P}_n = (\hat{P}_n^f)^{\eta} (\hat{P}_n^u)^{1-\eta} \tag{B27}$$

$$\hat{\lambda}_{zin,a_0a} = \frac{\left(\hat{L}_{n,a}^{\beta} \hat{y}_{zn,a} \hat{P}_n^{-1} \hat{\mu}_{in,a_0a}^{-1}\right)^{\kappa}}{\sum_{n'} \sum_{a'} \lambda_{zin',a_0a'} \left(\hat{L}_{n',a'}^{\beta} \hat{y}_{zn',a'} \hat{P}_{n'}^{-1} \hat{\mu}_{in',a_0a'}^{-1}\right)^{\kappa}} \tag{B28}$$

$$\hat{y}_{zn,a} = \hat{r}_n \left(\frac{L_{1n,r}\hat{L}_{1n,r} + \psi^{\epsilon-1}L_{2n,r}\hat{L}_{2n,r}}{L_{1n,r} + \psi^{\epsilon-1}L_{2n,r}}\right)^{-1} \cdot \mathbb{I}\{a = r\} + \hat{w}_n \cdot \mathbb{I}\{a = u\} \tag{B29}$$

$$\hat{L}_{zn,a} = \sum_i \sum_{a_0} \frac{L_{zin,a_0a}}{L_{zn,a}} \hat{\lambda}_{zin,a_0a} \tag{B30}$$

$$\hat{L}_{n,a} = \frac{L_{1n,a}}{L_{n,a}} \hat{L}_{1n,a} + \frac{L_{2n,a}}{L_{n,a}} \hat{L}_{2n,a} \tag{B31}$$

$$R_i^f \hat{r}_i = \sum_n \pi_{ni}^f \hat{\pi}_{ni}^f R_n^f \hat{r}_n \tag{B32}$$

$$w_i \hat{w}_i L_{i,u} \hat{L}_{i,u} = \sum_n \pi_{ni}^u \hat{\pi}_{ni}^u w_n \hat{w}_n L_{n,u} \hat{L}_{n,u} \tag{B33}$$

$$\hat{\bar{U}}_{zi,a_0} = \left[\sum_{n'} \sum_{a'} \lambda_{zin',a_0a'} \left(\hat{L}_{n',a'}^{\beta} \hat{y}_{zn',a'} \hat{P}_{n'}^{-1} \hat{\mu}_{in',a_0a'}^{-1}\right)^{\kappa}\right]^{1/\kappa} \tag{B34}$$

B.4 Parameter Calibration

Migration flows by skill type We construct bilateral migration flows by skill type using the 2000 Population Census microdata, focusing on the working-age population (ages 15–64). Individuals are categorized into two skill groups (z): Type 1 (Urban CA) with high school education or above, and Type 2 (Forest CA) with middle school education or below. A key step in our procedure is mapping observed census characteristics to the model's spatial and sectoral definitions. The origin location i is defined by the individual's registered *hukou*, where we map agricultural and non-agricultural *hukou* to the rural ($a_0 = r$) and urban

($a_0 = u$) origin sectors, respectively. Since the 2000 Census does not provide city-level *hukou* registration, we proxy the specific *hukou* city using the individual's reported place of origin. The destination location n and sector a are identified by the current residence and industry, where agriculture-related industries (AFHF) are classified as the rural sector ($a = r$) and all others as urban ($a = u$).

A central identification challenge arises because the census data does not explicitly distinguish prior employment history for rural migrants; specifically, for individuals originating from rural areas ($a_0 = r$), we observe the aggregate flow but not the specific sub-sector (forestry versus general agriculture). To isolate forestry-specific labor flows, we apply a structural adjustment based on provincial labor composition. We calculate a forestry-to-agriculture labor ratio, ρ_{pz}, for each province p and skill type z:

$$\rho_{pz} = \frac{\sum_{i \in p} L_{iz}^{\text{Forestry}}}{\sum_{i \in p} L_{iz}^{\text{Agriculture}}} \tag{B35}$$

where L_{iz}^{Forestry} denotes workers strictly in the forestry industry (code 20) and $L_{iz}^{\text{Agriculture}}$ represents the broad agricultural sector. We then adjust the baseline rural-origin flows ($L_{zin,ra}^{\text{Base}}$) by this ratio to derive the estimated forestry migration flow: $L_{zin,ra}^{\text{Forestry}} = L_{zin,ra}^{\text{Base}} \times \rho_{pz}$.

Finally, to ensure our estimated flows are consistent with aggregate population statistics, we apply a macro-level scaling factor. We compute a scaling factor λ_n for each destination city n by comparing the total working-age population in official census tabulations (L_n^{Macro}) with the sample count in the microdata (L_n^{Micro}), such that $\lambda_n = L_n^{\text{Macro}} / L_n^{\text{Micro}}$. The final adjusted bilateral migration flow matrix, L_{zin,a_0a}, is obtained by:

$$L_{zin,a_0a} = L_{zin,a_0a}^{\text{Forestry}} \times \lambda_n \tag{B36}$$

This procedure yields a geographically consistent $2N \times 2N$ matrix of migration flows that accounts for both the unobserved origin sector of rural migrants and the sampling weights of the census microdata.

Trade flow of forestry. We construct bilateral trade flows for forestry products by integrating multiple datasets to overcome the absence of direct city-to-city forestry statistics. The core dataset is the 2007 inter-provincial input-output (IO) table compiled by Chen et al. (2023). Since the IO table aggregates forestry into the broader agricultural sector(AFHF), we isolate forestry-specific flows using a three-step procedure that leverages sector-specific input intensities. First, using the 2007 National IO Table, we compute the forestry intensity ζ_k for each downstream sector k, defined as the share of forestry intermediate inputs in total

agricultural intermediate inputs: ζ_k = Forestry Input$_k$/Agricultural Input$_k$. This captures the heterogeneity in forestry dependence across industries, such as the high intensity in wood processing versus low intensity in food manufacturing.

Next, we apply these intensity coefficients to the provincial agricultural flows to reconstruct the forestry-specific trade network. The forestry trade flow from province p to province q is estimated by summing the forestry content embedded in agricultural inputs supplied by p across all sectors k in q:

$$\text{Flow}_{pq}^{\text{Forestry}} = \sum_k \left(\text{Flow}_{pq,k}^{\text{Agri}} \times \zeta_k \right) \tag{B37}$$

where $\text{Flow}_{pq,k}^{\text{Agri}}$ represents the value of agricultural intermediate inputs from province p used by sector k in province q. This step effectively separates forestry trade from the aggregate agricultural flows based on the specific industrial composition of the destination province.

Finally, we disaggregate these provincial flows to city pairs (i, n) using supply and demand weights derived from local economic data. The bilateral forestry trade flow X_{ni}^f is calculated as:

$$X_{ni}^f = \text{Flow}_{pq}^{\text{Forestry}} \times \frac{S_i^{\text{supply}} \times D_n^{\text{demand}}}{\sum_{i' \in p} S_{i'}^{\text{supply}} \times \sum_{n' \in q} D_{n'}^{\text{demand}}} \tag{B38}$$

where the supply weight S_i^{supply} is proportional to the forest land area in origin city i (from National Forest Inventory data), and the demand weight D_n^{demand} is proportional to the derived demand for forestry inputs in destination city n. The latter is constructed using 2004 Economic Census microdata by summing the output of local firms weighted by their sector's forestry intensity ζ_k. This approach ensures that trade flows reflect both the comparative advantage of supply origins and the specific industrial demand of destinations.

Reduction in travel times. We begin by estimating bilateral city-level travel times for 2000 and 2010. The source data are inter-county travel times constructed from the Transportation Networks Dataset in Ma and Tang (2024), as described in Section 2.2.2. Since counties are the administrative subdivisions of cities in China, we then aggregate these county-to-county travel times to the city level for each year. Specifically, we calculate the population-weighted average travel time between all county pairs connecting two cities, where the weight for each pair is the product of their respective populations in 2000. Finally, we compute the relative change in city-to-city travel times as the ratio of the aggregated values in 2010 to those in 2000, which captures the relative change in travel time between cities resulting from transport improvements.

Reduction in migration costs and trade costs. We calibrate the migration costs by combining the auxiliary structural equation implied by the model with the calibrated bilateral migration flows, following Tombe and Zhu (2019) and Fan (2019). The migration cost to move from (i, a_0) to (n, a) is parameterized as

$$\ln \mu_{in,a_0a} = \beta_0 + \beta_1 \ln times_{in} + \beta_2 I_{in} \ln times_{in} + \beta_3 I_{in} + \nu_{a_0a}. \tag{B39}$$

The indicator I_{in} equals one if cities i and n are located in the same province. The variable $times_{in}$ measures geographic barriers, defined as the travel time between cities i and n. Finally, ν_{a_0a} denotes a migration type fixed effect, capturing whether the move is rural-rural, rural-urban, urban-rural, or urban-urban.

Second, we infer the migration cost from the migration decision (λ_{zin,a_0a}) as shown in equation 15. Dividing λ_{zin,a_0a} by λ_{zii,a_0a_0} and taking the logarithm of the ratio yields

$$\ln \frac{\lambda_{zin,a_0a}}{\lambda_{zii,a_0a_0}} = \kappa \times (\ln \nu_{zn,a} - \ln \nu_{zi,a_0} - \ln \mu_{in,a_0a} + \ln \mu_{ii,a_0a}) \tag{B40}$$

We then substitute equation 18 into equation B40, aggregate migration flows across worker types z to mitigate measurement error, and estimate the key parameters using the regression

$$\ln \frac{\lambda_{in,a_0a}}{\lambda_{ii,a_0a_0}} = \tilde{\beta}_0 + \tilde{\beta}_1 \ln times_{in} + \tilde{\beta}_2 I_{in} \ln times_{in} + \tilde{\beta}_3 I_{in} + \tilde{\nu}_{a_0,a} + \tilde{\nu}_{n,a} - \tilde{\nu}_{i,a_0} + \epsilon_{in,a_0a}. \tag{B41}$$

Here, we control for an extensive set of fixed effects. The term $\tilde{\nu}_{a_0a}$ denotes migration type fixed effects, which take four possible values depending on whether the origin-destination pair is rural-rural, rural-urban, urban-rural, or urban-urban. We also include fixed effects ($\tilde{\nu}_{i,a_0}$) and destination fixed effects ($\tilde{\nu}_{n,a}$). The structural parameters β_k can then be recovered from the estimated coefficients via the mapping $\tilde{\beta}_k = -\kappa\beta_k$. Migration costs are subsequently obtained from equation 18. In particular, we get the hat change of migration costs as

$$\ln \hat{\mu}_{in,a_0a} = \beta_1 \ln \widehat{times}_{in} + \beta_2 I_{in} \ln \widehat{times}_{in}. \tag{B42}$$

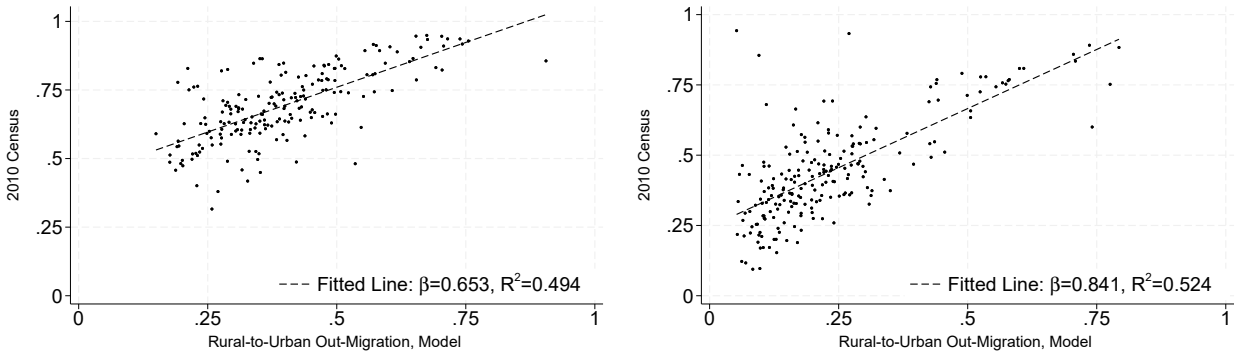
Since our bilateral trade data are only available at the province level, we do not directly estimate the trade gravity equation. Instead, we adopt the gravity estimates from Fan (2019). Specifically, we assume that for city pairs within the same province, each additional 1,000 km increases trade costs by 1 log point, whereas for city pairs across provinces, each additional 1,000 km increases trade costs by 21 log points. That is,

$$\ln \hat{\tau}_{in} = 0.021 \Delta \text{Distance}_{i,n} - 0.02 I_{in} \Delta \text{Distance}_{i,n}. \tag{B43}$$

Table B1: Summary of Key Parameter Calibration in the Model

Parameter	Description	Source	Value
L_{zin,a_0a}	Migration flow by types	Population Census	
X^f_{ni}, X^u_{ni}	Trade in forest and urban goods	Input–output table from Chen et al. (2023)	
$\widehat{times}_{in}$	Hat change in travel times	Transportation Networks Dataset from Ma and Tang (2024)	
$\hat{\mu}_{in,a_0a}$	Hat change in migration costs	Authors' own calculation	
$\hat{\tau}_{i,n}$	Hat change in trade costs	Authors' own calculation	
ψ	Comparative advantage parameter	Authors' own calculation	2.5
υ	Cost share of land in forest production	Rural Fixed Observation Point survey	0.68
η	Expenditure share on forest products	National input–output table, 2007	0.1
κ	Migration elasticity	Tombe and Zhu (2019)	1.5
ϵ	Elasticity of substitution	Hsieh and Klenow (2009)	3
θ	Trade elasticity	Simonovska and Waugh (2014)	4
α	Agglomeration elasticity	Allen and Arkolakis (2022)	0.1
β	Congestion elasticity	Allen and Arkolakis (2022)	-0.3

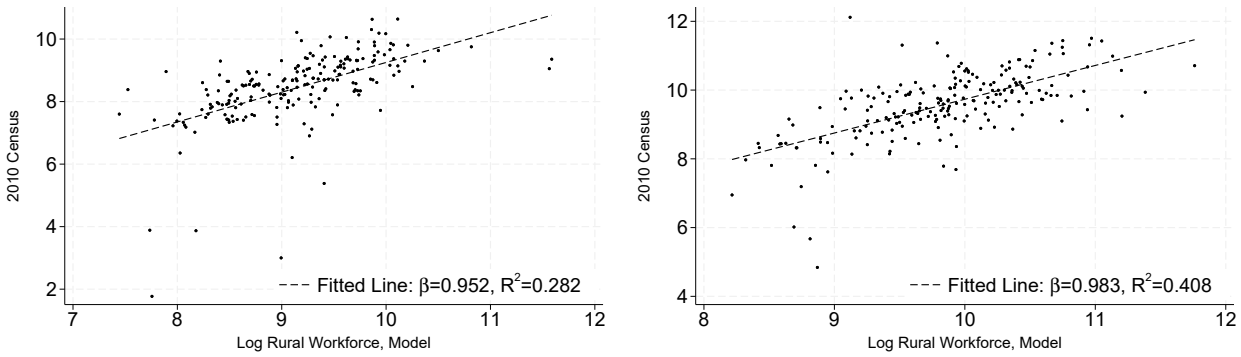
B.5 Model Fit



(a) Urban CA Individuals (b) Forest CA Individuals

Figure B1: Model-Predicted Rural-to-Urban Out-Migration

Notes: The figure compares the model-predicted out-migration probability—driven by highway expansion—with the probabilities observed in the 2010 Census, with each point representing a city. The probability measure is defined as the share of all migrants in a city who move to urban areas. Panel (a) reports results for Urban CA (type-1) individuals, while panel (b) reports results for Forest CA (type-2) individuals.



(a) Urban CA Individuals (b) Forest CA Individuals

Figure B2: Model-Predicted Rural Workforce in the Forest Sector

Notes: The figure compares the model-predicted rural workforce in each city with the rural workforce observed in the 2010 Census, measured in logarithmic form. Panel (a) reports results for Urban CA (type-1) individuals, while panel (b) reports results for Forest CA (type-2) individuals.

B.6 Effective Forest Productivity

Note that from effective forest productivity specified in equation 11, taking logs gives

$$\ln \bar{\psi}_i = \frac{1}{\epsilon - 1}\Big(\ln(L_{1i,r} + \psi^{\epsilon-1} L_{2i,r}) - \ln(L_{1i,r} + L_{2i,r})\Big) \tag{B44}$$

For simplicity, we suppress the subscripts i and r. Taking the total differential yields

$$d\ln \bar{\psi} = \frac{1}{\epsilon - 1}\left(\frac{d\tilde{N}}{\tilde{N}} - \frac{dN}{N}\right), \quad \tilde{N} = L_1 + \psi^{\epsilon-1} L_2,\ N = L_1 + L_2. \tag{B45}$$

Using $dL_j = L_j\, d\ln L_j$ $(j = 1, 2)$, this becomes

$$d\ln \bar{\psi} = \frac{1}{\epsilon - 1}\left[\left(\frac{L_1}{\tilde{N}} - \frac{L_1}{N}\right) d\ln L_1 + \left(\frac{\psi^{\epsilon-1} L_2}{\tilde{N}} - \frac{L_2}{N}\right) d\ln L_2\right]. \tag{B46}$$

Simplifying the coefficients of $d\ln L_1$ and $d\ln L_2$, we obtain

$$d\ln \bar{\psi} = \Theta\,(d\ln L_2 - d\ln L_1), \tag{B47}$$

where Θ is given by

$$\Theta \equiv \frac{(\psi^{\epsilon-1} - 1) L_1 L_2}{(\epsilon - 1)(L_1 + \psi^{\epsilon-1} L_2)(L_1 + L_2)}. \tag{B48}$$

Since $\psi > 1$ and $\epsilon > 1$, we have $\Theta > 0$. Finally, replacing differentials with proportional (hat) changes gives the approximation

$$\ln \hat{\bar{\psi}} \approx \Theta\Big(\ln \hat{L}_2 - \ln \hat{L}_1\Big). \tag{B49}$$

B.7 Decomposing the Sources of Forest Sales

According to equations B6 and B8, the total revenue of farms in the rural area of region i can be expressed as

$$R_i^f = \sum_n \left(\frac{\epsilon}{\epsilon-1}\frac{\chi_i}{\bar{A}_i\bar{\psi}_i}\tau_{ni}^f\right)^{1-\epsilon}(P_n^f)^{\epsilon-1}X_n^f = \tilde{\epsilon}\left(\frac{\chi_i}{\bar{A}_i}\right)^{1-\epsilon}\left(\bar{\psi}_i\right)^{\epsilon-1}M_i^f \tag{B50}$$

where $\tilde{\epsilon} \equiv \left(\frac{\epsilon}{\epsilon-1}\right)^{1-\epsilon}$ is a constant, and $M_i^f \equiv \sum_n \left(\tau_{ni}^f\right)^{1-\epsilon}(P_n^f)^{\epsilon-1}X_n^f$ represents the market access for farmers operating in location i. Substituting the input cost bundle from equation B1, we decompose the impact of highway expansion on forest product sales as follows:

$$\hat{R}_i^f = \underbrace{(\hat{r}_i)^{\upsilon(1-\epsilon)}}_{\text{Land Rent}} \times \underbrace{\left(\hat{P}_i^u\right)^{(1-\upsilon)(1-\epsilon)}}_{\text{Urban Input}} \times \underbrace{\left(\hat{\bar{\psi}}_i\right)^{\epsilon-1}}_{\text{Effective Forest Productivity}} \times \underbrace{\hat{M}_i^f}_{\text{Market Access}}. \tag{B51}$$